Monday, August 29, 2022

1 A Bit of Notation

In this course $\mathbb{R}^N$ is the Euclidean space. Elements of $\mathbb{R}^N$ are points (or vectors) $\boldsymbol{x} = (x_1, \ldots, x_N)$, with $x_n \in \mathbb{R}$, $n = 1, \ldots, N$. The *inner product* of $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^N$ is given by

$$\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{n=1}^{N} x_n y_n$$

and the *Euclidean norm* is

$$\|\boldsymbol{x}\| := \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{x_1^2 + \cdots + x_N^2}.$$

The *open ball* centered at $\boldsymbol{x} \in \mathbb{R}^N$ and radius $r > 0$ is given by $B(\boldsymbol{x}, r) := \{\boldsymbol{y} \in \mathbb{R}^N : \|\boldsymbol{y} - \boldsymbol{x}\| < r\}$ and the *open cube* centered at $\boldsymbol{x}$ and side-length $r > 0$ is given by $Q(\boldsymbol{x}, r) := \{\boldsymbol{y} \in \mathbb{R}^N : |y_n - x_n| < r/2 \text{ for every } n = 1, \ldots, N\}$.

A set $I \subseteq \mathbb{R}$ is an *interval* if for every x, $y \in I$, we have that $tx + (1-t)y \in I$ for every $t \in (0, 1)$. The *length* of I is given by $\operatorname{length} I := \sup I - \inf I$. The empty set $\emptyset$, the real line $\mathbb{R}$ are intervals. Given N bounded intervals $I_1, \ldots, I_N \subset \mathbb{R}$, a *rectangle* in $\mathbb{R}^N$ is a set of the form

$$R := I_1 \times \cdots \times I_N.$$

The set $\mathbb{N}$ of natural numbers starts from 1, while $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

2 Outer Measures

Consider a set X, for example the Euclidean space $\mathbb{R}^N$ or an interval $I \subseteq \mathbb{R}$. We want to measure an arbitrary set $E \subseteq X$.

The idea is to try to approximate E as closely as possible with unions of "nice" sets whose measure we know, for example in $\mathbb{R}^N$ we could use cubes or rectangles or balls, in an interval I we could use intervals (a, b) or $(a, b]$.

So let's take a family $\mathcal{G} \subseteq \mathcal{P}(X)$. An element of $\mathcal{G}$ will be called an *elementary set*. What are the properties that we need on the family $\mathcal{G}$? We want to be able to cover every set of X. This is possible if we can cover X. Thus, let us assume that there exists sequence $\{X_n\}_n$ in $\mathcal{G}$ such that $X = \bigcup_{n=1}^{\infty} X_n$, and let's throw in $\mathcal{G}$ also the empty set.

Then we need a way to measure our elementary sets. So let's consider a function $\rho : \mathcal{G} \to [0, \infty]$ such that $\rho(\emptyset) = 0$. For every set $E \subseteq X$ we try to cover E in the best possible way, that is, we define

$$\mu^*(E) := \inf \left\{ \sum_{n=1}^{\infty} \rho(E_n) : E_n \in \mathcal{G} \text{ for every } n \in \mathbb{N}, E \subseteq \bigcup_{n=1}^{\infty} E_n \right\}. \quad (1)$$

Let's see some important examples. The most important is given by the Lebesgue outer measure.

Example 1 (Lebesgue Outer Measure) *In the Euclidean space* $\mathbb{R}^N$ *we take as family of elementary sets* $\mathcal{G}$ *the family of all rectangles and we define the* elementary measure *of a rectangle* R *as*

$$\operatorname{meas}_N R := \operatorname{length} I_1 \cdot \cdots \cdot \operatorname{length} I_N.$$

For each set $E \subseteq \mathbb{R}^N$ *define*

$$\mathcal{L}_o^N(E) := \inf\left\{\sum_{n=1}^{\infty} \operatorname{meas}_N R_n : \ R_n \textit{ rectangles, } E \subseteq \bigcup_{n=1}^{\infty} R_n\right\}. \tag{2}$$

Another important example is given by *Lebesgue–Stieltjes outer measures.*

Example 2 (Lebesgue–Stieltjes outer measure) *Let* $I \subseteq \mathbb{R}$ *be an open interval and let* $f : I \to \mathbb{R}$ *be an increasing function. Take* $\mathcal{G}$ *to be the family of all intervals* $(a, b]$*, where* $a, b \in I$*, with* $a \leq b$*, and define the elementary measure* $\rho : \mathcal{G} \to [0, \infty)$ *by*

$$\rho((a, b]) := f(b) - f(a).$$

Given $E \subseteq I$ *the* Lebesgue-Stieltjes outer measure *of* E *generated by* f *is given by*

$$\mu_f^*(E) := \inf\left\{\sum_{n=1}^{\infty}(f(b_n) - f(a_n)) : \ a_n, b_n \in I,\ a_n \leq b_n,\ E \subseteq \bigcup_{n=1}^{\infty}(a_n, b_n]\right\}. \tag{3}$$

More generally, we are interested in the case of functions $f : I \to \mathbb{R}^N$ or $f : I \to \mathbb{C}$ or $f : I \to Y$, where Y is a normed space. In this case we can still define $\rho((a, b]) := f(b) - f(a)$ but (3) makes no sense now. We will need to do something else.

Another important example is given by the Hausdorff outer measure in $\mathbb{R}^N$. Loosely speaking the Hausdorff outer measure is a measure that is adapted to measure sets of lower dimensions in $\mathbb{R}^N$, say a curve in the plane or a surface in $\mathbb{R}^3$. It is also used to measure fractals.

Example 3 (The Hausdorff Outer Measure) *Let* $0 \leq s < \infty$*. For* $0 < \delta \leq \infty$ *consider the family of elementary sets*

$$\mathcal{G}_\delta := \left\{F \subset \mathbb{R}^N : \operatorname{diam} F < \delta\right\}$$

and for every $F \in \mathcal{G}_\delta$ *define the elementary measure*

$$\rho_s(F) := \alpha_s \left(\frac{\operatorname{diam} F}{2}\right)^s,$$

where $\alpha_s > 0$ *is a constant. For each set* $E \subseteq \mathbb{R}^N$ *we define*

$$\mathcal{H}_\delta^s(E) := \inf\left\{\sum_{n=1}^{\infty} \alpha_s \left(\frac{\operatorname{diam} E_n}{2}\right)^s : \ E \subseteq \bigcup_{n=1}^{\infty} E_n,\ \operatorname{diam} E_n < \delta\right\}, \tag{4}$$

where, when $s = 0$, we only sum only over those $E_n \neq \emptyset$.

Since for each set $E \subseteq \mathbb{R}^N$ the function $\delta \mapsto \mathcal{H}^s_\delta(E)$ is decreasing, there exists

$$\mathcal{H}^s_o(E) := \lim_{\delta \to 0^+} \mathcal{H}^s_\delta(E) = \sup_{\delta > 0} \mathcal{H}^s_\delta(E). \tag{5}$$

*$\mathcal{H}^s_o$ is called the s-*dimensional Hausdorff outer measure *of E.*

The particular value of the constant α_s is not important and in a lot of books it is taken to be 1. Here, we define

$$\alpha_s := \frac{\pi^{s/2}}{\Gamma\left(\frac{s}{2} + 1\right)},$$

where Γ is the Euler *Gamma function*

$$\Gamma(t) := \int_0^\infty e^{-x} x^{t-1} dx, \quad 0 < t < \infty.$$

Note that $\Gamma(n) = (n-1)!$ for all $n \in \mathbb{N}$. The reason for this choice of constants is that when $N \in \mathbb{N}$, then α_N is the Lebesgue measure of the unit ball in $\mathbb{R}^N$, so that $\mathcal{L}^N(B(x,r)) = \alpha_N r^N$ for every open ball $B(x,r) \subset \mathbb{R}^N$.

Exercise 4 *Prove that in the definition (5) it is possible to restrict the class of admissible sets in the covers $\{E_n\}$ to closed and convex sets (open and convex, respectively), and that the condition $\operatorname{diam} E_n < \delta$ can be replaced by $\operatorname{diam} E_n \leq \delta$, without changing the value of $\mathcal{H}^s_o(E)$.*

What are the properties of the function μ^* defined in (1)? It turns out that μ^* is an outer measure.

Definition 5 *Let X be a nonempty set. A map $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is an* outer measure *if*

(i) $\mu^*(\emptyset) = 0$*;*

(ii) $\mu^*(E) \leq \mu^*(F)$ *for all* $E \subseteq F \subseteq X$*;*

(iii) $\mu^*\left(\bigcup_{n=1}^\infty E_n\right) \leq \sum_{n=1}^\infty \mu^*(E_n)$ *for every countable collection* $\{E_n\} \subseteq \mathcal{P}(X)$ *(countable subadditivity).*

Remark 6 *In several books, outer measures are called measures.*

Wednesday, August 31, 2022

Let's prove that the function μ^* defined in (1) is an outer measure.

Proposition 7 *Let X be a nonempty set and let $\mathcal{G} \subseteq \mathcal{P}(X)$ be such that $\emptyset \in \mathcal{G}$ and there exists $\{X_n\} \subseteq \mathcal{G}$ with $X = \bigcup_{n=1}^\infty X_n$. Let $\rho : \mathcal{G} \to [0, \infty]$ be such that $\rho(\emptyset) = 0$. Then the map $\mu^* : \mathcal{P}(X) \to [0, \infty]$ defined in (1) is an outer measure. Moreover,*

$$\mu^*(E) \leq \rho(E) \tag{6}$$

for every $E \in \mathcal{G}$.

Proof. Since $\emptyset \in \mathcal{G}$ we have that $\mu^*(\emptyset) = 0$. If $E \subseteq F \subseteq X$ then any sequence $\{E_n\}_n$ of elements of $\mathcal{G}$ admissible for F in (1) is also admissible for E, and so $\mu^*(E) \leq \mu^*(F)$. Finally, let $\{F_k\}_k$ be a sequence of subsets of X. Fix $\varepsilon > 0$ and for each k find a sequence $\{E_{n,k}\}_k$ in $\mathcal{G}$ admissible for F_k in (1) and such that

$$\sum_{n=1}^{\infty} \rho(E_{n,k}) \leq \mu^*(F_k) + \frac{\varepsilon}{2^k}.$$

Since $\mathbb{N} \times \mathbb{N}$ is countable, we may write $\{E_{n,k}\}_{k,n\in\mathbb{N}} = \{R_j\}_{j\in\mathbb{N}}$. Note that

$$\bigcup_{k=1}^{\infty} F_k \subseteq \bigcup_{j=1}^{\infty} R_j = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} E_{n,k},$$

and so (see Exercise 8 below)

$$\mu^*\left(\bigcup_{k=1}^{\infty} F_k\right) \leq \sum_{j=1}^{\infty} \rho(R_j) = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \rho(E_{n,k}) \leq \sum_{k=1}^{\infty} \mu^*(F_k) + \varepsilon.$$

By letting $\varepsilon \to 0^+$ we conclude the proof of (iii).

Finally, if $E \in \mathcal{G}$, then taking $E_1 := E$, $E_n := \emptyset$ for all $n \geq 2$, it follows from the definition of μ^* that $\mu^*(E) \leq \rho(E)$. ∎

Exercise 8 *Double series.*

(i) Let $a_{n,k} \geq 0$, for $k, n \in \mathbb{N}$. Prove that

$$\sum_{k=1}^{\infty} \sum_{n=1}^{\infty} a_{n,k} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} a_{n,k}.$$

(ii) Let $a_{nk} \geq 0$, for $k, n \in \mathbb{N}$ and define $c_m := \sum_{n+k=m+1} a_{n,k} = a_{1,m} + \cdots + a_{m,1}$. Prove that

$$\sum_{m=1}^{\infty} c_m = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} a_{n,k}.$$

(iii) Let

$$a_{nk} := \begin{cases} 1 & \text{if } k = n, \\ -1 & \text{if } k = n+1, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that

$$\sum_{k=1}^{\infty} \sum_{n=1}^{\infty} a_{nk} \neq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} a_{nk}.$$

Corollary 9 *The set functions $\mathcal{L}_o^N$, μ_f^*, and $\mathcal{H}_o^s$ defined in (2), (3), (5), respectively, are outer measures.*

Proof. The fact that $\mathcal{L}_o^N$, μ_f^*, and $\mathcal{H}_\delta^s$ are outer measures follow from Proposition 7. It remains to show that $\mathcal{H}_o^s$ is an outer measure. Since $\mathcal{H}_\delta^s(\emptyset) = 0$ for every $\delta > 0$, letting $\delta \to 0^+$ gives $\mathcal{H}_0^s(\emptyset) = 0$.

If $E \subseteq F$, then $\mathcal{H}_\delta^s(E) \leq \mathcal{H}_\delta^s(F)$, and so letting $\delta \to 0^+$ gives $\mathcal{H}_o^s(E) \leq \mathcal{H}_o^s(F)$.

To prove countable subadditivity, let $\{E_n\} \subseteq \mathbb{R}^N$. Since $\mathcal{H}_\delta^s$ is an outer measure, we have that

$$\mathcal{H}_\delta^s\left(\bigcup_{n=1}^\infty E_n\right) \leq \sum_{n=1}^\infty \mathcal{H}_\delta^s(E_n) \leq \sum_{n=1}^\infty \mathcal{H}_o^s(E_n),$$

where in the last inequality we have used (5). Letting $\delta \to 0^+$ and using (5) once more gives the desired inequality. ■

Next we discuss under what conditions the elementary measure ρ coincides with the outer measure μ^* on elementary sets, that is, when we have equality in (6).

Proposition 10 *Let X be a nonempty set and let $\mathcal{G} \subseteq \mathcal{P}(X)$ be such that $\emptyset \in \mathcal{G}$ and there exists $\{X_n\}_n$ in $\mathcal{G}$ with $X = \bigcup_{n=1}^\infty X_n$. Let $\rho : \mathcal{G} \to [0, \infty]$ be such that $\rho(\emptyset) = 0$ and let $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be defined in (1). Then*

$$\mu^*(E) = \rho(E)$$

for every $E \in \mathcal{G}$ if and only if ρ is countably subadditive, that is,

$$\rho(E) \leq \sum_{n=1}^\infty \rho(E_n)$$

for all $E \subseteq \bigcup_{n=1}^\infty E_n$ with $E, E_n \in \mathcal{G}$, $n \in \mathbb{N}$.

Proof. Let $E \in \mathcal{G}$ and assume that

$$\rho(E) \leq \sum_{n=1}^\infty \rho(E_n)$$

for all $E_n \in \mathcal{G}$, $n \in \mathbb{N}$, with $E \subseteq \bigcup_{n=1}^\infty E_n$. Taking the infimum over all such $\{E_n\}_n$ we get $\rho(E) \leq \mu^*(E)$, which, together with (6) implies that $\mu^*(E) = \rho(E)$.

Conversely, if $\mu^*(E) = \rho(E)$ for every for every $E \in \mathcal{G}$, then by properties (ii) and (iii) of an outer measure,

$$\rho(E) = \mu^*(E) \leq \mu^*\left(\bigcup_{n=1}^\infty E_n\right) \leq \sum_{n=1}^\infty \mu^*(E_n) = \sum_{n=1}^\infty \rho(E_n)$$

for all $E_n \in \mathcal{G}$, $n \in \mathbb{N}$, with $E \subseteq \bigcup_{n=1}^\infty E_n$. ■

Exercise 11 *Prove that*

$$\operatorname{meas}_N R \leq \sum_{n=1}^{\infty} \operatorname{meas}_N R_n$$

for every R, R_n rectangles in $\mathbb{R}^N$, with $R \subseteq \bigcup_{n=1}^{\infty} R_n$. Conclude that

$$\mathcal{L}_o^N(R) = \operatorname{meas}_N R$$

for every rectangle in $\mathbb{R}^N$.

In general we have strict inequality in (6) for Lebesgue–Stiljies outer measures.

Theorem 12 *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a monotone function. Then for every $x \in I^\circ$ there exist the left and right limit*

$$f^-(x) := \lim_{y \to x^-} f(y), \quad f^+(x) := \lim_{y \to x^+} f(y).$$

Moreover $f = f^+ = f^-$ for all but countably many x. In particular, f has at most countably many discontinuity points.

Proof. It's in the real analysis notes. ■

Theorem 13 *Let $I \subseteq \mathbb{R}$ be an open interval, let $f : I \to \mathbb{R}$ be increasing. Then for all $a, b \in I$, with $a \leq b$,*

$$\mu_f^*((a, b]) = f(b) - f^+(a) - \sum_{x \in (a,b)} (f^+(x) - f(x)). \tag{7}$$

Moreover,

$$\mu_f^*(\{a\}) = f(a) - f^-(a). \tag{8}$$

Friday, September 2, 2022

3 σ-Algebras and Measures

In the previous section we have given the definition of outer measures and provided a general method for constructing outer measures. The next question is what to do with an outer measure. If we want to measure sets, an important property that is desirable is that if we take two disjoint sets, then the measure of the union should be the sum of the measures.

Unfortunately, in general an outer measure does not have this property. To circumvent this problem Carathéodory proposed to restrict an outer measure $\mu^* : \mathcal{P}(X) \to [0, \infty]$ to a smaller class of subsets for which additivity of disjoint sets holds. The class that we chose is the following:

Definition 14 *Let X be a nonempty set and let $\mu^* : \mathcal{P}(X) \to [0,\infty]$ be an outer measure. A set $E \subseteq X$ is said to be μ^*-*measurable *if*

$$\mu^*(F) = \mu^*(F \cap E) + \mu^*(F \setminus E)$$

for all sets $F \subseteq X$.

Remark 15 *By the subadditivity of μ^* the inequality*

$$\mu^*(F) \leq \mu^*(F \cap E) + \mu^*(F \setminus E)$$

holds for all sets $F \subseteq X$. Hence, to prove that a set $E \subseteq X$ is μ^*-measurable, it suffices to show that

$$\mu^*(F) \geq \mu^*(F \cap E) + \mu^*(F \setminus E) \tag{9}$$

for all sets $F \subseteq X$. Moreover, it is enough to consider sets $F \subseteq X$ such that $\mu^*(F) < \infty$, since otherwise the inequality (9) is automatically satisfied.

We will see below in Theorem 20 that the restriction of μ^* to the class

$$\mathfrak{M}^* := \{E \subseteq X : \ E \text{ is } \mu^*\text{-measurable}\}$$

is additive, actually countably additive and that the class $\mathfrak{M}^*$ has some important properties, precisely it is a σ-*algebra*.

Definition 16 *Let X be a nonempty set. A collection $\mathfrak{M} \subseteq \mathcal{P}(X)$ is an* algebra *if*

(i) $\emptyset \in \mathfrak{M}$;

(ii) if $E \in \mathfrak{M}$ then $X \setminus E \in \mathfrak{M}$;

(iii) if E_1, $E_2 \in \mathfrak{M}$ then $E_1 \cup E_2 \in \mathfrak{M}$.

$\mathfrak{M}$ *is said to be a* σ-algebra *if it satisfies (i)–(ii) and*

(iii)′ if $\{E_n\} \subseteq \mathfrak{M}$ then $\bigcup_{n=1}^{\infty} E_n \in \mathfrak{M}$.

To highlight the dependence of the σ-algebra $\mathfrak{M}$ on X we will sometimes use the notation $\mathfrak{M}(X)$. If $\mathfrak{M}$ is a σ-algebra then the pair $(X, \mathfrak{M})$ is called a *measurable space*. For simplicity we will often apply the term measurable space only to X.

Using De Morgan's laws and (ii) and (iii)′, it follows that a σ-algebra is closed under countable intersection.

Definition 17 *Let X be a nonempty set and let $\mathfrak{M} \subseteq \mathcal{P}(X)$ be an algebra. A map $\mu : \mathfrak{M} \to [0,\infty]$ is called a* (positive) finitely additive measure *if*

$$\mu(\emptyset) = 0, \quad \mu(E_1 \cup E_2) = \mu(E_1) + \mu(E_2)$$

for all E_1, $E_2 \in \mathfrak{M}$ with $E_1 \cap E_2 = \emptyset$.

Definition 18 *Let X be a nonempty set, let $\mathfrak{M} \subseteq \mathcal{P}(X)$ be a σ-algebra. A map $\mu : \mathfrak{M} \to [0, \infty]$ is called a* (positive) measure *if*

$$\mu(\emptyset) = 0, \quad \mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$$

for every countable collection $\{E_n\}_n$ in $\mathfrak{M}$ of pairwise disjoint sets. The triple $(X, \mathfrak{M}, \mu)$ is said to be a measure space.

Definition 19 *Given a measure space $(X, \mathfrak{M}, \mu)$, the measure μ is said to be* complete *if for every $E \in \mathfrak{M}$ with $\mu(E) = 0$ it follows that every $F \subseteq E$ belongs to $\mathfrak{M}$.*

Theorem 20 (Carathéodory) *Let X be a nonempty set and let $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be an outer measure. Then*

$$\mathfrak{M}^* := \{E \subseteq X : \ E \text{ is } \mu^*\text{-measurable}\} \tag{10}$$

is a σ-algebra and $\mu^* : \mathfrak{M}^* \to [0, \infty]$ is a complete measure.

Proof. Step 1: Since $\mu^*(\emptyset) = 0$, for any $F \subseteq X$,

$$\mu^*(F) = \mu^*(F \cap \emptyset) + \mu^*(F \setminus \emptyset),$$

thus $\emptyset \in \mathfrak{M}^*$.
Step 2: To prove that if $E \in \mathfrak{M}^*$, then $X \setminus E \in \mathfrak{M}^*$, it suffices to observe that

$$F \cap (X \setminus E) = F \setminus E, \qquad F \setminus (X \setminus E) = F \cap E.$$

Step 3: We show that if $E_1, E_2 \in \mathfrak{M}^*$, then $E_1 \cup E_2 \in \mathfrak{M}^*$. Fix a set $F \subseteq X$ with $\mu^*(F) < \infty$. Using the fact that E_1, $E_2 \in \mathfrak{M}^*$ we have that

$$\infty > \mu^*(F) = \mu^*(F \cap E_1) + \mu^*(F \setminus E_1),$$
$$\mu^*(F \setminus E_1) = \mu^*((F \setminus E_1) \cap E_2) + \mu^*((F \setminus E_1) \setminus E_2).$$

We now add these two inequalities and cancel $\mu^*(F \setminus E_1) < \infty$ from both sides. We get

$$\begin{aligned} \mu^*(F) &= \mu^*(F \cap E_1) + \mu^*((F \setminus E_1) \cap E_2) + \mu^*((F \setminus E_1) \setminus E_2) \\ &\geq \mu^*((F \cap E_1) \cup (F \setminus E_1) \cap E_2) + \mu^*((F \setminus E_1) \setminus E_2) \\ &= \mu^*(F \cap (E_1 \cup E_2)) + \mu^*(F \setminus (E_1 \cup E_2)), \end{aligned}$$

where in the second inequality we have used the subadditivity of μ^*.

Thus $\mathfrak{M}^*$ is an algebra.
Step 4: To prove that $\mu^* : \mathfrak{M}^* \to [0, \infty]$ is a finitely additive measure, let E_1, $E_2 \in \mathfrak{M}^*$ be disjoint sets and let $F \subseteq X$. Since $E_1 \in \mathfrak{M}^*$ and E_1, E_2 are sets, we obtain

$$\begin{aligned} \mu^*(F \cap (E_1 \cup E_2)) &= \mu^*((F \cap (E_1 \cup E_2)) \cap E_1) + \mu^*((F \cap (E_1 \cup E_2)) \setminus E_1) \\ &= \mu^*(F \cap E_1) + \mu^*(F \cap E_2), \end{aligned}$$

which implies finite additivity (take $F := X$).

Using an induction argument we have that if $E_1, \ldots, E_m \in \mathfrak{M}^*$, $m \in \mathbb{N}$, are pairwise disjoint and $F \subseteq X$, then $\bigcup_{n=1}^m E_n \in \mathfrak{M}^*$ and

$$\mu^* \left(F \cap \bigcup_{n=1}^m E_n \right) = \sum_{n=1}^m \mu^* (F \cap E_n) . \tag{11}$$

■

Remark 21 ***Adam's alternative proof of Step 3:*** *We show that if $E_1, E_2 \in \mathfrak{M}^*$, then $E_1 \cup E_2 \in \mathfrak{M}^*$. Fix a set $F \subseteq X$ with $\mu^*(F) < \infty$. Using the fact that E_1, $E_2 \in \mathfrak{M}^*$ we have that*

$$\infty > \mu^*(F) = \mu^*(F \cap E_1) + \mu^*(F \setminus E_1),$$
$$\mu^*(F \setminus E_1) = \mu^*((F \setminus E_1) \cap E_2) + \mu^*((F \setminus E_1) \setminus E_2)$$

We now add these two inequalities and cancel $\mu^*(F \setminus E_1) < \infty$ from both sides. We get

$$\mu^*(F) = \mu^*(F \cap E_1) + \mu^*((F \setminus E_1) \cap E_2) + \mu^*((F \setminus E_1) \setminus E_2) . \tag{12}$$

Using again the fact that $E_1 \in \mathfrak{M}^*$, we have that

$$\begin{aligned} \mu^*(F \cap (E_1 \cup E_2)) &= \mu^*((F \cap (E_1 \cup E_2)) \cap E_1) + \mu^*((F \cap (E_1 \cup E_2)) \setminus E_1) \\ &= \mu^*(F \cap E_1) + \mu^*((F \setminus E_1) \cap E_2), \end{aligned}$$

where we used the fact that $(F \cap (E_1 \cup E_2)) \cap E_1 = F \cap E_1$ and $(F \cap (E_1 \cup E_2)) \setminus E_1 = (F \setminus E_1) \cap E_2$. Using this identity in (12), we obtain We now add these two inequalities and cancel $\mu^*(F \setminus E_1) < \infty$ from both sides. We get

$$\mu^*(F) = \mu^*(F \cap (E_1 \cup E_2)) + \mu^*(F \setminus (E_1 \cup E_2)) .$$

Monday, September 5, 2022

Labor Day, no classes

Wednesday, September 7, 2022

Proof.

Step 5: We are now ready to prove that $\mu^* : \mathfrak{M}^* \to [0, \infty]$ is a countably additive measure. Let $\{E_n\} \subseteq \mathfrak{M}^*$ be any sequence of pairwise disjoint sets and let $F \subseteq X$. Since $\bigcup_{n=1}^m E_n \in \mathfrak{M}^*$ for any $m \in \mathbb{N}$, we have that

$$\begin{aligned} \mu^*(F) &= \mu^* \left(F \cap \bigcup_{n=1}^m E_n \right) + \mu^* \left(F \setminus \left(\bigcup_{n=1}^m E_n \right) \right) \\ &= \sum_{n=1}^m \mu^*(F \cap E_n) + \mu^* \left(F \setminus \left(\bigcup_{n=1}^m E_n \right) \right) \\ &\geq \sum_{n=1}^m \mu^*(F \cap E_n) + \mu^* \left(F \setminus \left(\bigcup_{n=1}^\infty E_n \right) \right), \end{aligned}$$

by (11) and the subadditivity of μ^*. Letting $m \to \infty$ in the previous inequality yields

$$\mu^* (F) \geq \sum_{n=1}^{\infty} \mu^* (F \cap E_n) + \mu^* \left(F \setminus \left(\bigcup_{n=1}^{\infty} E_n \right) \right). \tag{13}$$

By the properties of outer measures, the right-hand side of the previous inequality is greater than or equal to

$$\mu^* \left(F \cap \left(\bigcup_{n=1}^{\infty} E_n \right) \right) + \mu^* \left(F \setminus \left(\bigcup_{n=1}^{\infty} E_n \right) \right),$$

and so

$$\mu^* (F) \geq \mu^* \left(F \cap \left(\bigcup_{n=1}^{\infty} E_n \right) \right) + \mu^* \left(F \setminus \left(\bigcup_{n=1}^{\infty} E_n \right) \right),$$

which implies that $\bigcup_{n=1}^{\infty} E_n \in \mathfrak{M}^*$. On the other hand, taking $F := \bigcup_{n=1}^{\infty} E_n$ in (13) gives

$$\mu^* \left(\bigcup_{n=1}^{\infty} E_n \right) \geq \sum_{n=1}^{\infty} \mu^* (E_n),$$

and so $\mu^* : \mathfrak{M}^* \to [0, \infty]$ is a countably additive measure.
Step 6: To prove that $\mathfrak{M}^*$ is a σ-algebra, let $\{E_n\} \subseteq \mathfrak{M}^*$. Then the sets

$$F_1 := E_1, \quad F_{n+1} := E_{n+1} \setminus \bigcup_{k=1}^{n} E_k$$

belong to $\mathfrak{M}^*$ and are pairwise disjoint. Hence,

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n \in \mathfrak{M}^*.$$

Step 6: Finally, if $\mu^* (E) = 0$, then by the monotonicity of the outer measure, $\mu^* (F \cap E) = 0$ for all sets $F \subseteq X$. Hence E is μ^*-measurable and $\mu^* : \mathfrak{M}^* \to [0, \infty]$ is a complete measure. ■

Example 22

(i) *The class of all $\mathcal{L}_o^N$-measurable subsets of $\mathbb{R}^N$ is called the* σ-algebra of Lebesgue measurable sets*, and by Carathéodory's theorem, $\mathcal{L}_o^N$ restricted to this σ-algebra is a complete measure, called the N–dimensional* Lebesgue measure *and denoted by $\mathcal{L}^N$. Given a Lebesgue measurable set $E \subseteq \mathbb{R}^N$, we will write indifferently*

$$\mathcal{L}^N (E)$$

for the Lebesgue measure of E.

(ii) *By Carathéodory's theorem, $\mathcal{H}_o^s$ restricted to the σ-algebra of all $\mathcal{H}_o^s$-measurable subsets of $\mathbb{R}^N$ is a complete measure denoted $\mathcal{H}^s$ and called s-*dimensional Hausdorff measure*.*

(iii) *Let $I \subseteq \mathbb{R}$ be an open interval and let $f : I \to \mathbb{R}$ be an increasing function. By Carathéodory's theorem, μ_f^* restricted to the σ-algebra of all μ_f^*-measurable subsets of I is a complete measure denoted μ_f and called the* Lebesgue-Stieltjes measure *generated by f.*

Using Carathéodory's theorem, we have created a large class of complete measures. The next problem is to understand the class $\mathfrak{M}^*$ of the μ^*-measurable sets. For instance, in the case of the Lebesgue measure $\mathcal{L}^N$, it is important to determine if a ball, or a cube, or an open set is Lebesgue measurable.

Let X be a nonempty set. Given any subset $\mathcal{F} \subseteq \mathcal{P}(X)$ the smallest (in the sense of inclusion) σ-algebra that contains $\mathcal{F}$ is given by the intersection of all σ-algebras on X that contain $\mathcal{F}$.

If X is a topological space, then the *Borel σ-algebra $\mathcal{B}(X)$* is the smallest σ-algebra containing all open subsets of X.

Definition 23 *Let X be a metric space and let $\mu^* : \mathcal{P}(X) \to [0, \infty]$ be an outer measure. Then μ^* is said to be a* metric outer measure *if*

$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$$

for all sets E, $F \subseteq X$, with

$$\operatorname{dist}(E, F) := \inf\{d(x, y) : x \in E, y \in F\} > 0.$$

Proposition 24 *The outer measures $\mathcal{H}_o^s$,$0 \leq s < \infty$, μ_f^*, and $\mathcal{L}_o^N$ are metric outer measures.*

Proof. We only prove it for $\mathcal{L}_o^N$. Let $E, F \subset \mathbb{R}^N$ be such that $d := \operatorname{dist}(E, F) > 0$ and consider a sequence $\{R_n\}_n$ of rectangles such that

$$E \cup F \subseteq \bigcup_{n=1}^{\infty} R_n.$$

By partitioning each rectangle R_n into smaller rectangles, if necessary, we can assume that $\operatorname{diam} R_n < \frac{d}{2}$ for every n. Hence, if $R_n \cap E \neq \emptyset$, then necessarily, $R_n \cap F = \emptyset$, while if $R_n \cap F \neq \emptyset$, then necessarily, $R_n \cap E = \emptyset$. Thus, we can divide the sequence $\{R_n\}_n$ into two subsequences, one covering E and one F. It follows that

$$\sum_{n=1}^{\infty} \operatorname{meas}_N R_n = \sum_{R_n \cap E \neq \emptyset} \operatorname{meas}_N R_n + \sum_{R_n \cap F \neq \emptyset} \operatorname{meas}_N R_n \geq \mathcal{L}_o^N(E) + \mathcal{L}_o^N(F).$$

Taking the infimum over all sequences $\{R_n\}_n$ covering $E \cup F$ gives

$$\mathcal{L}_o^N(E \cup F) \geq \mathcal{L}_o^N(E) + \mathcal{L}_o^N(F).$$

The other inequality follows from the fact that $\mathcal{L}_o^N$ is an outer measure. Thus, we have shown that $\mathcal{L}_o^N$ is a metric outer measure. ■

Friday, September 9, 2022

Proposition 25 *Let X be a metric space and let $\mu^* : \mathcal{P}(X) \to [0,\infty]$ be a metric outer measure. Then every Borel set is μ^*-measurable.*

Proof. Since closed sets generate the Borel σ-algebra $\mathcal{B}(X)$, to show that $\mathfrak{M}^*$ contains $\mathcal{B}(X)$, it is enough to prove that $\mathfrak{M}^*$ contains all closed sets. Thus let $C \subseteq X$ be a closed set and let $F \subseteq X$ be such that $\mu^*(F) < \infty$. For $n \in \mathbb{N}$ define

$$E_0 := \{x \in F \setminus C : \operatorname{dist}(x, C) \geq 1\}$$
$$E_n := \left\{x \in F \setminus C : \frac{1}{n+1} \leq \operatorname{dist}(x, C) < \frac{1}{n}\right\}.$$

Note that the sets E_n are disjoint. Moreover, since C is closed we have that

$$\bigcup_{n=0}^{\infty} E_n = F \setminus C.$$

Indeed, if $x \in F \setminus C$, then $\operatorname{dist}(x, C) > 0$, and so we may find $n \in \mathbb{N}$ such that $x \in E_n$.

If $x \in E_{2k}$ and $y \in E_n$, where $n \geq 2k+2$, then

$$\frac{1}{2k+1} \leq \operatorname{dist}(x, C) \leq d(x, y) + \operatorname{dist}(y, C) < d(x, y) + \frac{1}{n}$$

and so $d(x, y) \geq \frac{1}{2k+1} - \frac{1}{n} > 0$, which implies that $\operatorname{dist}(E_{2k}, E_n) > 0$ for all $k \geq 0$ and all $n \geq 2k+2$. By the fact that μ^* is a metric outer measure, for all $k \in \mathbb{N}$,

$$\sum_{j=0}^{k} \mu^*(E_{2j}) = \mu^*\left(\bigcup_{j=0}^{k} E_{2j}\right) \leq \mu^*(F) < \infty$$

Similarly

$$\sum_{j=1}^{k} \mu^*(E_{2j-1}) \leq \mu^*(F) < \infty.$$

Thus the series $\sum_{j=0}^{\infty} \mu^*(E_{2j})$ and $\sum_{j=1}^{\infty} \mu^*(E_{2j-1})$ are convergent. In turn, the series $\sum_{n=0}^{\infty} \mu^*(E_n)$ is convergent.

Next observe that the sets $F \cap C$ and $\bigcup_{j=0}^{n} E_j$ have positive distance, since if $x \in F \cap C$ and $y \in \bigcup_{j=0}^{n} E_j$, then

$$\frac{1}{n+1} \leq \operatorname{dist}(y, C) \leq d(x, y).$$

Hence, using again the fact that μ^* is a metric outer measure, we have that

$$\begin{aligned}\mu^*(F\cap C)+\mu^*(F\setminus C) &= \mu^*(F\cap C)+\mu^*\left(\bigcup_{j=0}^{\infty}E_j\right)\\ &= \mu^*(F\cap C)+\mu^*\left(\bigcup_{j=0}^{n-1}E_j\cup\bigcup_{j=n}^{\infty}E_j\right)\\ &\le \mu^*(F\cap C)+\mu^*\left(\bigcup_{j=0}^{n-1}E_j\right)+\mu^*\left(\bigcup_{j=n}^{\infty}E_j\right)\\ &\le \mu^*(F\cap C)+\mu^*\left(\bigcup_{j=0}^{n-1}E_j\right)+\sum_{j=n}^{\infty}\mu^*(E_j)\\ &= \mu^*\left((F\cap C)\cup\bigcup_{j=0}^{n-1}E_j\right)+\sum_{j=n}^{\infty}\mu^*(E_j)\\ &\le \mu^*(F)+\sum_{j=n}^{\infty}\mu^*(E_j).\end{aligned}$$

Letting $n\to\infty$, we conclude that $\mu^*(F\cap C)+\mu^*(F\setminus C)\le\mu^*(F)$ and the proof is complete. ■

It follows from the previous two propositions that open sets, closed sets, and Borel sets are Lebesgue measurable and $\mathcal{H}^s_o$-measurable, and, when $N=1$, μ^*_f-measurable.

Proposition 26 *Let $(X,\mathfrak{M},\mu)$ be a measure space.*

(i) If $\{E_n\}_n$ is an increasing sequence of subsets of $\mathfrak{M}$ then

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\lim_{n\to\infty}\mu(E_n).$$

(ii) If $\{E_n\}_n$ is a decreasing sequence of subsets of $\mathfrak{M}$ and $\mu(E_1)<\infty$ then

$$\mu\left(\bigcap_{n=1}^{\infty}E_n\right)=\lim_{n\to\infty}\mu(E_n).$$

Example 27 *Without the hypothesis $\mu(E_1)<\infty$, property (ii) may be false. Indeed, let $E_n:=[n,\infty)$. Then $\{E_n\}_n$ is a decreasing sequence, $\mathcal{L}^1(E_n)=\infty$ for all $n\in\mathbb{N}$, but*

$$\mathcal{L}^1\left(\bigcap_{n=1}^{\infty}E_n\right)=\mathcal{L}^1(\emptyset)=0\neq\lim_{n\to\infty}\mathcal{L}^1(E_n)=\infty.$$

Monday, September 12, 2022

Let's prove the proposition.

Proof. (i) Define

$$F_n := E_n \setminus E_{n-1},$$

where $E_0 := \emptyset$. Since $\{E_n\}_n$ is an increasing sequence, it follows that the sets F_n are pairwise disjoint with $\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n$, and so by the properties of measures we have

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} E_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mu(F_n) = \lim_{l\to\infty} \sum_{n=1}^{l} \mu(F_n) \\ &= \lim_{l\to\infty} \mu\left(\bigcup_{n=1}^{l} F_n\right) = \lim_{l\to\infty} \mu(E_l).\end{aligned}$$

(ii) Apply part (i) to the increasing sequence $\{E_1 \setminus E_n\}_n$ to get

$$\begin{aligned}\lim_{n\to\infty} (\mu(E_1) - \mu(E_n)) &= \lim_{n\to\infty} \mu(E_1 \setminus E_n) = \mu\left(\bigcup_{n=1}^{\infty} (E_1 \setminus E_n)\right) \\ &= \mu\left(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\right) = \mu(E_1) - \mu\left(\bigcap_{n=1}^{\infty} E_n\right).\end{aligned}$$

Since $\mu(E_1) < \infty$, we get

$$\lim_{n\to\infty} \mu(E_n) = \mu\left(\bigcap_{n=1}^{\infty} E_n\right).$$

■

4 Lebesgue integration of nonnegative functions

We are now in a position to introduce the notion of integral. Let $(X, \mathfrak{M}, \mu)$ be a measure space. Given a set $E \subseteq X$ the *characteristic function of* E is the function χ_E, defined by

$$\chi_E(x) := \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Let $E, F \subseteq X$ belong to the σ-algebra $\mathfrak{M}$. We define the *integral* of χ_E over F as

$$\int_F \chi_E \, d\mu := \mu(F \cap E).$$

Definition 28 *Let $(X, \mathfrak{M})$ be a measurable space and $E \in \mathfrak{M}$. A* simple function *is a function $s : E \to \mathbb{R}$ that can be written as*

$$s = \sum_{n=1}^{\ell} c_n \chi_{E_n},$$

where $c_1, \ldots, c_\ell \in \mathbb{R}$ and the sets E_n are measurable.

Let $(X, \mathfrak{M}, \mu)$ be a measure space, $E \in \mathfrak{M}$, and $s : E \to [0, \infty)$ be a nonnegative simple function. If $s \neq 0$, we can write

$$s = \sum_{n=1}^{\ell} c_n \chi_{E_n},$$

where the sets $E_n \subseteq E$ are measurable, pairwise disjoint, and $c_n > 0$ for all $n = 1, \ldots, \ell$. Given $F \in \mathfrak{M}$ with $F \subseteq E$, we define the *Lebesgue integral* of s over F as

$$\int_F s \, d\mu := \sum_{n=1}^{\ell} c_n \mu(F \cap E_n). \tag{14}$$

We leave as an exercise to show that the integral does not depend on the particular representation of s, that is, that if

$$s = \sum_{k=1}^{m} d_k \chi_{F_k},$$

where the sets $F_k \subseteq E$ are measurable, pairwise disjoint, and $d_k > 0$, then

$$\int_F s \, d\mu = \sum_{k=1}^{m} d_k \mu(F \cap F_k).$$

We set $\int_F 0 \, d\mu := 0$.

Proposition 29 *Let $(X, \mathfrak{M}, \mu)$ be a measure space, $E, F \in \mathfrak{M}$, with $F \subseteq E$, and s, $t : E \to [0, \infty)$ be simple functions. Then*

$$\int_F (s + t) \, d\mu = \int_F s \, d\mu + \int_F t \, d\mu$$

and

$$\int_F cs \, d\mu = c \int_F s \, d\mu$$

for every $c \geq 0$, where we set $0 \cdot \infty := 0$.

Proof. Exercise. ■

Proposition 30 *Let $(X, \mathfrak{M}, \mu)$ be a measure space and $s : X \to [0, \infty)$ be a simple function. Then the set function*

$$\nu(E) := \int_E s \, d\mu, \quad E \in \mathfrak{M},$$

is a measure.

Proof. If $s = 0$, then $\nu = 0$ and there is nothing to prove. Assume $s \neq 0$ and let $s = \sum_{n=1}^{\ell} c_n \chi_{E_n}$, where the sets E_n are measurable, pairwise disjoint, and $c_n > 0$ for all $n = 1, \ldots, \ell$. Let $\{F_k\}_k$ be a sequence of measurable, pairwise disjoint sets. Then $\nu(\emptyset) = \int_\emptyset s \, d\mu = 0$ and

$$\begin{aligned}
\nu\left(\bigcup_{k=1}^{\infty} F_k\right) &= \int_{\bigcup_{k=1}^{\infty} F_k} s \, d\mu = \sum_{n=1}^{\ell} c_n \mu\left(E_n \cap \bigcup_{k=1}^{\infty} F_k\right) = \sum_{n=1}^{\ell} c_n \mu\left(\bigcup_{k=1}^{\infty} (F_k \cap E_n)\right) \\
&= \sum_{n=1}^{\ell} c_n \sum_{k=1}^{\infty} \mu(F_k \cap E_n) = \sum_{k=1}^{\infty} \sum_{n=1}^{\ell} c_n \mu(F_k \cap E_n) \\
&= \sum_{k=1}^{\infty} \int_{F_k} s \, d\mu = \sum_{k=1}^{\infty} \nu(F_k),
\end{aligned}$$

where we used the fact that μ is a measure, the sets F_k are pairwise disjoint, and Exercise 8. ■

Definition 31 *Let $(X, \mathfrak{M})$ be a measurable space, $E \in \mathfrak{M}$, and $f : E \to [0, \infty]$. We say that f is* Lebesgue measurable *if there exists a sequence of simple functions $s_n : E \to [0, \infty)$ such that $s_n \leq f$ for every n and $s_n \to f$ pointwise in E.*

Proposition 32 *Let $(X, \mathfrak{M})$ be a measurable space, $E \in \mathfrak{M}$, and $f : E \to [0, \infty]$ and $g : E \to [0, \infty]$ be two measurable functions. Then $f + g$, fg, $\min\{f, g\}$, $\max\{f, g\}$ are measurable.*

Proof. Exercise. ■

Proposition 33 *Let $(X, \mathfrak{M})$ be a measurable space, $E \in \mathfrak{M}$, and $f_n : E \to [0, \infty]$, $n \in \mathbb{N}$, be measurable functions. Then $\sup_n f_n$, $\inf_n f_n$, $\liminf_{n\to\infty} f_n$, and $\limsup_{n\to\infty} f_n$ are measurable.*

Proof. Exercise. ■

Wednesday, September 14, 2022

Remark 34 *Let $(X, \mathfrak{M})$ be a measurable space, $E \in \mathfrak{M}$, and $f : E \to [0, \infty]$ be* Lebesgue measurable. *Then there exists a sequence of simple functions $s_n : E \to [0, \infty)$ such that $s_n \leq f$ for every n and $s_n \to f$ pointwise in E. By taking $t_n = \max\{s_1, \ldots, s_n\}$, we have that t_n is still a simple function, $0 \leq t_n \leq f$, and $t_n \to f$ pointwise in E by the squeeze theorem. Thus, in what follows, without loss of generality, we can assume that the sequence $\{s_n\}_n$ of simple functions approximating a measurable function has the property that $s_n \leq s_{n+1}$ in E.*

Definition 35 *Let $(X, \mathfrak{M}, \mu)$ be a measure space. Given $E, F \in \mathfrak{M}$ with $F \subseteq E$ and a measurable function $f : E \to [0, \infty]$, the* (Lebesgue) integral *of f over F is*

$$\int_F f \, d\mu := \sup \left\{ \int_F s \, d\mu : \, s \text{ simple, } 0 \leq s \leq f \text{ in } F \right\}.$$

We list below some basic properties of Lebesgue integration for nonnegative functions.

Proposition 36 *Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $E, F \in \mathfrak{M}$ with $F \subseteq E$ and let f, $g : E \to [0, \infty]$ be two measurable functions.*

(i) If $f \leq g$, then $\int_F f \, d\mu \leq \int_F g \, d\mu$.

(ii) If $c \in [0, \infty)$, then $\int_F cf \, d\mu = c \int_F f \, d\mu$ (here we set $0\infty := 0$).

(iii) $\int_F f \, d\mu = 0$ if and only if $f(x) = 0$ for μ a.e. $x \in F$ (even if $\mu(F) = \infty$).

(iv) If $\mu(F) = 0$, then $\int_F f \, d\mu = 0$, (even if $f \equiv \infty$ in E).

(v) If $\int_F f \, d\mu < \infty$ then $f(x) < \infty$ for μ a.e. $x \in F$.

(vi) $\int_F f \, d\mu = \int_E \chi_F f \, d\mu$.

Proof. (i) If s is a simple function with $0 \leq s \leq f$ in F, then $s \leq f \leq g$ in F and so

$$\int_F s \, d\mu \leq \sup \left\{ \int_F t \, d\mu : \, t \text{ simple, } 0 \leq t \leq g \text{ in } F \right\} = \int_F g \, d\mu.$$

Taking the supremum over all such s gives $\int_F f \, d\mu \leq \int_F g \, d\mu$.

(ii) Assume that $c > 0$. If s is a simple function with $\sum_{n=1}^{\ell} c_n \chi_{E_n}$ and $0 \leq s \leq f$ in F, then cs is also a simple function and

$$c \int_F s \, d\mu = c \sum_{n=1}^{\ell} c_n \mu(E_n \cap F) = \sum_{n=1}^{\ell} cc_n \mu(E_n \cap F) = \int_F cs \, d\mu.$$

Since $0 \leq cs \leq cf$ we get

$$\int_F cs \, d\mu \leq \int cf \, d\mu.$$

But

$$c \int_F s \, d\mu = \int_F cs \, d\mu \leq \int_F cf \, d\mu$$

or equivalently $\int_F s \, d\mu \leq \frac{1}{c} \int_F cf \, d\mu$. Taking the supremum over all such s gives $\int_F f \, d\mu \leq \frac{1}{c} \int_F cf \, d\mu$, that is, $c \int_F f \, d\mu \leq \int_F cf \, d\mu$. Since what we proved holds for every f and c, to obtain the converse inequality it suffices to apply what we just proved to the function $h = cf$, and with c replaced by $\frac{1}{c}$, namely $\frac{1}{c} \int_F h \, d\mu \leq \int_F \frac{1}{c} h \, d\mu$. Then $\frac{1}{c} \int_F cf \, d\mu = \frac{1}{c} \int_F h \, d\mu \leq \int_F \frac{1}{c} h \, d\mu = \int_F \frac{1}{c} cf \, d\mu$ which gives $\int_F cf \, d\mu \leq c \int_F f \, d\mu$. The case $c = 0$ is left as an exercise.

(iii) Assume that $\int_F f\,d\mu = 0$. For $n \in \mathbb{N}$ define

$$F_n := \left\{x \in F :\ f(x) \geq \frac{1}{n}\right\}.$$

Then $f \geq \frac{1}{n}\chi_{F_n}$, and so

$$0 = \int_F f\,d\mu \geq \int_F \frac{1}{n}\chi_{F_n}\,d\mu = \frac{1}{n}\mu(F_n).$$

Since

$$F_+ := \{x \in F :\ f(x) > 0\} = \bigcup_{n=1}^{\infty} F_n,$$

it follows that $\mu(F_+) = 0$. Thus $f(x) = 0$ for μ a.e. $x \in F$.

Conversely, assume that there exists a set $F_0 \in \mathfrak{M}$ with $\mu(F_0) = 0$ such that $f(x) = 0$ for all $x \in F \setminus F_0$. Given a simple function $0 \leq s \leq f$, we have that $s = 0$ in $F \setminus F_0$ and so if $s = \sum_{n=1}^{\ell} c_n\chi_{E_n}$, then we can assume that $c_1 = 0$ and $F \setminus F_0 \subseteq E_1$ and so

$$\int_F s\,d\mu = \sum_{n=1}^{\ell} c_n\mu(E_n \cap F) = 0 + \sum_{n=2}^{\ell} c_n\mu(E_n \cap F_0) = 0.$$

Since this is true for all simple functions s below f, we get $\int_F f\,d\mu = 0$.

(iv) Given a simple function $0 \leq s \leq f$, if $s = \sum_{n=1}^{\ell} c_n\chi_{E_n}$, then

$$\int_F s\,d\mu = \sum_{n=1}^{\ell} c_n\mu(E_n \cap F) = 0.$$

Since this is true for all simple functions s below f, we get $\int_F f\,d\mu = 0$.

(v) Take $s = n\chi_{E_\infty}$, where $E_\infty := \{x \in F :\ f(x) = \infty\}$. Then s is simple and $0 \leq s \leq f$. It follows that

$$n\mu(E_\infty) = \int_F s\,d\mu \leq \int_F f\,d\mu < \infty.$$

Letting $n \to \infty$ we have that $\mu(E_\infty) = 0$.

(vi) Note that if $s = \sum_{n=1}^{\ell} c_n\chi_{E_n}$, then $s\chi_F$ is a simple function with $s\chi_F = \sum_{n=1}^{\ell} c_n\chi_{E_n \cap F}$ and so

$$\int_F s\,d\mu = \sum_{n=1}^{\ell} c_n\mu(E_n \cap F) = \int_E s\chi_F\,d\mu.$$

If s is a simple function with $0 \leq s \leq f$ in F, then $s\chi_F$ is a simple function with $s\chi_F \leq f\chi_F$ in E. Hence,

$$\int_F s\,d\mu = \int_E s\chi_F\,d\mu \leq \int_E f\chi_F\,d\mu.$$

Taking the supremum over all such s gives $\int_F f\,d\mu \leq \int_E f\chi_F\,d\mu$. Conversely, if s is a simple function with $0 \leq s \leq f\chi_F$ in E, then $s = 0$ outside F and so $s\chi_F = s$ and $s \leq f$ in F and so

$$\int_E s\,d\mu = \int_E \chi_F s\,d\mu = \int_F s\,d\mu \leq \int_F f\,d\mu.$$

Taking the supremum over all such s gives $\int_E f\chi_F\,d\mu \leq \int_F f\,d\mu$. ■

Remark 37 *Note that the only place where we used that $\mathfrak{M}$ is a σ-algebra is in property (iii).*

The next two results are central in the theory of integration of nonnegative functions.

Theorem 38 (Lebesgue monotone convergence theorem) *Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $E \in \mathfrak{M}$ and let $f_n : E \to [0, \infty]$ be a sequence of measurable functions such that*

$$0 \leq f_1(x) \leq f_2(x) \leq \ldots \leq f_n(x) \to f(x)$$

for every $x \in E$. Then $f : E \to [0, \infty]$ is measurable and

$$\lim_{n\to\infty} \int_E f_n\,d\mu = \int_E f\,d\mu.$$

Friday, September 16, 2022

Proof. By Proposition 33 the function f is measurable, and since $f_n \leq f_{n+1} \leq f$ we have that $\int_E f_n\,d\mu \leq \int_E f_{n+1}\,d\mu \leq \int_E f\,d\mu$. In particular there exists

$$\lim_{n\to\infty} \int_E f_n\,d\mu =: \ell \in [0, \infty]$$

and $\ell \leq \int_E f\,d\mu$. To prove the opposite inequality, let s be a simple function, with $0 \leq s \leq f$ in E. Fix $0 < c < 1$ and for $n \in \mathbb{N}$ define

$$E_n := \{x \in E : \ f_n(x) \geq cs(x)\}.$$

Since f_n and s are measurable and $f_n \leq f_{n+1}$, it follows that E_n is measurable and $E_n \subseteq E_{n+1}$. We claim that

$$E = \bigcup_{n=1}^{\infty} E_n.$$

To see this, fix $x \in E$. If $f(x) = 0$, then $f_n(x) = 0$ for all $n \in \mathbb{N}$ and $s(x) = 0$, and so $x \in E_n$ for all $n \in \mathbb{N}$. If $f(x) > 0$, then $f(x) > cs(x)$ and since $f_n(x) \to f(x)$, we may find $n \in \mathbb{N}$ so large that $f_n(x) > cs(x)$. Thus $x \in E_n$ and the claim is proved.

Using the fact that $f_n \geq 0$ and the definition of E_n and we have that

$$\int_E f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} cs \, d\mu = c \int_{E_n} s \, d\mu.$$

By Exercise 30, the set function

$$\nu(F) := \int_F s \, d\mu, \quad F \in \mathfrak{M},$$

is a measure, and so by Proposition 26,

$$\int_E f_n \, d\mu \geq c \int_{E_n} s \, d\mu = c\nu(E_n) \to c\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = c\nu(E).$$

Thus

$$\ell \geq c\nu(E) = c \int_E s \, d\mu.$$

Letting $c \nearrow 1$ we obtain that

$$\ell \geq \int_E s \, d\mu,$$

and given the arbitrariness of the simple function s below f, taking the supremum over all such admissible functions s yields

$$\ell \geq \int_E f \, d\mu.$$

This concludes the proof. ■

Remark 39 *The previous theorem continues to hold if we assume that $f_n(x) \to f(x)$ for μ a.e. $x \in E$. Indeed, in view of Proposition 36(iv), it suffices to redefine f_n and f to be zero in the set of measure zero in which there is no pointwise convergence.*

Example 40 *The Lebesgue monotone convergence theorem does not hold in general for decreasing sequences. Indeed, consider $X = \mathbb{R}$ and let μ be the Lebesgue measure $\mathcal{L}^1$. Define*

$$f_n := \frac{1}{n} \chi_{[n,\infty)}.$$

Then $f_n \geq f_{n+1}$ and

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n \, dx = \infty \neq 0 = \int_{\mathbb{R}} \lim_{n\to\infty} f_n \, dx.$$

Corollary 41 *Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $f, g : X \to [0,\infty]$ be two measurable functions. Then*

$$\int_X (f+g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu.$$

Proof. By Remark 34 there exist two sequences $\{s_n\}$ and $\{t_n\}$of simple functions such that

$$0 \leq s_1(x) \leq s_2(x) \leq \ldots \leq s_n(x) \to f(x),$$
$$0 \leq t_1(x) \leq t_2(x) \leq \ldots \leq t_n(x) \to g(x)$$

for every $x \in X$. By Proposition 29,

$$\int_X (s_n + t_n)\, d\mu = \int_X s_n\, d\mu + \int_X t_n\, d\mu.$$

The conclusion follows from Lebesgue's monotone convergence theorem. ■

Corollary 42 *Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $f_n : X \to [0, \infty]$ be a sequence of measurable functions. Then*

$$\sum_{n=1}^{\infty} \int_X f_n\, d\mu = \int_X \sum_{n=1}^{\infty} f_n\, d\mu.$$

Proof. Apply the Lebesgue monotone convergence theorem to the increasing sequence of partial sums and use linearity of the integral. ■

Example 43 *Given a doubly indexed sequence $\{a_{n,k}\}$, with $a_{n,k} \geq 0$ for all n, $k \in \mathbb{N}$, we have*

$$\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} a_{n,k} = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} a_{n,k}.$$

To see this, it suffices to consider $X = \mathbb{N}$ with counting measure and to define $f_n : \mathbb{N} \to [0, \infty]$ by $f_n(k) := a_{n,k}$. Then

$$\int_X f_n\, d\mu = \sum_{k=1}^{\infty} a_{n,k},$$

and the result now follows from the previous corollary.

Monday, September 19, 2022

Lemma 44 (Fatou lemma) *Let $(X, \mathfrak{M}, \mu)$ be a measure space. If $f_n : X \to [0, \infty]$ is a sequence of measurable functions, then*

$$f := \liminf_{n \to \infty} f_n$$

is a measurable function and

$$\int_X f\, d\mu \leq \liminf_{n \to \infty} \int_X f_n\, d\mu$$

Proof. For $n \in \mathbb{N}$ define

$$g_n := \inf_{k \geq n} f_k.$$

Then $g_n \leq f_n$, and so

$$\int_X g_n \, d\mu \leq \int_X f_n \, d\mu.$$

Since $g_n \leq g_{n+1}$, by Lebesgue's monotone convergence theorem

$$\begin{aligned} \int_X \liminf_{n\to\infty} f_n \, d\mu &= \int_X \lim_{n\to\infty} g_n \, d\mu = \lim_{n\to\infty} \int_X g_n \, d\mu \\ &\leq \liminf_{n\to\infty} \int_X f_n \, d\mu. \end{aligned}$$

■

Example 45 *Fatou's lemma fails for real-valued functions. Indeed, consider $X = \mathbb{R}$ and let μ be the Lebesgue measure $\mathcal{L}^1$. Define*

$$f_n := -\frac{1}{n}\chi_{[0,n]}.$$

Then

$$\liminf_{n\to\infty} \int_{\mathbb{R}} f_n \, dx = -1 < 0 = \int_{\mathbb{R}} \lim_{n\to\infty} f_n \, dx.$$

5 Lebesgue Integration of Functions of Arbitrary Sign

Definition 46 *Let $(X, \mathfrak{M})$ be a measurable space, $E \in \mathfrak{M}$, and $f : E \to [-\infty, \infty]$. We say that f is* Lebesgue measurable *if f^+ and f^- are Lebesgue measurable, where*

$$f^+ := \max\{f, 0\}, \quad f^- := \max\{-f, 0\}.$$

Note that $f = f^+ - f^-$, $|f| = f^+ + f^-$.

Definition 47 *Let $(X, \mathfrak{M}, \mu)$ be a measure space. Given $E, F \in \mathfrak{M}$ with $F \subseteq E$ and a measurable function $f : E \to [-\infty, \infty]$, if at least one of the two integrals $\int_F f^+ \, d\mu$ and $\int_F f^- \, d\mu$ is finite, then we define the* Lebesgue integral of f *over F to be*

$$\int_F f \, d\mu := \int_F f^+ \, d\mu - \int_F f^- \, d\mu.$$

If both $\int_F f^+ \, d\mu$ and $\int_E f^- \, d\mu$ are both finite, then f is said to be Lebesgue integrable *over F.*

Proposition 48 *Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $E, F \in \mathfrak{M}$ with $F \subseteq E$ and let $f,\, g : E \to [-\infty, \infty]$ be two measurable functions.*

(i) *If f and g are integrable over F and α, $\beta \in \mathbb{R}$, then $\alpha f + \beta g$ is integrable and*
$$\int_F (\alpha f + \beta g)\, d\mu = \alpha \int_F f\, d\mu + \beta \int_F g\, d\mu.$$

(ii) $\left|\int_F f\, d\mu\right| \leq \int_F |f|\, d\mu$.

(iii) *If f is Lebesgue integrable, then the set $\{x \in F : |f(x)| = \infty\}$ has measure zero.*

(iv) *If $f(x) = g(x)$ for μ a.e. $x \in F$, then $\int_F f^{\pm}\, d\mu = \int_F g^{\pm}\, d\mu$, so that $\int_F f\, d\mu$ is well-defined if and only if $\int_F g\, d\mu$ is well-defined, and in this case we have*
$$\int_F f\, d\mu = \int_F g\, d\mu. \tag{15}$$

Exercise 49 *Prove the previous proposition.*

Property (15) shows that the Lebesgue integral does not distinguish functions that coincide μ a.e. in F. This motivates the next definition.

Definition 50 *Let $(X, \mathfrak{M}, \mu)$ be a measure space. Given $E, E_0, F \in \mathfrak{M}$, with $E_0, F \subseteq E$ and $\mu(E_0) = 0$, and $f : E \setminus E_0 \to [-\infty, \infty]$ a measurable function, then we define the* (Lebesgue) integral of f *over the measurable set F as the Lebesgue integral of the function*
$$g(x) := \begin{cases} f(x) & \text{if } x \in E \setminus E_0, \\ 0 & \text{otherwise,} \end{cases}$$
provided $\int_F g\, d\mu$ is well-defined. Note that in this case
$$\int_F g\, d\mu = \int_F H\, d\mu,$$
where
$$H(x) := \begin{cases} f(x) & \text{if } x \in E \setminus E_0, \\ w(x) & \text{otherwise,} \end{cases}$$
and w is an arbitrary measurable function defined on E_0. If the measure μ is complete, then $\int_E g\, d\mu$ is well-defined if and only if $\int_{E \setminus F} f\, d\mu$ is well-defined.

For functions of arbitrary sign we have the following convergence result.

Theorem 51 (Lebesgue dominated convergence theorem) *Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $E \in \mathfrak{M}$ and let $f_n : E \to \mathbb{R}$ be a sequence of measurable functions such that*
$$\lim_{n \to \infty} f_n(x) = f(x)$$
for all $x \in E$. If there exists a Lebesgue integrable function $g : E \to [0, \infty]$ such that
$$|f_n(x)| \leq g(x)$$

for all $x \in E$ and all $n \in \mathbb{N}$, then f is Lebesgue integrable

$$\lim_{n\to\infty} \int_E f_n \, d\mu = \int_E f \, d\mu.$$

Wednesday, September 21, 2022

Proof. The function f is measurable by Proposition 33, and by Fatou's lemma

$$\int_E |f| \, d\mu \le \liminf_{n\to\infty} \int_E |f_n| \, d\mu \le \int_E |g| \, d\mu < \infty.$$

Thus f is integrable. Since $g \pm f_n \ge 0$, again by Fatou's lemma we have

$$\begin{aligned} \int_E g \, d\mu \pm \int_E f \, d\mu &= \int_E (g \pm f) \, d\mu \le \liminf_{n\to\infty} \int_E (g \pm f_n) \, d\mu \\ &= \int_E g \, d\mu + \liminf_{n\to\infty} \int_E (\pm f_n) \, d\mu. \end{aligned}$$

Using the fact that $\int_E g \, d\mu \in \mathbb{R}$, we can rewrite the previous two inequalities as

$$\int_E f \, d\mu \le \liminf_{n\to\infty} \int_E f_n \, d\mu \le \limsup_{n\to\infty} \int_E f_n \, d\mu \le \int_E f \, d\mu,$$

and so the theorem holds. ■

Example 52 *If g is not integrable then the theorem fails in general. Indeed, consider $E = [0, 1]$ and let μ be the Lebesgue measure $\mathcal{L}^1$. Define*

$$f_n := n\chi_{[0, \frac{1}{n}]}.$$

Then

$$\lim_{n\to\infty} \int_0^1 f_n \, dx = 1 \ne 0 = \int_0^1 \lim_{n\to\infty} f_n \, dx.$$

Exercise 53 *Use Lebesgue dominated convergence theorem to calculate the limit*

$$\lim_{n\to\infty} \int_0^n \left(1 - \frac{x}{n}\right)^n x^{a-1} \, dx,$$

where $a > 0$.

Corollary 54 *Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $E \in \mathfrak{M}$ and let $f_n : E \to [-\infty, \infty]$ be a sequence of measurable functions. If*

$$\sum_{n=1}^{\infty} \int_E |f_n| \, d\mu < \infty,$$

then there exists a set $E_0 \in \mathfrak{M}$ with $\mu(E_0) = 0$ such that the series $\sum_{n=1}^{\infty} f_n(x)$ converges for all $x \in E \setminus E_0$, the function

$$f(x) := \sum_{n=1}^{\infty} f_n(x),$$

defined for $x \in E \setminus E_0$, is integrable, and

$$\sum_{n=1}^{\infty} \int_E f_n \, d\mu = \int_{E \setminus E_0} f \, d\mu.$$

Proof. Define

$$g(x) := \sum_{n=1}^{\infty} |f_n(x)|, \quad x \in E.$$

Then $g : E \to [0, \infty]$. By Corollary 42,

$$\int_E g \, d\mu = \sum_{n=1}^{\infty} \int_E |f_n| \, d\mu < \infty$$

and so g is Lebesgue integral. In particular, there exists a set $E_0 \in \mathfrak{M}$ with $\mu(E_0) = 0$ such that $g(x) < \infty$ for all $x \in E \setminus E_0$. If $g(x) := \sum_{n=1}^{\infty} |f_n(x)| < \infty$, then $\sum_{n=1}^{\infty} f_n(x)$ converges. Thus, there exists

$$\lim_{n \to \infty} \sum_{k=1}^{n} f_k(x) = \sum_{k=1}^{\infty} f_k(x) \in \mathbb{R}$$

for all $x \in E \setminus E_0$. Moreover,

$$\left| \sum_{k=1}^{n} f_k(x) \right| \leq \sum_{k=1}^{n} |f_k(x)| \leq \sum_{k=1}^{\infty} |f_k(x)| = g(x).$$

Hence, we can apply the Lebesgue dominated convergence theorem in $E \setminus E_0$ to obtain

$$\lim_{n \to \infty} \int_{E \setminus E_0} \sum_{k=1}^{n} f_k \, d\mu = \int_{E \setminus E_0} \lim_{n \to \infty} \sum_{k=1}^{n} f_k \, d\mu.$$

By the linearity of the integral, the left-hand side equals

$$\lim_{n \to \infty} \sum_{k=1}^{n} \int_{E \setminus E_0} f_k \, d\mu.$$

Hence,

$$\sum_{n=1}^{\infty} \int_{E \setminus E_0} f_k \, d\mu = \int_{E \setminus E_0} \sum_{k=1}^{\infty} f_k \, d\mu.$$

Finally, observe that since $\mu(E_0) = 0$, we have $\int_{E \setminus E_0} f_k \, d\mu = \int_E f_k \, d\mu$. ■

For simplicity in what follows we write that a property $\mathcal{P}$ holds for μ-a.e. $x \in E$ to mean that there is a measurable set E_0 with $\mu(E_0) = 0$ such that the property $\mathcal{P}$ holds for all $x \in E \setminus E_0$.

6 Product Spaces

Definition 55 *Given a measure space $(X, \mathfrak{M}, \mu)$, a set $E \in \mathfrak{M}$ has σ-*finite measure *if it can be written as a countable union of sets of finite measure, that is, if there exist $E_n \in \mathfrak{M}$, $n \in \mathbb{N}$, such that $\mu(E_n) < \infty$ and $\bigcup_{n=1}^{\infty} E_n = E$. If the entire space X has σ-finite measure, we say that the measure μ is σ-*finite.

We recall that, given two measurable spaces $(X, \mathfrak{M})$ and $(Y, \mathfrak{N})$ we denote by $\mathfrak{M} \otimes \mathfrak{N} \subseteq \mathcal{P}(X \times Y)$ the smallest σ-algebra that contains all sets of the form $E \times F$, where $E \in \mathfrak{M}$, $F \in \mathfrak{N}$. Then $\mathfrak{M} \otimes \mathfrak{N}$ is called the *product σ-algebra* of $\mathfrak{M}$ and $\mathfrak{N}$.

Exercise 56 *Let X and Y be metric spaces and let $\mathcal{B}(X)$ and $\mathcal{B}(Y)$ be their respective Borel σ-algebras. Prove that*

$$\mathcal{B}(X) \otimes \mathcal{B}(Y) \subseteq \mathcal{B}(X \times Y).$$

Prove that

$$\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2).$$

Let $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \nu)$ be two measure spaces. For every $E \in X \times Y$ define

$$(\mu \times \nu)^* (E) := \inf \left\{ \sum_{n=1}^{\infty} \mu(F_n) \nu(G_n) : \{F_n\}_n \text{ in } \mathfrak{M}, \; \{G_n\}_n \text{ in } \mathfrak{N}, \right. \\ \left. E \subseteq \bigcup_{n=1}^{\infty} (F_n \times G_n) \right\}. \tag{16}$$

By Proposition 7, $(\mu \times \nu)^* : \mathcal{P}(X) \to [0, \infty]$ is an outer measure, and it is called the *product outer measure* of μ and ν. By Carathéodory's theorem, the restriction of $(\mu \times \nu)^*$ to the σ-algebra $\mathfrak{M} \times \mathfrak{N}$ of $(\mu \times \nu)^*$-measurable sets is a complete measure, denoted by $\mu \times \nu$ and called the *product measure* of μ and ν.

Note that $\mathfrak{M} \boxtimes \mathfrak{N}$ is, in general, larger than the product σ-algebra $\mathfrak{M} \otimes \mathfrak{N}$.

Theorem 57 *Let $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \nu)$ be two measure spaces.*

(i) If $F \in \mathfrak{M}$ and $G \in \mathfrak{N}$, then $F \times G$ is $(\mu \times \nu)^*$-measurable and

$$(\mu \times \nu)(F \times G) = \mu(F) \nu(G); \tag{17}$$

(ii) if μ and ν are complete and E has σ-finite $\mu \times \nu$ measure, then for μ a.e. $x \in X$ the section

$$E_x := \{y \in Y : (x, y) \in E\}$$

belongs to the σ-algebra $\mathfrak{N}$, and for ν a.e. $y \in Y$ the section

$$E_y := \{x \in X : (x, y) \in E\}$$

belongs to the σ-algebra $\mathfrak{M}$. Moreover, the functions $y \mapsto \mu(E_y)$ and $x \mapsto \nu(E_x)$ are measurable and

$$(\mu \times \nu)(E) = \int_Y \mu(E_y)\, d\nu(y) = \int_X \nu(E_x)\, d\mu(x).$$

Friday, September 23, 2022

Remark 58 *If $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \nu)$ are two measure spaces, then $\mu \times \nu : \mathfrak{M} \boxtimes \mathfrak{N} \to [0, \infty]$ is complete. On the other hand, $\mu \times \nu : \mathfrak{M} \otimes \mathfrak{N} \to [0, \infty]$ is not complete in general even if and μ and ν are complete. Indeed, if there exists a nonempty set $F \in \mathfrak{M}$ such that $\mu(F) = 0$ and a set $G \in \mathcal{P}(Y) \setminus \mathfrak{N}$, then the set $F \times G$ belongs to $\mathfrak{M} \boxtimes \mathfrak{N}$ since $F \times G \subseteq F \times Y$ and $(\mu \times \nu)(F \times Y) = \mu(F)\nu(Y) = 0$. On the other hand, by the previous exercise we have that $F \times G$ does not belong $\mathfrak{M} \otimes \mathfrak{N}$, since for every $x \in F$ the section*

$$(F \times G)_x = G$$

does not belong to $\mathfrak{N}$. In particular this can be applied to $\mathcal{L}^1 \times \mathcal{L}^1$ since we have shown that there exist sets that are not Lebesgue measurable.

Exercise 59 *Let $N = m + k$, where N, n, $m \in \mathbb{N}$. Prove that $(\mathcal{L}^n \times \mathcal{L}^m)^* = \mathcal{L}_o^N$.*

The previous result is a particular case of Tonelli's theorem in the case that $f = \chi_E$.

Theorem 60 (Tonelli) *Let $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \nu)$ be two measure spaces. Assume that μ and ν are complete and σ-finite, and let $f : X \times Y \to [0, \infty]$ be an $\mathfrak{M} \boxtimes \mathfrak{N}$ measurable function. Then for μ a.e. $x \in X$ the function $f(x, \cdot)$ is measurable and the function $\int_Y f(\cdot, y)\, d\nu(y)$ is measurable. Similarly, for ν a.e. $y \in Y$ the function $f(\cdot, y)$ is measurable and the function $\int_X f(x, \cdot)\, d\mu(x)$ is measurable. Moreover,*

$$\begin{aligned}\int_{X \times Y} f(x, y)\, d(\mu \times \nu)(x, y) &= \int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x) \\ &= \int_Y \left(\int_X f(x, y)\, d\mu(x) \right) d\nu(y).\end{aligned}$$

Proof. If $f = \chi_E$ or, more generally, if

$$f = \sum_{n=1}^{\ell} c_n \chi_{E_n},$$

then the result follows from the previous theorem. If $f : X \times Y \to [0, \infty]$ is an arbitrary $\mathfrak{M} \boxtimes \mathfrak{N}$ measurable function, then by Remark 34 there exists a sequence $\{s_n\}_n$ of simple functions $s_n : X \times Y \to [0, \infty)$ such that

$$0 \le s_1(x, y) \le s_2(x, y) \le \ldots \le s_n(x, y) \to f(x, y)$$

for every $(x, y) \in X \times Y$. By the Lebesgue monotone convergence theorem (applied twice) we have

$$\begin{aligned}
\int_{X\times Y} f(x,y)\, d(\mu\times\nu)(x,y) &= \lim_{n\to\infty}\int_{X\times Y} s_n(x,y)\, d(\mu\times\nu)(x,y)\\
&= \lim_{n\to\infty}\int_X \left(\int_Y s_n(x,y)\, d\nu(y)\right) d\mu(x)\\
&= \int_X \left(\lim_{n\to\infty}\int_Y s_n(x,y)\, d\nu(y)\right) d\mu(x).
\end{aligned}$$

Since by the previous theorem for all $n \in \mathbb{N}$ and for μ a.e. $x \in X$ the functions

$$y \in Y \mapsto s_n(x,y)$$

are measurable, we may apply again Lebesgue monotone convergence theorem to conclude that for μ a.e. $x \in X$,

$$\lim_{n\to\infty}\int_Y s_n(x,y)\, d\nu(y) = \int_Y f(x,y)\, d\nu(y),$$

and so

$$\int_{X\times Y} f(x,y)\, d(\mu\times\nu)(x,y) = \int_X \left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x).$$

Similarly, we have

$$\int_{X\times Y} f(x,y)\, d(\mu\times\nu)(x,y) = \int_Y \left(\int_X f(x,y)\;\, d\mu(x)\right) d\nu(y).$$

■

Exercise 61 *Prove that in the case that $f : X \times Y \to [0, \infty]$ is $\mathfrak{M} \otimes \mathfrak{N}$ measurable, then Tonelli's theorem still holds even if the measures μ and ν are not complete, and the statements are satisfied for every $x \in X$ and $y \in Y$ (as opposed to for μ a.e. $x \in X$ and for ν a.e. $y \in Y$).*

The version of Tonelli's theorem for integrable functions of arbitrary sign is the well–known Fubini's theorem:

Theorem 62 (Fubini) *Let $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \nu)$ be two measure spaces. Assume that μ and ν are complete, and let $f : X\times Y \to [-\infty, \infty]$ be $\mu\times\nu$-integrable. Then for μ a.e. $x \in X$ the function $f(x, \cdot)$ is ν-integrable, and the function $\int_Y f(\cdot, y)\, d\nu(y)$ is μ-integrable.*

Similarly, for ν a.e. $y \in Y$ the function $f(\cdot, y)$ is μ-integrable, and the function $\int_X f(x, \cdot)\; d\mu(x)$ is ν-integrable. Moreover,

$$\begin{aligned}
\int_{X\times Y} f(x,y)\, d(\mu\times\nu)(x,y) &= \int_X \left(\int_Y f(x,y)\, d\nu(y)\right) d\mu(x)\\
&= \int_Y \left(\int_X f(x,y)\;\, d\mu(x)\right) d\nu(y).
\end{aligned}$$

Proof. The proof is very similar to that of Tonelli's theorem. We consider first the case in which f is a characteristic function, then a simple function, then an nonnegative integrable function, and finally use the fact that $f = f^+ - f^-$. Note that, since f is $\mu \times \nu$-integrable, by Remark 48 the set

$$E := \{(x, y) \in E : |f(x, y)| > 0\}$$

has σ-finite $\mu \times \nu$ measure. Thus, we are in a position to apply Theorem 57(ii). ■

Exercise 63 *Prove that in the case that $f : X \times Y \to [-\infty, \infty]$ is $\mathfrak{M} \otimes \mathfrak{N}$ measurable, then Fubini's theorem still holds even if the measures μ and ν are not complete.*

Example 64 *The next example shows that Fubini's theorem fails without assuming the integrability of the function f. Consider the function*

$$f(x, y) := \frac{x^2 - y^2}{(x^2 + y^2)^2}, \quad (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

We showed this in 21-269 that

$$\int_0^1 \left(\int_0^1 \int \frac{x^2 - y^2}{(x^2 + y^2)^2} \, dx \right) dy = -\int_0^1 \frac{1}{y^2 + 1} \, dy = -\frac{1}{4}\pi,$$

while

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} \, dy \right) dx = \int_0^1 \frac{1}{x^2 + 1} \, dx = \frac{1}{4}\pi.$$

Since, by Tonelli's theorem

$$\begin{aligned} \int_{[0,1] \times [0,1]} \left(\frac{x^2 - y^2}{(x^2 + y^2)^2} \right)^+ dxdy &= \int_0^1 \left(\int_0^1 \left(\frac{x^2 - y^2}{(x^2 + y^2)^2} \right)^+ dx \right) dy \\ &= \int_0^1 \left(\int_0^1 \left(\frac{x^2 - y^2}{(x^2 + y^2)^2} \right)^+ dy \right) dx \end{aligned}$$

and

$$\begin{aligned} \int_{[0,1] \times [0,1]} \left(\frac{x^2 - y^2}{(x^2 + y^2)^2} \right)^- dxdy &= \int_0^1 \left(\int_0^1 \left(\frac{x^2 - y^2}{(x^2 + y^2)^2} \right)^- dx \right) dy \\ &= \int_0^1 \left(\int_0^1 \left(\frac{x^2 - y^2}{(x^2 + y^2)^2} \right)^- dy \right) dx, \end{aligned}$$

this implies that

$$\int_{[0,1]\times[0,1]} \left(\frac{x^2-y^2}{(x^2+y^2)^2}\right)^+ dxdy = \int_{[0,1]\times[0,1]} \left(\frac{x^2-y^2}{(x^2+y^2)^2}\right)^- dxdy = \infty,$$

so that the Lebesgue integral of f is not defined.

Exercise 65 *Prove that the function*

$$f(x,y) := \frac{\sin^3 x}{x^4+y^2}, \quad (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$$

is Lebesgue integrable over the set $E = \left\{(x,y) \in \mathbb{R}^2 : \, y > x^2, \, x > 0\right\}$ and compute

$$\int_E \frac{\sin^3 x}{x^4+y^2}\, dxdy.$$

Monday, September 26, 2022

7 Lebesgue's Differentiation Theorem

In this section we prove that a monotone function is differentiable at all points except at most a set of Lebesgue measure zero.

Theorem 66 (Lebesgue's Differentiation Theorem) *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a monotone function. Then there exists a set $E \subset I$ of Lebesgue measure zero such that f is differentiable in $I \setminus E$.*

The proof relies on the following covering lemmas.

Lemma 67 *Let $E \subset \mathbb{R}$ be a bounded set and let $\mathcal{F}$ be a family of open intervals with the property that each $x \in E$ is the left endpoint of an interval $(x, x+h_x)$ in $\mathcal{F}$. Then for every $\varepsilon > 0$ there exist disjoint intervals $I_1, \ldots, I_n \in \mathcal{F}$ such that*

$$\mathcal{L}^1_o\left(E \cap \bigcup_{k=1}^n I_k\right) \geq \mathcal{L}^1_o(E) - \varepsilon.$$

Proof. Define

$$E_n := \left\{x \in E : \; h_x > \frac{1}{n}\right\}.$$

Then $E_n \subseteq E_{n+1}$ and

$$\bigcup_{n=1}^\infty E_n = E.$$

One of the properties of the Lebesgue outer measures is that (Exercise)

$$\lim_{n\to\infty} \mathcal{L}^1_o(E_n) = \mathcal{L}^1_o(E).$$

Hence we can find m so large that

$$\mathcal{L}_o^1(E_m) > \mathcal{L}_o^1(E) - \frac{\varepsilon}{2}.$$

Let $a_1 := \inf E_m$, $b_1 := \sup E_m$, and let $\ell := b_1 - a_1 > 0$. Given

$$\eta = \frac{\varepsilon}{2(m\ell + 1)},$$

by the definition of infimum we can find $x_1 \in E_m$ with $a_1 \leq x_1 < a_1 + \eta$. By definition of E_m there exists an interval $I_1 = (x_1, x_1 + h_1)$ in $\mathcal{F}$, with $h_1 > \frac{1}{m}$. If $x_1 + h_1 \geq b_1$, then we stop.

If $x_1 + h_1 < b_1$, let

$$a_2 := \inf\{x \in E_m : \ x \geq x_1 + h_1\}.$$

Then by the definition of infimum we can find $x_2 \in E_m$ with $a_2 \leq x_2 < a_2 + \eta$. By definition of E_m there exists an interval $I_2 = (x_2, x_2 + h_2)$ in $\mathcal{F}$, with $h_2 > \frac{1}{m}$. If $x_2 + h_2 \geq b_1$, we stop, while if $x_2 + h_2 < b_1$ we define

$$a_3 := \inf\{x \in E_m : \ x \geq x_2 + h_2\}.$$

We continue in this way constructing intervals I_k until $x_k + h_k < b_1$. Since each interval I_k has length larger than $\frac{1}{m}$, we have that we will find at most n intervals with $n < m\ell + 1$. Let

$$S := \bigcup_{k=1}^{n} I_k, \quad T := \bigcup_{k=1}^{n} (x_k - \eta, x_k].$$

Then $x_k - \eta \leq a_k$ and so $E_m \subseteq S \cup T$. Moreover the intervals I_k are disjoint by construction. Now

$$\begin{aligned}
\mathcal{L}_o^1(E) - \frac{\varepsilon}{2} &< \mathcal{L}_o^1(E_m) \leq \mathcal{L}_o^1(E_m \cap S) + \mathcal{L}_o^1(E_m \cap T) \\
&\leq \mathcal{L}_o^1(E_m \cap S) + \mathcal{L}_o^1(T) \leq \mathcal{L}_o^1(E_m \cap S) + \sum_{k=1}^{n} \mathcal{L}_o^1((x_k - \eta, x_k]) \\
&= \mathcal{L}_o^1(E_m \cap S) + n\eta \leq \mathcal{L}_o^1(E_m \cap S) + \frac{\varepsilon}{2},
\end{aligned}$$

and so $\mathcal{L}_o^1(E_m \cap S) \geq \mathcal{L}_o^1(E) - \varepsilon$. In turn,

$$\mathcal{L}_o^1\left(E \cap \bigcup_{k=1}^{n} I_k\right) \geq \mathcal{L}_o^1(E_m \cap S) \geq \mathcal{L}_o^1(E) - \varepsilon,$$

which concludes the proof. ■

Lemma 68 *Let $E \subset \mathbb{R}$ be a bounded set and let $\mathcal{F}$ be a family of open intervals with the property that each $x \in E$ is the left endpoint of an interval $(x, x + h_x)$*

in $\mathcal{F}$ with h_x arbitrarily small (that is, for every $\eta > 0$ there is one interval with $h_x < \eta$). Then for every $\varepsilon > 0$ there there exist disjoint intervals $I_1, \ldots, I_n \in \mathcal{F}$ such that

$$\mathcal{L}_o^1\left(E \cap \bigcup_{k=1}^n I_k\right) \geq \mathcal{L}_o^1(E) - \varepsilon, \quad \sum_{k=1}^n \text{length}\, I_k \leq \mathcal{L}_o^1(E) + \varepsilon.$$

Proof. Consider an open set $U \supseteq E$ such that

$$\mathcal{L}_o^1(U) \leq \mathcal{L}_o^1(E) + \varepsilon.$$

Let $\mathcal{F}'$ be the subfamily of intervals $(x, x + h_x)$ in $\mathcal{F}$ contained in U. Note that for each $x \in E \subseteq U$ there must exist at least one such interval, since U contains a ball centered at x and there are intervals of arbitrarily small length.

Apply the previous lemma to the family $\mathcal{F}'$ to find disjoint intervals $I_1, \ldots, I_n \in \mathcal{F}'$ such that

$$\mathcal{L}_o^1\left(E \cap \bigcup_{k=1}^n I_k\right) \geq \mathcal{L}_o^1(E) - \varepsilon.$$

Since the intervals are disjoint and contained in U, it follows that

$$\sum_{k=1}^n \text{length}\, I_k \leq \mathcal{L}_o^1(U) \leq \mathcal{L}_o^1(E) + \varepsilon.$$

This concludes the proof. ■

Wednesday, September 28, 2022

The last lemma is of interest in itself.

Lemma 69 *Let $I \subseteq \mathbb{R}$ be an interval, let $f : I \to \mathbb{R}$, and let*

$$E := \{x \in I^\circ : \text{ there exist } f'_+(x) \text{ and } f'_-(x) \text{ and } f'_+(x) \neq f'_-(x)\}.$$

Then E is countable.

Proof. Write $\mathbb{Q} = \{r_n : n \in \mathbb{N}\}$ and consider the set $E_- := \{x \in E : f'_-(x) < f'_+(x)\}$. By the density of the rationals, there exist countably many rationals in the interval $(f'_-(x), f'_+(x))$. Let $m \in \mathbb{N}$ be the smallest integer such that $f'_-(x) < r_m < f'_+(x)$. Since

$$\lim_{y \to x^-} \frac{f(y) - f(x)}{y - x} < r_m$$

let $p \in \mathbb{N}$ be the smallest integer such that

$$\frac{f(y) - f(x)}{y - x} < r_m$$

for all $r_p < y < x$. On the other hand, since

$$\lim_{y \to x^+} \frac{f(y) - f(x)}{y - x} > r_m$$

let $q \in \mathbb{N}$ be the smallest integer such that

$$\frac{f(y) - f(x)}{y - x} > r_m$$

for all $x < y < r_q$. It follows that

$$f(y) - f(x) > r_m(y - x)$$

for all $r_p < y < r_q$ with $y \neq x$.

Thus we have shown that to each $x \in E_-$ we can uniquely associate three natural numbers (m_x, p_x, q_x) for which $f(y) - f(x) > r_{m_x}(y - x)$ for all $r_{p_x} < y < r_{q_x}$ with $y \neq x$. Next we claim that if $x, z \in E_-$ with $x \neq z$, then

$$(m_x, p_x, q_x) \neq (m_z, p_z, q_z).$$

Indeed, if $(m_x, p_x, q_x) = (m_z, p_z, q_z)$, then $f(y) - f(x) > r_{m_x}(y - x)$ for all $r_{p_x} < y < r_{q_x}$ with $y \neq x$. In particular, taking $y = z$ gives $f(z) - f(x) > r_{m_x}(z - x)$. But since $(m_x, p_x, q_x) = (m_z, p_z, q_z)$, we also have $f(x) - f(z) > r_{m_x}(x - z)$. Adding these two inequalities gives a contradiction. Thus the claim holds.

Hence, the function $x \in E_- \mapsto (m_x, p_x, q_x)$ is injective, which shows that the cardinality of E_- is at most the cardinality of $\mathbb{Q} \times \mathbb{Q} \times \mathbb{Q}$, that is, E_- is countable. ■

Next we recall the definitions of liminf and limsup. Let (X, d) be a metric space, $E \subseteq X$, and $f : E \to \mathbb{R}$. Assume that $x_0 \in X$ is an accumulation point of E. For every $r > 0$ define

$$g(r) := \inf_{E \cap (B(x_0, r) \setminus \{x_0\})} f.$$

Note that $g(r)$ is $-\infty$ if f is not bounded from below in $E \cap (B(x_0, r) \setminus \{x_0\})$. If $r_1 < r_2$ then $g(r_1) \geq g(r_2)$. Hence the function $g : (0, \infty)$ is decreasing. It follows that there exists

$$\lim_{r \to 0^+} g(r) = \sup_{(0,\infty)} g = \ell \in \overline{\mathbb{R}}.$$

This limit is called the *limit inferior of f as x approaches x_0* and is denoted

$$\liminf_{x \to x_0} f(x) \quad \text{or} \quad \varliminf_{x \to x_0} f(x).$$

On the other hand, for every $r > 0$ define

$$h(r) := \sup_{E \cap (B(x_0, r) \setminus \{x_0\})} f.$$

Note that $h(r)$ is ∞ if f is not bounded from above in $E \cap (B(x_0, r) \setminus \{x_0\})$. If $r_1 < r_2$ then $h(r_1) \leq h(r_2)$. Hence the function $h : (0, \infty)$ is increasing. It follows that there exists

$$\lim_{r \to 0^+} h(r) = \inf_{(0,\infty)} h = L \in \overline{\mathbb{R}}.$$

This limit is called the *limit superior of f as x approaches x_0* and is denoted

$$\limsup_{x\to x_0} f(x) \quad \text{or} \quad \overline{\lim_{x\to x_0}} f(x).$$

The following theorem is left as an exercise.

Theorem 70 *Let (X,d) be a metric space, let $E \subseteq X$ and let $f : E \to \mathbb{R}$. Assume that $x_0 \in X$ is an accumulation point of E. Then*

$$\liminf_{x\to x_0} f(x) \leq \limsup_{x\to x_0} f(x). \tag{18}$$

Moreover, there exists $\lim_{x\to x_0} f(x) = \ell \in \overline{\mathbb{R}}$ if and only if

$$\liminf_{x\to x_0} f(x) = \limsup_{x\to x_0} f(x) = \ell. \tag{19}$$

Given a set $E \subseteq \mathbb{R}$ and function $f : E \to \mathbb{R}$, for every $x_0 \in E$ such that $x_0 \in \operatorname{acc}(E \cap (-\infty, x_0))$ and $x_0 \in \operatorname{acc}(E \cap (x_0, \infty))$, the four *Dini's derivatives* of f are given by

$$\overline{D}_- f(x_0) := \limsup_{x\to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}, \quad \overline{D}_+ f(x_0) := \limsup_{x\to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}$$
$$\underline{D}_- f(x_0) := \liminf_{x\to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}, \quad \underline{D}_+ f(x_0) := \liminf_{x\to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}.$$

Friday, September 28, 2022

We now turn to the proof of Lebesgue's differentiation theorem.

Proof. Step 1: Assume that f is increasing and that I is bounded and let

$$E := \{x \in I^\circ : \ \underline{D}_+ f(x) < \overline{D}_+ f(x)\}.$$

We claim that E has Lebesgue measure zero. To see this we write E as a countable union of sets

$$E = \bigcup_{r,s\in\mathbb{Q},\ 0<r<s} E_{r,s}, \quad E_{r,s} := \left\{x \in E : \ \underline{D}_+ f(x) < r < s < \overline{D}_+ f(x)\right\}.$$

It is enough to prove that each set $E_{r,s}$ has Lebesgue measure zero. Since

$$\underline{D}_+ f(x) = \liminf_{y\to x^+} \frac{f(y) - f(x)}{y - x} = \lim_{R\to 0^+} \inf_{y\in(x,x+R)\cap I} \frac{f(y) - f(x)}{y - x} < r,$$

for each $x \in E_{r,s}$ there exist $R_1 > 0$ (depending on x) such that

$$\inf_{y\in(x,x+R)\cap I} \frac{f(y) - f(x)}{y - x} < r$$

for all $0 < R \leq R_1$. Since $x \in I^\circ$, by taking R_1 smaller, we can assume that $(x, x+R_1) \subseteq I$, so $(x, x+R) \cap I = (x, x+R)$ for all $0 < R \leq R_1$. Since r is not a lower bound, for every $0 < R \leq R_1$, we can find $y_R \in (x, x+R)$ such that

$$\frac{f(y_R) - f(x)}{y_R - x} < r.$$

Write $y_R = x + h_R$. In conclusion for each $x \in E_{r,s}$ we found countably many an open intervals $(x, x+h)$, where $h > 0$ is arbitrarily small, such that

$$\frac{f(x+h) - f(x)}{h} < r. \tag{20}$$

Let $\mathcal{F}$ be the family of all such intervals as x varies in $E_{r,s}$. By Lemma 68 for every $\varepsilon > 0$ there exist disjoint intervals $I_1, \ldots, I_n \in \mathcal{F}$ such that

$$\mathcal{L}_o^1\left(E_{r,s} \cap \bigcup_{k=1}^n I_k\right) \geq \mathcal{L}_o^1(E_{r,s}) - \varepsilon, \quad \sum_{k=1}^n \operatorname{length} I_k \leq \mathcal{L}_o^1(E_{r,s}) + \varepsilon. \tag{21}$$

Write $I_k = (x_k, x_k + h_k)$. Then by (20) and (21),

$$\sum_{k=1}^n f(x_k + h_k) - f(x_k) < r \sum_{k=1}^n h_k \leq r\mathcal{L}_o^1(E_{r,s}) + r\varepsilon. \tag{22}$$

Let $V := \bigcup_{k=1}^n I_k$. Note that V is open. Setting $F_{r,s} := E_{r,s} \cap V$, for each $x \in F_{r,s}$ we have that

$$\overline{D}_+ f(x) = \limsup_{y \to x^+} \frac{f(y) - f(x)}{y - x} = \lim_{R \to 0^+} \sup_{y \in (x, x+R) \cap I} \frac{f(y) - f(x)}{y - x} > s.$$

Hence, for each $x \in F_{r,s}$ there exist $R_2 > 0$ (depending on x) such that

$$\sup_{y \in (x, x+R) \cap I} \frac{f(y) - f(x)}{y - x} > s$$

for all $0 < R \leq R_2$. Since $x \in I^\circ$, by taking R_2 smaller, we can assume that $(x, x+R_2) \subseteq V \subseteq I$, so $(x, x+R) \cap I = (x, x+R)$ for all $0 < R \leq R_2$. Since s is not an upper bound, for every $0 < R \leq R_2$, we can find $y_R \in (x, x+R)$ such that

$$\frac{f(y_R) - f(x)}{y_R - x} > s.$$

Write $y_R = x + t_R$. In conclusion for each $x \in F_{r,s}$ we were able to find infinitely many open interval $(x, x+t) \subseteq V$, where $t > 0$ is arbitrarily small, such that

$$\frac{f(x+t) - f(x)}{t} > s. \tag{23}$$

Let $\mathcal{G}$ be the family of all such intervals. By Lemma 68 for every $\varepsilon > 0$ there there exist disjoint intervals $J_1, \ldots, J_m \in \mathcal{G}$ such that

$$\mathcal{L}_o^1\left(F_{r,s} \cap \bigcup_{i=1}^m J_i\right) \geq \mathcal{L}_o^1(F_{r,s}) - \varepsilon. \tag{24}$$

Write $J_i = (y_i, y_i + t_i)$. Then by (21), (23), and (24),

$$\sum_{i=1}^m f(y_i+t_i)-f(y_i) > s\sum_{i=1}^m t_i \geq s\mathcal{L}_o^1\left(F_{r,s} \cap \bigcup_{i=1}^m J_i\right) \geq s\mathcal{L}_o^1(F_{r,s})-s\varepsilon \geq s\mathcal{L}_o^1(E_{r,s})-2s\varepsilon. \tag{25}$$

But since each J_i is contained in $V = \bigcup_{k=1}^n I_k$, and since the intervals I_k are disjoint, it follows that each interval J_i is contained in some interval I_k. Since f is increasing it follows that

$$\sum_{i=1}^m f(y_i + t_i) - f(y_i) \leq \sum_{k=1}^n f(x_k + h_k) - f(x_k).$$

Combining this inequality with (22) and (25) gives

$$s\mathcal{L}_o^1(E_{r,s})-2s\varepsilon < \sum_{i=1}^m f(y_i+t_i)-f(y_i) \leq \sum_{k=1}^n f(x_k+h_k)-f(x_k) < r\mathcal{L}_o^1(E_{r,s})+r\varepsilon,$$

that is

$$(s-r)\mathcal{L}_o^1(E_{r,s}) < 2s\varepsilon + r\varepsilon.$$

Since $s - r > 0$, letting $\varepsilon \to 0^+$ we conclude that $\mathcal{L}_o^1(E_{r,s}) = 0$.

Hence, we have shown that $\mathcal{L}_o^1(E) = 0$. It follows that for all $x \in I^\circ \setminus E$ there exists the right derivative $f'_+(x)$ (possibly infinite).

With a similar proof we can show that the left derivative exist (possibly infinite) for all $x \in I^\circ$ except for a set of Lebesgue measure zero. It follows from Lemma 69 that there exists $f'(x)$ (possibly infinite) for all $x \in I$ except for a set of Lebesgue measure zero.

Step 2: Let

$$F := \{x \in I^\circ : \ f'_+(x) = \infty\}.$$

We leave as an exercise to prove that $\mathcal{L}_o^1(F) = 0$. ■

Monday, October 3, 2022

Corollary 71 *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be an increasing function. Then for every $a, b \in I$ with $a < b$,*

$$\int_a^b f'(x)\, dx \leq f(b) - f(a).$$

Proof. Let $a, b \in I$ with $a < b$. Consider the function $g : [a, \infty) \to \mathbb{R}$ given by

$$g(x) = \begin{cases} f(x) & \text{if } x \in [a, b], \\ f(b) & \text{if } x \geq b, \end{cases}$$

and define

$$g_n(x) := \frac{g(x + \frac{1}{n}) - g(x)}{\frac{1}{n}}, \quad x \in [a, b].$$

Then $g_n \geq 0$, $g_n(x) \to g'(x)$ for $\mathcal{L}^1$ a.e. $x \in [a, b]$. Moreover, g_n is measurable, since monotone functions are measurable and differences of measurable functions are measurable. By Fatou's lemma

$$\int_a^b g'(x)\, dx \leq \liminf_{n \to \infty} \int_a^b g_n(x)\, dx$$

On the other hand, for every $h > 0$,

$$\begin{aligned} \frac{1}{h} \int_a^b [g(x+h) - g(x)]\; dx &= \frac{1}{h} \left\{ \int_b^{b+h} g(x)\; dx - \int_a^{a+h} g(x)\; dx \right\} \\ &\leq \frac{1}{h} \{(g(b) - g(a))\, h\} = g(b) - g(a). \end{aligned} \tag{26}$$

Hence, taking $h = \frac{1}{n}$ gives the result for g. To conclude, observe that $g'(x) = f'(x)$ for all $x \in (a, b)$ where the derivative exists, and that $f(a) = g(a)$ and $f(b) = g(b)$. ■

In what follows, given an interval $I \subseteq \mathbb{R}$, a *partition* of I is a finite set $P := \{x_0, \ldots, x_n\} \subset I$, where

$$x_0 < x_1 < \cdots < x_n.$$

Definition 72 *Let $I \subseteq \mathbb{R}$ be an interval and $\boldsymbol{f} : I \to \mathbb{R}^N$. The* pointwise variation *of $\boldsymbol{f}$ on the interval I is*

$$\operatorname{Var} \boldsymbol{f} := \sup \left\{ \sum_{i=1}^{n} \| \boldsymbol{f}(x_i) - \boldsymbol{f}(x_{i-1}) \| \right\},$$

where the supremum is taken over all partitions $P := \{x_0, \ldots, x_n\}$ of I, $n \in \mathbb{N}$. A function $\boldsymbol{f} : I \to \mathbb{R}^N$ has finite *or* bounded pointwise variation *if* $\operatorname{Var} F < \infty$.

The space of all functions $\boldsymbol{f} : I \to \mathbb{R}^N$ of bounded pointwise variation is denoted by $BV\left(I; \mathbb{R}^N\right)$.

Remark 73 *We can give a similar definition for functions $f : I \to X$, where (X, d) is a metric space. The only difference is that*

$$\operatorname{Var} f := \sup \left\{ \sum_{i=1}^{n} d(f(x_i), f(x_{i-1})) \right\}$$

When $X = \mathbb{R}$ we write $BV(I)$ for $BV(I;\mathbb{R}^N)$.

To highlight the dependence on the interval I, we will sometimes write $\operatorname{Var}_I \boldsymbol{f}$.

A function $\boldsymbol{f} : I \to \mathbb{R}^N$ has *locally finite* or *locally bounded pointwise variation* if $\operatorname{Var}_{[a,b]} \boldsymbol{f} < \infty$ for all intervals $[a,b] \subset I$. The space of all functions $\boldsymbol{f} : I \to \mathbb{R}^N$ of locally bounded pointwise variation is denoted by $BV_{\text{loc}}(I;\mathbb{R}^N)$.

It almost goes without saying that if $I = [a,b]$, then

$$BV_{\text{loc}}\left([a,b];\mathbb{R}^N\right) = BV\left([a,b];\mathbb{R}^N\right).$$

Theorem 74 (Indefinite pointwise variation) *Let $I \subseteq \mathbb{R}$ be an interval, $c \in I$, and $\boldsymbol{f} \in BV_{\text{loc}}(I;\mathbb{R}^N)$. For every $x \in I$ define*

$$v(x) := \begin{cases} \operatorname{Var}_{[c,x]} \boldsymbol{f} & \text{if } x \geq c, \\ -\operatorname{Var}_{[x,c]} \boldsymbol{f} & \text{if } x < c. \end{cases} \tag{27}$$

Then for all $x, y \in I$, with $x < y$,

$$\|\boldsymbol{f}(y) - \boldsymbol{f}(x)\| \leq v(y) - v(x) = \operatorname{Var}_{[x,y]} \boldsymbol{f}. \tag{28}$$

In particular v is increasing and $\boldsymbol{f}$ is continuous at all but countably many points of I. Moreover, there exist

$$\boldsymbol{f}^-(x) = \lim_{y \to x^-} \boldsymbol{f}(y), \quad \boldsymbol{f}^+(x) = \lim_{y \to x^+} \boldsymbol{f}(y)$$

for all $x \in I^\circ$. Finally, if $N = 1$, the functions $v \pm f$ are increasing.

Proof. You proved the first part in 21-269. ■

Theorem 75 *Let $I \subseteq \mathbb{R}$ be an interval. Then every function in $BV_{\text{loc}}(I)$ is differentiable for $\mathcal{L}^1$-a.e.$x \in I$.*

Proof.

■

Wednesday, October 5, 2022

We now give an example of a continuous, nowhere differentiable function.

Theorem 76 *Let $f(x) = |x|$ for $x \in [-1,1]$ and extend f to $\mathbb{R}$ as a periodic function of period 2. Then the function*

$$g(x) = \sum_{n=1}^{\infty} \left(\frac{3}{4}\right)^n f(4^n x), \quad x \in \mathbb{R},$$

is real-valued, continuous, and nowhere differentiable.

Proof. Let $f_n(x) = \left(\frac{3}{4}\right)^n f(4^n x)$, $x \in \mathbb{R}$. Note that f_n is continuous. Consider the series

$$\sum_{n=1}^{\infty} \sup_{x \in \mathbb{R}} |f_n(x)|.$$

Since f is periodic, $\sup_{y\in\mathbb{R}}|f(y)| = \sup_{y\in[-1,1]}|f(y)| = 1$, and so $|f_n(x)| \le \left(\frac{3}{4}\right)^n$. In turn $\sum_{n=1}^{\infty}\sup_{x\in\mathbb{R}}|f_n(x)| \le \sum_{n=1}^{\infty}\left(\frac{3}{4}\right)^n$. It follows from stuff done in 21-269 that the series $\sum_{n=1}^{\infty} f_n$ converges uniformly to a continuous function g.

Next we prove that g is nowhere differentiable. Fix $x \in \mathbb{R}$. We are going to construct a sequence $h_m \to 0$ such that $\frac{g(x+h_m)-g(x)}{h_m} \to \infty$ as $m \to \infty$. We take $h_m = \pm\frac{1}{2}\frac{1}{4^m}$, where the sign is chosen in such a way that in the open interval of endpoints $4^m x$ and $4^m(x+h_m)$ there is no integer. Let's prove that we can always do this. We have $4^m(x+\frac{1}{2}\frac{1}{4^m}) - 4^m(x-\frac{1}{2}\frac{1}{4^m}) = 1$. If both $4^m(x+\frac{1}{2}\frac{1}{4^m})$ and $4^m(x-\frac{1}{2}\frac{1}{4^m})$ are integers, then in the interval $(4^m(x-\frac{1}{2}\frac{1}{4^m}), 4^m(x+\frac{1}{2}\frac{1}{4^m}))$ there is no integer and so we can take the sign of h_m as we like. If $4^m(x+\frac{1}{2}\frac{1}{4^m})$ and $4^m(x-\frac{1}{2}\frac{1}{4^m})$ are not both integers, then in the interval $(4^m(x-\frac{1}{2}\frac{1}{4^m}), 4^m(x+\frac{1}{2}\frac{1}{4^m}))$ there is exactly one integer. If this integer is $4^m x$ then we the sign of h_m as we like. If the integer is in $(4^m(x-\frac{1}{2}\frac{1}{4^m}), 4^m x)$, then we take $h_m = \frac{1}{2}\frac{1}{4^m}$, while if the integer is in $(4^m x, 4^m(x+\frac{1}{2}\frac{1}{4^m}))$, then we take $h_m = -\frac{1}{2}\frac{1}{4^m}$.

We now study

$$\begin{aligned}\frac{f_n(x+h_m)-f_n(x)}{h_m} &= \frac{\left(\frac{3}{4}\right)^n f(4^n(x+h_m)) - \left(\frac{3}{4}\right)^n f(4^n x)}{h_m} \\ &= \left(\frac{3}{4}\right)^n \frac{f(4^n x \pm \frac{1}{2}4^{n-m}) - f(4^n x)}{\pm\frac{1}{2}\frac{1}{4^m}}.\end{aligned}$$

If $n > m$ then $\frac{1}{2}4^{n-m}$ is an even integer and so by the periodicity of f the difference quotient is zero. If $n = m$ then since in the open interval of endpoints $4^m x$ and $4^m(x+h_m)$ there is no integer we have that the points $(x+h_m, f(4^m(x+h_m)))$ and $(x, f(4^m x))$ lie in the same line of the graph of f with slope either 1 or -1. Hence,

$$\left|\frac{f_m(x+h_m)-f_m(x)}{h_m}\right| = \left(\frac{3}{4}\right)^m \frac{|f(4^m(x+h_m)) - f(4^m x)|}{|h_m|} = \left(\frac{3}{4}\right)^m \frac{4^m|h_m|}{|h_m|} = 3^m.$$

Finally, if $n < m$, then using the fact that f is Lipschitz continuous with Lipschitz constant 1 we get

$$\left|\frac{f_n(x+h_m)-f_n(x)}{h_m}\right| = \left(\frac{3}{4}\right)^n \frac{|f(4^n(x+h_m)) - f(4^n x)|}{|h_m|} \le \left(\frac{3}{4}\right)^n \frac{4^n|h_m|}{|h_m|} = 3^n.$$

Hence,

$$\frac{g(x+h_m)-g(x)}{h_m} = \sum_{n=1}^{m}\frac{f_n(x+h_m)-f_n(x)}{h_m}$$

and using the inequality $|a+b| \geq |b| - |a|$ we get

$$\begin{aligned}\left|\frac{g(x+h_m)-g(x)}{h_m}\right| &= \left|\frac{f_m(x+h_m)-f_m(x)}{h_m} + \sum_{n=1}^{m-1}\frac{f_n(x+h_m)-f_n(x)}{h_m}\right| \\ &\geq \left|\frac{f_m(x+h_m)-f_m(x)}{h_m}\right| - \sum_{n=1}^{m-1}\left|\frac{f_n(x+h_m)-f_n(x)}{h_m}\right| \\ &\geq 3^m - \sum_{n=1}^{m-1} 3^n = 3^m - \frac{1}{2}3^m + \frac{3}{2} = \frac{1}{2}3^m + \frac{3}{2} \to \infty\end{aligned}$$

as $m \to \infty$. This concludes the proof. ■

8 The Fundamental Theorem of Calculus

Next we study the fundamental theorem of calculus for Lebesgue's integration. The Cantor function $f : [0,1] \to \mathbb{R}$ is a continuous, increasing function with derivative $f'(x) = 0$ for $\mathcal{L}^1$ a.e. $x \in [0,1]$, which does not satisfy the fundamental theorem of calculus since

$$1 - 0 = f(1) - f(0) > \int_0^1 f'(x)\,dx = 0.$$

It turns out that the functions which satisfy the fundamental theorem of calculus for the Lebesgue integration are absolutely continuous.

Definition 77 *Let $I \subseteq \mathbb{R}$ be an interval. A function $\boldsymbol{f} : I \to \mathbb{R}^N$ is said to be* absolutely continuous *on I if for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\sum_{k=1}^{\ell} \|\boldsymbol{f}(b_k) - \boldsymbol{f}(a_k)\| \leq \varepsilon \tag{29}$$

for every finite number of nonoverlapping intervals (a_k, b_k), $k = 1, \ldots, \ell$, with $[a_k, b_k] \subseteq I$ and

$$\sum_{k=1}^{\ell} (b_k - a_k) \leq \delta.$$

The space of all absolutely continuous functions $\boldsymbol{f} : I \to \mathbb{R}^N$ is denoted by $AC\left(I;\mathbb{R}^N\right)$. When $N = 1$ we simply write $AC(I)$.

Friday, October 7, 2022

Remark 78 *Note that since ℓ is arbitrary, we can also take $\ell = \infty$, namely, replace finite sums by series.*

Example 79 *Let $I \subseteq \mathbb{R}$ be an interval. If $\boldsymbol{f} : I \to \mathbb{R}^N$ is Lipschitz continuous with Lipschitz constant L, then*

$$\sum_{k=1}^{\ell} \|\boldsymbol{f}(b_k) - \boldsymbol{f}(a_k)\| \leq L \sum_{k=1}^{\ell} (b_k - a_k) \leq \varepsilon$$

provided we take $\delta = \frac{\varepsilon}{L+1}$.

Next we show that absolutely continuous functions have bounded variation.

Proposition 80 *Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^N$ be absolutely continuous. Then $\boldsymbol{f}$ has finite variation. In particular, $\boldsymbol{f}$ is differentiable for $\mathcal{L}^1$-a.e. $x \in [a,b]$.*

Proof. Take $\varepsilon = 1$, and let $\delta > 0$ be as in Definition 77. Let n be the integer part of $\frac{2(b-a)}{\delta}$ and partition $[a,b]$ into n intervals $[x_{i-1}, x_i]$ of equal length $\frac{b-a}{n}$,

$$a = x_0 < x_1 < \cdots < x_n = b.$$

Since $\frac{b-a}{n} \leq \delta$, in view of (29), on each interval $[x_{i-1}, x_i]$ we have that $\operatorname{Var}_{[x_{i-1},x_i]} \boldsymbol{f} \leq 1$, and so by the previous exercise

$$\operatorname{Var}_{[a,b]} \boldsymbol{f} = \sum_{i=1}^{n} \operatorname{Var}_{[x_{i-1},x_i]} \boldsymbol{f} \leq n \leq \frac{2(b-a)}{\delta} < \infty,$$

where we have used the fact that $\frac{b-a}{n} \geq \frac{\delta}{2}$.

The last part of the statement follows from the fact that any function $\boldsymbol{f} : [a,b] \to \mathbb{R}$ of bounded variation is differentiable for $\mathcal{L}^1$-a.e. $x \in [a,b]$ by Theorem 75. ■

Theorem 81 *Let $I \subseteq \mathbb{R}$ be an open interval and let $\boldsymbol{f} : I \to \mathbb{R}^N$ be an absolutely continuous function such that there exists $\boldsymbol{f}'(x) = 0$ for $\mathcal{L}^1$ a.e. $x \in I$. Then $\boldsymbol{f}$ is constant.*

Proof. Given $\varepsilon > 0$, let $\delta > 0$ be the number given in the definition of absolute continuity. Let $a, b \in I$ with $a < b$. We claim that $\boldsymbol{f}(a) = \boldsymbol{f}(b)$. Let $E := \left\{x \in (a,b) : \boldsymbol{f}'(x) = 0\right\}$. Then $\mathcal{L}_o^1(E) = b - a$.

For every $x \in E$, we have that

$$\lim_{y \to x} \frac{\boldsymbol{f}(y) - \boldsymbol{f}(x)}{y - x} = \boldsymbol{f}'(x) = 0,$$

and so there exists $h_x > 0$ such that $[x - h_x, x + h_x] \subset (a,b)$ and

$$\left\| \frac{\boldsymbol{f}(y) - \boldsymbol{f}(x)}{y - x} \right\| \leq \varepsilon \tag{30}$$

for all $y \in I$ with $|x-y| \le h_x$. Consider the family $\mathcal{F}$ of intervals $(x, x+h)$, where $x \in E$ and $0 < h \le h_x$. By Lemma 68 there exist disjoint intervals $(x_1, x_1+h_1), \ldots, (x_n, x_n+h_n) \in \mathcal{F}$ such that

$$\mathcal{L}_o^1\left(E \cap \bigcup_{k=1}^{n}(x_n, x_n+h_n)\right) \ge b-a-\delta, \quad \sum_{k=1}^{n} h_k \le b-a+\delta.$$

Without loss of generality assume that $x_1 < x_2 < \cdots < x_n$. Since

$$\sum_{k=1}^{n} h_k \ge \mathcal{L}_o^1\left(E \cap \bigcup_{k=1}^{n}(x_n, x_n+h_n)\right) \ge b-a-\delta,$$

the sum of the length of the intervals $[a, x_1]$, $[x_1+h_1, x_2]$, $\ldots$, $[x_n+h_n, b]$ is less than or equal δ. Since $\boldsymbol{f}$ is absolutely continuous, we have that

$$\|\boldsymbol{f}\,(a)-\boldsymbol{f}\,(x_1)\|+\sum_{k=1}^{n-1}\|\boldsymbol{f}\,(x_{k+1})-\boldsymbol{f}\,(x_k+h_k)\|+\|\boldsymbol{f}\,(b)-\boldsymbol{f}\,(x+h_n)\| \le \varepsilon.$$

On the other hand by (30),

$$\|\boldsymbol{f}\,(x_k+h_k)-\boldsymbol{f}\,(x_k)\| \le \varepsilon h_k$$

and so

$$\begin{aligned}\|\boldsymbol{f}\,(a)-\boldsymbol{f}\,(b)\| &\le \|\boldsymbol{f}\,(a)-\boldsymbol{f}\,(x_1)\|+\sum_{k=1}^{n-1}\|\boldsymbol{f}\,(x_{k+1})-\boldsymbol{f}\,(x_k+h_k)\| \\ &\quad+\sum_{k=1}^{n}\|\boldsymbol{f}\,(x_k)-\boldsymbol{f}\,(x_k+h_k)\|+\|\boldsymbol{f}\,(b)-\boldsymbol{f}\,(x_n+h_n)\| \\ &\le \varepsilon+\varepsilon\sum_{k=1}^{n} h_k \le \varepsilon+\varepsilon(b-a).\end{aligned}$$

Letting $\varepsilon \to 0^+$ gives $\boldsymbol{f}\,(a)=\boldsymbol{f}\,(b)$. Hence, $\boldsymbol{f}$ is constant. ∎

Theorem 82 *Let $g : [a, b] \to \mathbb{R}$ be a Lebesgue integrable function and let*

$$f(x) := \int_a^x g(t)\; dt.$$

Then f is absolutely continuous.

We begin with an auxiliary result.

Lemma 83 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set and let $g : E \to \mathbb{R}$ be a Lebesgue integrable function. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $F \subseteq E$ is a Lebesgue measurable set with $\mathcal{L}^N(F) \le \delta$, then*

$$\int_F |g(\boldsymbol{x})|\, d\boldsymbol{x} \le \varepsilon.$$

Proof. Consider the set $E_n := \{\boldsymbol{x} \in E : |g(\boldsymbol{x})| \geq n\}$. Then $g_n(\boldsymbol{x}) := |g(\boldsymbol{x})|\chi_{E_n}(\boldsymbol{x}) \to 0$ as $n \to \infty$ and $|g_n(\boldsymbol{x})| \leq |g(\boldsymbol{x})|$ for every $\boldsymbol{x} \in E$. Thus, by the Lebesgue dominated convergence theorem

$$\lim_{n\to\infty} \int_{E_n} |g(\boldsymbol{x})| \, d\boldsymbol{x} = 0.$$

Let n_ε be so large that $\int_{E_{n_\varepsilon}} |g(\boldsymbol{x})| \, d\boldsymbol{x} \leq \varepsilon$ and take $\delta = \varepsilon/n_\varepsilon$. If $F \subseteq E$ is a Lebesgue measurable set with $\mathcal{L}^N(F) \leq \delta$, then

$$\begin{aligned}\int_F |g(\boldsymbol{x})| \, d\boldsymbol{x} &= \int_{F\cap E_{n_\varepsilon}} |g(\boldsymbol{x})| \, d\boldsymbol{x} + \int_{F\setminus E_{n_\varepsilon}} |g(\boldsymbol{x})| \, d\boldsymbol{x} \leq \int_{E_{n_\varepsilon}} |g(\boldsymbol{x})| \, d\boldsymbol{x} + n_\varepsilon \mathcal{L}^N(F) \\ &\leq \varepsilon + n_\varepsilon \varepsilon/n_\varepsilon = 2\varepsilon,\end{aligned}$$

which concludes the proof. ■

Monday, October 10, 2022

We turn to the proof of Theorem 82,

Proof. Given $\varepsilon > 0$, let δ be the number given by in Lemma 83. Then if for (a_k, b_k), $k = 1, \ldots, \ell$, are nonoverlapping intervals with $[a_k, b_k] \subseteq [a, b]$ and

$$\sum_{k=1}^{\ell} (b_k - a_k) \leq \delta,$$

then the set $F = \bigcup_{k=1}^{\ell} (a_k, b_k)$ has Lebesgue measure less than or equal to δ and so

$$\sum_{k=1}^{\ell} |f(b_k) - f(a_k)| \leq \sum_{k=1}^{\ell} \int_{a_k}^{b_k} |g(x)| \, dx = \int_{\bigcup_{k=1}^{\ell}(a_k,b_k)} |g(x)| \, dx \leq \varepsilon.$$

■

We now prove that $f'(x) = g(x)$ for $\mathcal{L}^1$ a.e. $x \in [a, b]$.

Theorem 84 *Let $g : [a, b] \to \mathbb{R}$ be a Lebesgue integrable function and let*

$$f(x) := \int_a^x g(t) \; dt.$$

Then $f'(x) = g(x)$ for $\mathcal{L}^1$ a.e. $x \in [a, b]$.

The proof needs a few lemmas.

Theorem 85 (Fundamental theorem of calculus for Lipschitz continuous functions) *Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a Lipschitz continuous function. Then*

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

for all $a, b \in I$.

Proof. Since f is Lipschitz continuous, there exists $L \geq 0$ such that

$$|f(x) - f(y)| \leq L|x - y|$$

for all $x, y \in I$. Let $x \in I^\circ$ and define

$$f_n(x) := \frac{f(x + \frac{1}{n}) - f(x)}{\frac{1}{n}}.$$

Then $|f_n(x)| \leq nL(x + \frac{1}{n} - x) = L$. Since f is absolutely continuous, it is differentiable for all $x \in I$ except a set of measure zero and so $f_n(x) \to f'(x)$ as $n \to \infty$ for $\mathcal{L}^1$ -a.e. $x \in I$. Hence, by the Lebesgue dominated convergence theorem, for all $a, b \in I$, with $a < b$ and $b \in I^\circ$,

$$\int_a^b f'(x)\ dx = \lim_{n\to\infty} \int_a^b f_n(x)\ dx.$$

On the other hand,

$$\begin{aligned}\int_a^b f_n(x)\ dx &= n \int_a^b \left(f(x + \frac{1}{n}) - f(x) \right)\ dx \\ &= n \left[\int_b^{b+\frac{1}{n}} f(x)\ dx - \int_a^{a+\frac{1}{n}} f(x)\ dx \right] \to f(b) - f(a),\end{aligned}$$

where we used the fact that f is continuous. Hence,

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

with $a < b$ and $b \in I^\circ$. If $\sup I = b \in I$, we can take $b_n = b - \frac{1}{n}$ in what we just proved to get

$$\int_a^{b-\frac{1}{n}} f'(x)\ dx = f(b - \frac{1}{n}) - f(a).$$

By the Lebesgue dominated convergence theorem and the continuity of f,

$$\int_a^b f'(x)\ dx = \lim_{n\to\infty} \int_a^{b-\frac{1}{n}} f'(x)\ dx = \lim_{n\to\infty} f(b - \frac{1}{n}) - f(a) = f(b) - f(a).$$

■

Lemma 86 *Let $g : [a, b] \to \mathbb{R}$ be a Lebesgue integrable function such that*

$$\int_a^x g(t)\ dt = 0$$

for all $x \in [a, b]$. Then $g(x) = 0$ for $\mathcal{L}^1$ a.e. $x \in [a, b]$.

We now turn to the proof of Theorem 85.

Proof. Step 1: Assume that g is bounded, with $|g(x)| \le M$ for all $x \in [a,b]$. Then f is Lipschitz continuous and so by the previous theorem, for every $c \in [a,b]$,

$$\int_a^c f'(x)\ dx = f(c) = \int_a^c g(x)\ dx,$$

that is,

$$\int_a^c (f'(x) - g(x))\ dx = 0$$

for all $c \in [a,b]$. By Lemma 86, it follows that $f'(x) = g(x)$ for $\mathcal{L}^1$ a.e. $x \in [a,b]$.

Step 2: Assume that $g \ge 0$ and define

$$g_n(x) := \begin{cases} g(x) & \text{if } g(x) \le n, \\ 0 & \text{if } g(x) > n. \end{cases}$$

Then

$$f(x) = \int_a^x g(t)\ dt = \int_a^x g_n(t)\ dt + \int_a^x (g(t) - g_n(t))\ dt =: G_n(x) + H_n(x).$$

By Step 1 we have that $G_n'(x) = g_n(x)$ for all $x \in [a,b] \setminus E_n$, where $\mathcal{L}^1(E_n) = 0$. On the other hand, since $g \ge g_n$ we have that H_n is increasing and so, by the Lebesgue differentiation theorem, $H_n'(x) \ge 0$ for all $x \in [a,b] \setminus F_n$, where $\mathcal{L}^1(F_n) = 0$. Hence, since $f = G_n + H_n$, by differentiating, we obtain that

$$f'(x) = G_n'(x) + H_n'(x) = g_n(x) + H_n'(x) \ge g_n(x) + 0$$

for $\mathcal{L}^1$ a.e. $x \in [a,b] \setminus (E_n \cup F_n)$. Since countable union of sets of Lebesgue measure zero have Lebesgue measure zero, we have that $E := \bigcup_n (E_n \cup F_n)$ has Lebesgue measure zero. If $x \in [a,b] \setminus E$, then $f'(x) \ge g_n(x)$ for all n and so, letting $n \to \infty$ we obtain that $f'(x) \ge g(x)$. In turn,

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

On the other hand, by Corollary 71,

$$\int_a^b f'(x)\ dx \le f(b) - f(a),$$

which shows that

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

Hence,

$$\int_a^b (f'(x) - g(x))\ dx = 0,$$

but since $f' \ge g$, it follows that $f'(x) = g(x)$ for $\mathcal{L}^1$ a.e. $x \in [a,b]$.

Step 3: The general case follows by writing $g = g^+ - g^-$ and

$$f(x) = \int_a^x g^+(x)\,dx - \int_a^x g^-(x)\,dx$$

and applying Step 2 to each integral. ■

Wednesday, October 12, 2022

We are now ready to prove the fundamental theorem of calculus for Lebesgue integration.

Theorem 87 (Fundamental Theorem of Calculus) *Let $f : [a,b] \to \mathbb{R}$. Then f is absolutely continuous in $[a,b]$ if and only if f is differentiable $\mathcal{L}^1$-a.e. in $[a,b]$, f' is Lebesgue integrable, and the fundamental theorem of calculus is valid, that is, for all $x, x_0 \in [a,b]$,*

$$f(x) = f(x_0) + \int_{x_0}^x f'(t)\,dt. \tag{31}$$

Proof. Assume that f is differentiable $\mathcal{L}^1$-a.e. in $[a,b]$, f' is Lebesgue integrable, and the fundamental theorem of calculus is valid. Define

$$g(x) := \int_a^x f'(t)\,dt.$$

Then by the previous lemma g is absolutely continuous. In turn, since constant functions are absolutely continuous, it follows that the function $f = f(a) + g$ is absolutely continuous.

Conversely, assume that f is absolutely continuous. Then f has finite pointwise variation and so f is given by the difference of two increasing functions. Since by Corollary 71 the derivative of increasing functions is Lebesgue integrable, it follows that f' is Lebesgue integrable, since difference of Lebesgue integrable functions. In turn, by Theorem 82 the function

$$g(y) := \int_a^y f'(t)\,dt, \quad y \in [a,b],$$

belongs to $AC([a,b])$ with $g'(y) = f'(y)$ for $\mathcal{L}^1$ a.e. $y \in [a,b]$. Since $f - g \in AC([a,b])$ and

$$(f-g)'(y) = f'(y) - f'(y) = 0$$

for $\mathcal{L}^1$ a.e. $y \in [a,b]$, by Theorem 81, we have that $f - g$ is constant in $[a,b]$. Thus, there exists $c \in \mathbb{R}$ such that

$$(f-g)(y) = c$$

for all $y \in [a,b]$, that is,

$$f(y) = c + \int_a^y f'(t)\,dt$$

for all $y \in [a,b]$. ■

Corollary 88 (Integration by parts) *Let $f, g : [a, b] \to \mathbb{R}$ be absolutely continuous. Then*

$$\int_a^b fg'\,dx = -\int_a^b f'g\,dx + g(b)f(b) - f(a)g(a).$$

Proof. I leave as an exercise to check that fg is absolutely continuous. By the fundamental theorem of calculus,

$$\int_a^b (fg' + f'g)\,dx = \int_a^b (fg)'\,dx = g(b)f(b) - f(a)g(a).$$

■

9 The Area Formula

In this section, given $n \in \mathbb{N}$ we denote by $\|\cdot\|_n$ the Euclidean norm in $\mathbb{R}^n$.

Given $E \subseteq \mathbb{R}^k$ and $\boldsymbol{\varphi} : E \to \mathbb{R}^N$, assume that $\boldsymbol{\varphi}$ is differentiable at some point $\boldsymbol{y} \in E$. We recall that the *Jacobian matrix* of $\boldsymbol{\varphi}$ at $\boldsymbol{y}$ is the $N \times k$ matrix given by

$$J_{\boldsymbol{\varphi}}(\boldsymbol{y}) = \nabla\boldsymbol{\varphi}(\boldsymbol{y}) := \begin{pmatrix} \nabla\varphi_1(\boldsymbol{y}) \\ \vdots \\ \nabla\varphi_N(\boldsymbol{y}) \end{pmatrix}. \tag{32}$$

Definition 89 *Given $1 \le k \le N$, a nonempty set $M \subseteq \mathbb{R}^N$ is called a k-*dimensional differential parametrized surface *or* parametrized manifold *if there exists an open set $W \subseteq \mathbb{R}^k$ and a differentiable function $\varphi : W \to \mathbb{R}^N$ such that*

(i) $M = \varphi(W)$,

(ii) $\varphi : W \to M$ is a homeomorphism, that is, it is invertible and continuous together with its inverse $\varphi^{-1} : M \to W$,

(iii) the Jacobian matrix $J_\varphi(\boldsymbol{y})$ has maximum rank k for all $\boldsymbol{y} \in W$.

The function φ is called a chart *or a* system of coordinates *or a* parametrization. *We say that M is of class C^m, $m \in \mathbb{N}$, (respectively, C^∞) if φ is of class C^m (respectively, C^∞).*

Friday, October 14, 2022

Given $k, N \in \mathbb{N}$ and a linear function $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$, the *adjoint* of $\boldsymbol{L}$ is the linear function $\boldsymbol{L}^t : \mathbb{R}^N \to \mathbb{R}^k$ such that

$$\boldsymbol{y} \cdot \boldsymbol{L}^t(\boldsymbol{x}) = \boldsymbol{L}(\boldsymbol{y}) \cdot \boldsymbol{x} \tag{33}$$

for all $\boldsymbol{y} \in \mathbb{R}^k$ and $\boldsymbol{x} \in \mathbb{R}^N$. The matrix representing $\boldsymbol{L}^t$ is simply the transpose of the matrix representing $\boldsymbol{L}$.

Definition 90 *Given $k, N \in \mathbb{N}$ and a linear function $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$, we say that $\boldsymbol{L}$ is* orthogonal *if $\boldsymbol{L}(\boldsymbol{y}_1) \cdot \boldsymbol{L}(\boldsymbol{y}_2) = \boldsymbol{y}_1 \cdot \boldsymbol{y}_2$ for all $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^k$.*

Remark 91 *An orthogonal function $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$ preserves inner products and distances, since*

$$\begin{aligned}\|\boldsymbol{L}(\boldsymbol{y}_1) - \boldsymbol{L}(\boldsymbol{y}_2)\|_N &= \sqrt{\boldsymbol{L}(\boldsymbol{y}_1 - \boldsymbol{y}_2) \cdot \boldsymbol{L}(\boldsymbol{y}_1 - \boldsymbol{y}_2)} \\ &= \sqrt{(\boldsymbol{y}_1 - \boldsymbol{y}_2) \cdot (\boldsymbol{y}_1 - \boldsymbol{y}_2)} = \|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_k.\end{aligned}$$

Thus, $\boldsymbol{L}$ is Lipschitz continuous with Lipschitz constant one and it is injective with $\boldsymbol{L}^{-1} : \boldsymbol{L}(\mathbb{R}^k) \to \mathbb{R}^k$ Lipschitz continuous with Lipschitz constant one. Observe also that $\boldsymbol{L}^t \circ \boldsymbol{L} = I_k$.

Definition 92 *Given $N \in \mathbb{N}$ and a linear function $\boldsymbol{L} : \mathbb{R}^N \to \mathbb{R}^N$, we say that $\boldsymbol{L}$ is*

(i) symmetric *if $\boldsymbol{L} = \boldsymbol{L}^t$,*

(ii) diagonal *if the corresponding matrix is diagonal,*

(iii) positive definite *if $\boldsymbol{L}(\boldsymbol{x}) \cdot \boldsymbol{x} > 0$ for all $\boldsymbol{x} \in \mathbb{R}^N \setminus \{0\}$.*

Theorem 93 (Decomposition) *Let $1 \le k \le N$ and let $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$ be a linear function. Assume that the corresponding matrix has rank k. Then there exist an orthogonal linear function $\boldsymbol{P} : \mathbb{R}^k \to \mathbb{R}^k$, a positive definite, diagonal linear function $\boldsymbol{D} : \mathbb{R}^k \to \mathbb{R}^k$, and an orthogonal linear function $\boldsymbol{Q} : \mathbb{R}^k \to \mathbb{R}^N$ such that*

$$\boldsymbol{L} = \boldsymbol{Q} \circ \boldsymbol{D} \circ \boldsymbol{P}.$$

Proof. We claim that the function $\boldsymbol{L}^t \circ \boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^k$ is symmetric and positive definite. Indeed by (33),

$$\begin{aligned}(\boldsymbol{L}^t \circ \boldsymbol{L})(\boldsymbol{y}_1) \cdot \boldsymbol{y}_2 &= (\boldsymbol{L}^t(\boldsymbol{L}(\boldsymbol{y}_1))) \cdot \boldsymbol{y}_2 \\ &= \boldsymbol{L}(\boldsymbol{y}_1) \cdot \boldsymbol{L}(\boldsymbol{y}_2) = \boldsymbol{y}_1 \cdot (\boldsymbol{L}^t(\boldsymbol{L}(\boldsymbol{y}_2))) = \boldsymbol{y}_1 \cdot (\boldsymbol{L}^t \circ \boldsymbol{L})(\boldsymbol{y}_2)\end{aligned}$$

and

$$(\boldsymbol{L}^t \circ \boldsymbol{L})(\boldsymbol{y}) \cdot \boldsymbol{y} = (\boldsymbol{L}^t(\boldsymbol{L}(\boldsymbol{y}))) \cdot \boldsymbol{y} = \boldsymbol{L}(\boldsymbol{y}) \cdot \boldsymbol{L}(\boldsymbol{y}) = \|\boldsymbol{L}(\boldsymbol{y})\|_N^2 > 0$$

for all $\boldsymbol{y} \in \mathbb{R}^k \setminus \{0\}$, since the matrix corresponding to $\boldsymbol{L}$ has rank k. It follows that the eigenvalues μ_i of $\boldsymbol{L}^t \circ \boldsymbol{L}$ are all positive and that there exists an orthonormal basis $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_k\}$ of eigenvectors. Let $\lambda_i := \sqrt{\mu_i}$. Then $(\boldsymbol{L}^t \circ \boldsymbol{L})(\boldsymbol{b}_i) = \lambda_i^2 \boldsymbol{b}_i$ for all $i = 1, \dots, k$. Let $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$ be the canonical basis in $\mathbb{R}^k$.

Given any two vector spaces of dimension k each with a given basis, there is a linear function between these two vector spaces that maps one basis into the other. Let $\boldsymbol{P} : \mathbb{R}^k \to \mathbb{R}^k$ be the linear function that maps $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_k\}$ into $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$, let $\boldsymbol{D} : \mathbb{R}^k \to \mathbb{R}^k$ be the linear function that maps $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$

into $\{\lambda_1 \boldsymbol{e}_1, \ldots, \lambda_k \boldsymbol{e}_k\}$, and let $\boldsymbol{Q} : \mathbb{R}^k \to \mathbb{R}^N$ be the linear function that maps $\{\lambda_1 \boldsymbol{e}_1, \ldots, \lambda_k \boldsymbol{e}_k\}$ into $\{\boldsymbol{L}(\boldsymbol{b}_1), \ldots, \boldsymbol{L}(\boldsymbol{b}_k)\}$. Note that since $J_{\boldsymbol{L}}$ has rank k, the vector space $\boldsymbol{L}(\mathbb{R}^k) \subseteq \mathbb{R}^N$ has dimension k and $\{\boldsymbol{L}(\boldsymbol{b}_1), \ldots, \boldsymbol{L}(\boldsymbol{b}_k)\}$ is a basis in $\boldsymbol{L}(\mathbb{R}^k)$.

Since $\boldsymbol{P}(\boldsymbol{b}_i) = \boldsymbol{e}_i$, the function $\boldsymbol{P}$ is orthogonal. To verify that $\boldsymbol{Q}$ is orthogonal, note that

$$\boldsymbol{Q}(\boldsymbol{e}_i) \cdot \boldsymbol{Q}(\boldsymbol{e}_j) = \frac{1}{\lambda_i \lambda_j} \boldsymbol{L}(\boldsymbol{b}_i) \cdot \boldsymbol{L}(\boldsymbol{b}_j) = \frac{1}{\lambda_i \lambda_j} \boldsymbol{b}_i \cdot (\boldsymbol{L}^t(\boldsymbol{L}(\boldsymbol{b}_j))) = \frac{\lambda_j^2}{\lambda_i \lambda_j} \boldsymbol{b}_i \cdot \boldsymbol{b}_j = \delta_{i,j}.$$

It remains to show that $\boldsymbol{L} = \boldsymbol{Q} \circ \boldsymbol{D} \circ \boldsymbol{P}$. We have

$$(\boldsymbol{Q} \circ \boldsymbol{D} \circ \boldsymbol{P})(\boldsymbol{b}_i) = (\boldsymbol{Q} \circ \boldsymbol{D})(\boldsymbol{e}_i) = \boldsymbol{Q}(\lambda_i \boldsymbol{e}_i) = \boldsymbol{L}(\boldsymbol{b}_i),$$

and thus the result follows by linearity. ■

Monday, October 24, 2022

Given $E \subseteq \mathbb{R}^k$ and $\boldsymbol{\varphi} : E \to \mathbb{R}^N$, assume that $\boldsymbol{\varphi}$ is differentiable at some point $\boldsymbol{y} \in E$. The *Jacobian* of $\boldsymbol{\varphi}$ at $\boldsymbol{y}$ is the number

$$|||J_{\boldsymbol{\varphi}}(\boldsymbol{y})||| := \sqrt{\det(J_{\boldsymbol{\varphi}}^t(\boldsymbol{y}) J_{\boldsymbol{\varphi}}(\boldsymbol{y}))}, \tag{34}$$

where $J_{\boldsymbol{\varphi}}^t(\boldsymbol{y})$ is the transpose of $J_{\boldsymbol{\varphi}}(\boldsymbol{y})$. Note that when $k = N$,

$$|||J_{\boldsymbol{\varphi}}(\boldsymbol{y})||| = |\det J_{\boldsymbol{\varphi}}(\boldsymbol{y})|. \tag{35}$$

We recall that $\mathcal{H}_o^k$ stands for the k-dimensional Hausdorff outer measure and $\mathcal{H}^k$ is the k-dimensional Hausdorff measure obtained by restricting $\mathcal{H}_o^k$ to the σ-algebra of all $\mathcal{H}_o^k$-measurable sets (see the Carathéodory Theorem 20). We will use the following theorem, which we did not prove (maybe I will prove it at the end of the semester if I have time).

Theorem 94 *Let $\mathcal{H}_o^N$ be the N-th dimensional Hausdorff measure in $\mathbb{R}^N$. Then*

$$\mathcal{H}_o^N = \mathcal{L}_o^N.$$

Exercise 95 *Given $k, N \in \mathbb{N}$ with $1 \le k \le N$ and a linear function $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$, prove that if $E \subseteq \mathbb{R}^k$ is Lebesgue measurable, then $\boldsymbol{L}(E)$ is $\mathcal{H}_o^k$-measurable.*

Proposition 96 *Let $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$ be an orthogonal function. If $E \subseteq \mathbb{R}^k$, then*

$$\mathcal{H}_o^k(\boldsymbol{L}(E)) = \mathcal{L}_o^k(E). \tag{36}$$

Proof. By your homework and Theorem 94,

$$\mathcal{H}_o^k(\boldsymbol{L}(E)) \le \mathcal{H}_o^k(E) = \mathcal{L}_o^k(E).$$

On the other hand, since $\boldsymbol{L}^{-1} : \boldsymbol{L}(\mathbb{R}^k) \to \mathbb{R}^k$ is Lipschitz continuous with Lipschitz constant one,

$$\mathcal{L}_o^k(E) = \mathcal{H}_o^k(E) = \mathcal{H}_o^k(\boldsymbol{L}^{-1}(\boldsymbol{L}(E))) \le \mathcal{H}_o^k(\boldsymbol{L}(E))$$

and thus (36). ■

We are now ready to prove the area formula for injective linear functions.

Theorem 97 (Area formula for linear functions) *Let $1 \le k \le N$ and let $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$ be a linear function. Assume that $J_{\boldsymbol{L}}$ has rank k. Then for every Lebesgue measurable $E \subseteq \mathbb{R}^k$, $\boldsymbol{L}(E)$ is $\mathcal{H}^k_o$-measurable and*

$$\mathcal{H}^k(\boldsymbol{L}(E)) = \int_E |||J_{\boldsymbol{L}}|||\, d\boldsymbol{y} = |||J_{\boldsymbol{L}}|||\mathcal{L}^k(E).$$

Proof. Step 1: Assume first that $k = N$ and that $\boldsymbol{L}$ is a positive definite diagonal linear function $\boldsymbol{D} : \mathbb{R}^k \to \mathbb{R}^k$. Then

$$\boldsymbol{D}(\boldsymbol{y}) = (\lambda_1 y_1, \dots, \lambda_k y_k).$$

Consider a rectangle $R = I_1 \times \cdots \times I_k$. Then by Fubini's theorem and (35),

$$\mathcal{L}^k(\boldsymbol{D}(R)) = \lambda_1 \cdots \lambda_k \mathcal{L}^k(R) = \det J_{\boldsymbol{D}}\mathcal{L}^k(R) = |||J_{\boldsymbol{D}}|||\mathcal{L}^k(R).$$

If $V \subseteq \mathbb{R}^k$ is an open set, then we can write V as a countable union of disjoint rectangles R_n, and since $\boldsymbol{D}$ is injective, the sets $\boldsymbol{D}(R_n)$ are also disjoint and so $\mathcal{L}^k(\boldsymbol{D}(V)) = |||J_{\boldsymbol{D}}|||\mathcal{L}^k(V)$. By approximating Lebesgue measurable sets with open sets we obtain that $\mathcal{L}^k(\boldsymbol{D}(E)) = |||J_{\boldsymbol{D}}|||\mathcal{L}^k(E)$ for every Lebesgue measurable set $E \subseteq \mathbb{R}^k$.
Step 2: Given now $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$ such that $J_{\boldsymbol{L}}$ has rank k, by Theorem 93 there exist an orthogonal linear function $\boldsymbol{P} : \mathbb{R}^k \to \mathbb{R}^k$, a positive definite diagonal linear function $\boldsymbol{D} : \mathbb{R}^k \to \mathbb{R}^k$, and an orthogonal linear function $\boldsymbol{Q} : \mathbb{R}^k \to \mathbb{R}^N$ such that $\boldsymbol{L} = \boldsymbol{Q} \circ \boldsymbol{D} \circ \boldsymbol{P}$. For every Lebesgue measurable set $E \subseteq \mathbb{R}^k$, by Exercise 95, Theorem 94, (36), and Step 1,

$$\begin{aligned}\mathcal{H}^k(\boldsymbol{L}(E)) &= \mathcal{H}^k(\boldsymbol{Q}(\boldsymbol{D}(\boldsymbol{P}(E)))) = \mathcal{H}^k(\boldsymbol{D}(\boldsymbol{P}(E))) \\ &= \mathcal{L}^k(\boldsymbol{D}(\boldsymbol{P}(E))) = |||J_{\boldsymbol{D}}|||\mathcal{L}^k(\boldsymbol{P}(E)) = |||J_{\boldsymbol{D}}|||\mathcal{L}^k(E).\end{aligned} \tag{37}$$

Now since $\boldsymbol{Q}^t \circ \boldsymbol{Q} = I_k$,

$$\begin{aligned}\boldsymbol{L}^t \circ \boldsymbol{L} &= (\boldsymbol{Q} \circ \boldsymbol{D} \circ \boldsymbol{P})^t \circ (\boldsymbol{Q} \circ \boldsymbol{D} \circ \boldsymbol{P}) = \boldsymbol{P}^t \circ \boldsymbol{D}^t \circ \boldsymbol{Q}^t \circ \boldsymbol{Q} \circ \boldsymbol{D} \circ \boldsymbol{P} \\ &= \boldsymbol{P}^t \circ \boldsymbol{D}^t \circ \boldsymbol{D} \circ \boldsymbol{P},\end{aligned}$$

and so

$$\det(\boldsymbol{L}^t \circ \boldsymbol{L}) = \det \boldsymbol{P}^t \det(\boldsymbol{D}^t \circ \boldsymbol{D}) \det \boldsymbol{P}.$$

Since $\det \boldsymbol{P}^t = \det \boldsymbol{P} = \pm 1$ we have $\det(\boldsymbol{L}^t \circ \boldsymbol{L}) = \det(\boldsymbol{D}^t \circ \boldsymbol{D})$. Hence,

$$|||J_{\boldsymbol{L}}||| = \sqrt{\det(\boldsymbol{L}^t \circ \boldsymbol{L})} = \sqrt{\det(\boldsymbol{D}^t \circ \boldsymbol{D})} = |||J_{\boldsymbol{D}}|||,$$

which concludes the proof of the formula. ■

Wednesday, October 26, 2022

Consider the function

$$f(x) = x \sin \frac{1}{x}, \quad x \in (0, 1].$$

Let's study the Hölder continuity of f. We have

$$f'(x) = \sin\frac{1}{x} - x\left(\frac{1}{x^2}\right)\cos\frac{1}{x}$$

and so

$$|f'(x)| \le 1 + \frac{1}{x} \le \frac{2}{x}.$$

Let $0 < x < y \le 1$. We consider two cases. Assume that there exists $n \in \mathbb{N}$ such that $\frac{1}{2\pi(n+1)} \le x < y \le \frac{1}{2\pi n}$. Then by the mean value theorem

$$\begin{aligned}|f(x) - f(y)| = |f'(c)|(y-x) &\le \frac{2}{c}(y-x) \le 4\pi(n+1)(y-x)\\ &\le 8\pi n(y-x)\end{aligned}$$

and so, writing $x = \frac{1}{2\pi n+s}$, $y = \frac{1}{2\pi n+t}$, where $0 \le s, t \le 2\pi$, we have

$$\begin{aligned}\frac{|f(x)-f(y)|}{|x-y|^\alpha} &\le 8\pi n(y-x)^{1-\alpha} = 8\pi n\left(\frac{1}{2\pi n+t} - \frac{1}{2\pi n+s}\right)^{1-\alpha}\\ &= 8\pi n\left(\frac{s-t}{(2\pi n+t)(2\pi n+s)}\right)^{1-\alpha} \le C\frac{n}{n^{2(1-\alpha)}}\end{aligned}$$

and so we want $2(1-\alpha) = 1$, that is, $\alpha = \frac{1}{2}$.

On the other hand if $\frac{1}{2\pi(n+1)} \le x \le \frac{1}{2\pi n} \le \frac{1}{2\pi(\ell+1)} < y \le \frac{1}{2\pi\ell}$, then by the previous step

$$\begin{aligned}|f(x)-f(y)| &\le \left|f(x) - f\left(\frac{1}{2\pi n}\right)\right| + \left|f(y) - f\left(\frac{1}{2\pi(\ell+1)}\right)\right|\\ &\le C\left(\left|x - \frac{1}{2\pi n}\right|^{1/2} + \left|y - \frac{1}{2\pi(\ell+1)}\right|^{1/2}\right)\\ &\le C\,|x-y|^{1/2}.\end{aligned}$$

Next assume that $\alpha > \frac{1}{2}$ and consider two sequences x_n and y_n such that $\sin\frac{1}{x_n} = 0$ and $\sin\frac{1}{y_n} = 1$, that is,

$$x_n = \frac{1}{2\pi n} \text{ and } y_n = \frac{1}{\frac{\pi}{2} + 2n\pi},$$

then

$$\begin{aligned}\frac{|f(x_n)-f(y_n)|}{|x_n-y_n|^\alpha} &= \frac{1}{\frac{\pi}{2}+2n\pi}\frac{1}{\left|\frac{1}{2\pi n} - \frac{1}{\frac{\pi}{2}+2n\pi}\right|^\alpha} = \frac{1}{\frac{\pi}{2}+2n\pi}\frac{1}{\left|\frac{\frac{\pi}{2}}{2\pi n\left(\frac{\pi}{2}+2n\pi\right)}\right|^\alpha}\\ &= \frac{\left[2\pi n\left(\frac{\pi}{2}+2n\pi\right)\right]^\alpha}{\frac{\pi}{2}+2n\pi}\frac{1}{\left|\frac{\pi}{2}\right|^\alpha} \sim n^{2\alpha-1} \to \infty.\end{aligned}$$

The general case is more complicated. Let $a > 0$ and $b > 0$ and consider the function

$$f(x) = x^a \sin \frac{1}{x^b}, \quad x \in (0, 1].$$

We have

$$f'(x) = ax^{a-1} \sin \frac{1}{x^b} - bx^a \left(\frac{1}{x^{b+1}}\right) \cos \frac{1}{x^b}.$$

If $a \geq b+1$, then the derivative of f is bounded and so f is Lipschitz continuous. Since the domain is bounded, it follows that f is Hölder continuous of any exponent less than one. Thus assume that $a < b+1$. Then

$$|f'(x)| \leq \frac{C}{x^{b+1-a}}.$$

Let $0 < x < y \leq 1$. We consider two cases. Assume that there exists $n \in \mathbb{N}$ such that $\frac{1}{[2\pi(n+1)]^{1/b}} \leq x < y \leq \frac{1}{[2\pi n]^{1/b}}$. Then by the mean value theorem

$$|f(x) - f(y)| = |f'(c)|(y - x) \leq \frac{C}{c^{b+1-a}}(y - x) \leq Cn^{(b+1-a)/b}(y - x)$$

and so, writing $x = \frac{1}{(2\pi n+s)^{1/b}}$, $y = \frac{1}{(2\pi n+t)^{1/b}}$, where $0 \leq t < s \leq 2\pi$, we have

$$\frac{|f(x) - f(y)|}{|x - y|^\alpha} \leq Cn^{(b+1-a)/b}(y-x)^{1-\alpha} = Cn^{(b+1-a)/b} \left(\frac{1}{(2\pi n + t)^{1/b}} - \frac{1}{(2\pi n + s)^{1/b}}\right)^{1-\alpha}.$$

So we need to compute the limit

$$\lim_{n\to\infty} n^{(b+1-a)/b} \left(\frac{1}{(2\pi n + t)^{1/b}} - \frac{1}{(2\pi n + s)^{1/b}}\right)^{1-\alpha}.$$

Using the fact that $(1 + z)^\gamma = 1 + \gamma z + o(z)$, we have

$$\frac{1}{(2\pi n + s)^{1/b}} = (2\pi n + s)^{-1/b} = (2\pi n)^{-1/b} \left(1 + \frac{s}{2\pi n}\right)^{-1/b} = (2\pi n)^{-1/b} \left(1 - \frac{1}{b}\frac{s}{2\pi n} + o\left(\frac{1}{n}\right)\right)$$

$$\frac{1}{(2\pi n + t)^{1/b}} = (2\pi n)^{-1/b} \left(1 - \frac{1}{b}\frac{t}{2\pi n} + o\left(\frac{1}{n}\right)\right).$$

Hence,

$$\begin{aligned}\frac{1}{(2\pi n + t)^{1/b}} - \frac{1}{(2\pi n + s)^{1/b}} &= (2\pi n)^{-1/b} \left[1 - \frac{1}{b}\frac{s}{2\pi n} + o\left(\frac{1}{n}\right) - \left(1 - \frac{1}{b}\frac{t}{2\pi n} + o\left(\frac{1}{n}\right)\right)\right] \\ &= \frac{1}{(2\pi n)^{1/b}} \left[\frac{1}{b}\frac{t-s}{2\pi n} + o\left(\frac{1}{n}\right)\right]\end{aligned}$$

and so

$$\begin{aligned}n^{(b+1-a)/b} \left(\frac{1}{(2\pi n + t)^{1/b}} - \frac{1}{(2\pi n + s)^{1/b}}\right)^{1-\alpha} &= n^{(b+1-a)/b} \left(\frac{1}{(2\pi n)^{1/b}} \left[\frac{1}{b}\frac{t-s}{2\pi n} + o\left(\frac{1}{n}\right)\right]\right)^{1-\alpha} \\ &\sim \frac{n^{(b+1-a)/b}}{n^{(1+1/b)(1-\alpha)}}\end{aligned}$$

and so we want $(b+1-a)/b = (1+1/b)(1-\alpha)$, that is $\alpha = \frac{a}{b+1}$.

On the other hand if $\frac{1}{[2\pi(n+1)]^{1/b}} \leq x \leq \frac{1}{[2\pi n]^{1/b}} \leq \frac{1}{[2\pi(\ell+1)]^{1/b}} < y \leq \frac{1}{[2\pi\ell]^{1/b}}$, then by the previous step

$$\begin{aligned}|f(x)-f(y)| &\leq \left|f(x)-f\left(\frac{1}{[2\pi n]^{1/b}}\right)\right| + \left|f(y)-f\left(\frac{1}{2\pi(\ell+1)}\right)\right| \\ &\leq C\left(\left|x-\frac{1}{[2\pi n]^{1/b}}\right|^{\alpha} + \left|y-\frac{1}{2\pi(\ell+1)}\right|^{\alpha}\right) \\ &\leq C\,|x-y|^{\alpha}.\end{aligned}$$

Next assume that $\alpha > \frac{a}{b+1}$. Consider two sequences x_n and y_n such that $\sin\frac{1}{x_n} = 0$ and $\sin\frac{1}{y_n} = 1$, that is,

$$x_n = \frac{1}{(2\pi n)^{1/b}} \text{ and } y_n = \frac{1}{(\frac{\pi}{2}+2n\pi)^{1/b}},$$

then

$$\frac{|f(x_n)-f(y_n)|}{|x_n-y_n|^{\alpha}} = \frac{1}{\left(\frac{\pi}{2}+2n\pi\right)^{a/b}} \frac{1}{\left|\frac{1}{(2\pi n)^{1/b}} - \frac{1}{(\frac{\pi}{2}+2n\pi)^{1/b}}\right|^{\alpha}}$$

Using the fact that $(1+z)^{\gamma} = 1+\gamma z + o(z)$, we have

$$\begin{aligned}\frac{1}{(2\pi n)^{1/b}} - \frac{1}{(\frac{\pi}{2}+2n\pi)^{1/b}} &= \frac{(\frac{\pi}{2}+2n\pi)^{1/b} - (2\pi n)^{1/b}}{(2\pi n)^{1/b}(\frac{\pi}{2}+2n\pi)^{1/b}} = \frac{(2n\pi)^{1/b}\left[(\frac{\pi}{4n\pi}+1)^{1/b}-1\right]}{(2\pi n)^{1/b}(\frac{\pi}{2}+2n\pi)^{1/b}} \\ &= \frac{\frac{1}{b}\frac{\pi}{4n\pi} + o\left(\frac{\pi}{4n\pi}\right)}{(\frac{\pi}{2}+2n\pi)^{1/b}} \sim \frac{1}{n^{1+1/b}}.\end{aligned}$$

Hence,

$$\frac{|f(x_n)-f(y_n)|}{|x_n-y_n|^{\alpha}} = \frac{1}{\left(\frac{\pi}{2}+2n\pi\right)^{a/b}} \frac{1}{\left|\frac{1}{(2\pi n)^{1/b}} - \frac{1}{(\frac{\pi}{2}+2n\pi)^{1/b}}\right|^{\alpha}} \sim \frac{n^{(1+1/b)\alpha}}{n^{a/b}} = n^{(1+1/b)\alpha - a/b} \to \infty$$

provided $(1+1/b)\alpha - a/b = \frac{b+1}{b}\alpha - \frac{a}{b} = \frac{b+1}{b}\left(\alpha - \frac{a}{b+1}\right) > 0$, that is $\alpha > \frac{a}{b+1}$.

Monday, October 31, 2022

Next we extend the area formula to C^1 functions.

Theorem 98 (Area formula) *Let $1 \leq k \leq N$, let $V \subseteq \mathbb{R}^k$ be an open set and let $\boldsymbol{\varphi} : V \to \mathbb{R}^N$ be a function of class C^1 such that $J_{\boldsymbol{\varphi}}(\boldsymbol{y})$ has rank k for every $\boldsymbol{y} \in V$. Let $E \subseteq V$ be a Lebesgue measurable set and assume that $\boldsymbol{\varphi}$ is injective in E. Then*

$$\mathcal{H}^k(\boldsymbol{\varphi}(E)) = \int_E |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y}.$$

Exercise 99 *Let $\boldsymbol{f} : [a,b] \to \mathbb{R}^N$ be continuous. Prove that*

$$\left\|\int_a^b \boldsymbol{f}(x)\,dx\right\| \leq \int_a^b \|\boldsymbol{f}(x)\|\,dx.$$

In what follows given an $N\times k$ matrix A we define its norm as

$$\|A\|_{N\times k} := \sup\left\{\frac{\|A\boldsymbol{y}\|_N}{\|\boldsymbol{y}\|_k} : \boldsymbol{y}\in\mathbb{R}^k\setminus\{\boldsymbol{0}\}\right\}.$$

Note that

$$\|A\boldsymbol{y}\|_N \le \|A\|_{N\times k}\|\boldsymbol{y}\|_k \quad \text{for all } \boldsymbol{y}\in\mathbb{R}^k. \tag{38}$$

We divide the proof in a few lemmas.

Lemma 100 *Let $1\le k\le N$, let $V\subseteq\mathbb{R}^k$ be an open set and let $\boldsymbol{\varphi}: V\to\mathbb{R}^N$ be a function of class C^1 and let $\boldsymbol{y}_0\in V$. Assume that $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$ has rank k. Then for every $\varepsilon>0$ there exists $\delta>0$ such that*

$$(1-\varepsilon)\|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N \le \|\boldsymbol{\varphi}(\boldsymbol{y}_1)-\boldsymbol{\varphi}(\boldsymbol{y}_2)\|_N \le (1+\varepsilon)\|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N$$

for all $\boldsymbol{y}_1,\boldsymbol{y}_2\in B_k(\boldsymbol{y}_0,\delta)\subseteq V$, where $\boldsymbol{L}(\boldsymbol{y}) := J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)\boldsymbol{y}^t$, $\boldsymbol{y}\in\mathbb{R}^k$.

Proof. Since $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$ has rank k, $\boldsymbol{L}$ is injective, and so $\boldsymbol{L}(\boldsymbol{y})\neq\boldsymbol{0}$ for all $\boldsymbol{y}\in\mathbb{R}^k\setminus\{\boldsymbol{0}\}$. Define $g(\boldsymbol{y}) := \|\boldsymbol{L}(\boldsymbol{y})\|_N$. By Weierstrass theorem, there exists

$$\min_{\boldsymbol{y}\in\partial B_k(\boldsymbol{0},1)} g(\boldsymbol{y}) = g(\boldsymbol{y}_0) = c > 0.$$

Hence, for $\boldsymbol{y}\in\mathbb{R}^k\setminus\{\boldsymbol{0}\}$,

$$\left\|\boldsymbol{L}\left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_k}\right)\right\|_N \ge c.$$

It follows that

$$\|\boldsymbol{L}(\boldsymbol{y})\|_N \ge c\|\boldsymbol{y}\|_k \quad \text{for all } \boldsymbol{y}\in\mathbb{R}^k. \tag{39}$$

Since $\boldsymbol{\varphi}$ is of class C^1 there exists $\delta>0$ such that

$$\|J_{\boldsymbol{\varphi}}(\boldsymbol{y}) - J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)\|_{N\times k} \le c\varepsilon \tag{40}$$

for every $\boldsymbol{y}\in B_k(\boldsymbol{y}_0,\delta)$. Let $\boldsymbol{y}_1,\boldsymbol{y}_2\in B_k(\boldsymbol{y}_0,\delta)$. By the fundamental theorem of calculus applied to the function

$$h(t) := \boldsymbol{\varphi}(\boldsymbol{y}_1 t + (1-t)\boldsymbol{y}_2) - \boldsymbol{L}(\boldsymbol{y}_1 t + (1-t)\boldsymbol{y}_2)$$

we have

$$\boldsymbol{\varphi}(\boldsymbol{y}_1)-\boldsymbol{\varphi}(\boldsymbol{y}_2)-\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2) = \int_0^1 (J_{\boldsymbol{\varphi}}(\boldsymbol{y}_1 t+(1-t)\boldsymbol{y}_2) - J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0))(\boldsymbol{y}_1-\boldsymbol{y}_2)^t\,dt.$$

Hence, by (38), (39), (40), and Exercise 99,

$$\begin{aligned}\|\boldsymbol{\varphi}(\boldsymbol{y}_1)-\boldsymbol{\varphi}(\boldsymbol{y}_2)\|_N &\le \|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N + \|\boldsymbol{\varphi}(\boldsymbol{y}_1)-\boldsymbol{\varphi}(\boldsymbol{y}_2)-\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N\\ &\le \|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N + c\varepsilon\|\boldsymbol{y}_1-\boldsymbol{y}_2\|_k \le (1+\varepsilon)\|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N,\end{aligned}$$

while

$$\begin{aligned}\|\boldsymbol{\varphi}(\boldsymbol{y}_1)-\boldsymbol{\varphi}(\boldsymbol{y}_2)\|_N &\ge \|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N - \|\boldsymbol{\varphi}(\boldsymbol{y}_1)-\boldsymbol{\varphi}(\boldsymbol{y}_2)-\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N\\ &\ge \|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N - c\varepsilon\|\boldsymbol{y}_1-\boldsymbol{y}_2\|_k \ge (1-\varepsilon)\|\boldsymbol{L}(\boldsymbol{y}_1-\boldsymbol{y}_2)\|_N,\end{aligned}$$

which completes the proof. ■

Lemma 101 *Let $1 \leq k \leq N$, let $V \subseteq \mathbb{R}^k$ be an open set, let $\boldsymbol{\varphi} : V \to \mathbb{R}^N$ be a function of class C^1, and let $\boldsymbol{y}_0 \in V$. Assume that $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$ has rank k. Then for every $0 < \varepsilon < 1$ there exists $\delta > 0$ such that for every Lebesgue measurable set $E \subseteq B_k(\boldsymbol{y}_0, \delta) \subseteq V$, $\boldsymbol{\varphi}(E)$ is $\mathcal{H}^k_o$-measurable and*

$$(1-\varepsilon)^{k+1} \int_E |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y} \leq \mathcal{H}^k(\boldsymbol{\varphi}(E)) \leq (1+\varepsilon)^{k+1} \int_E |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y}.$$

Proof. Let $\boldsymbol{L}$ be as in the previous lemma. Since $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$ has rank k, the linear function $\boldsymbol{L} : \mathbb{R}^k \to \mathbb{R}^N$ is injective, and so there exists $\boldsymbol{L}^{-1} : \boldsymbol{L}(\mathbb{R}^k) \to \mathbb{R}^k$. Given $0 < \varepsilon < 1$, let $\delta > 0$ be so small that the conclusions of the previous lemma hold and also so that

$$(1+\varepsilon)^{-1} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})||| \leq |||J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)||| \leq (1+\varepsilon) |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})||| \tag{41}$$

for all $\boldsymbol{y} \in B_k(\boldsymbol{y}_0, \delta)$, where we used the fact that $\boldsymbol{\varphi}$ is of class C^1. Since by the previous lemma,

$$(1-\varepsilon)\|\boldsymbol{L}(\boldsymbol{y}_1 - \boldsymbol{y}_2)\|_N \leq \|\boldsymbol{\varphi}(\boldsymbol{y}_1) - \boldsymbol{\varphi}(\boldsymbol{y}_2)\|_N \leq (1+\varepsilon)\|\boldsymbol{L}(\boldsymbol{y}_1 - \boldsymbol{y}_2)\|_N \tag{42}$$

for all $\boldsymbol{y}_1, \boldsymbol{y}_2 \in B_k(\boldsymbol{y}_0, \delta) \subseteq V$, taking $\boldsymbol{y}_1 = \boldsymbol{L}^{-1}(\boldsymbol{x}_1)$ and $\boldsymbol{y}_1 = \boldsymbol{L}^{-1}(\boldsymbol{x}_2)$ we get

$$\|(\boldsymbol{\varphi} \circ \boldsymbol{L}^{-1})(\boldsymbol{x}_1) - (\boldsymbol{\varphi} \circ \boldsymbol{L}^{-1})(\boldsymbol{x}_2)\|_N \leq (1+\varepsilon)\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_N$$

for all $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in \boldsymbol{L}(B_k(\boldsymbol{y}_0, \delta))$. ■

Wednesday, November 2, 2022

Proof. Thus $\boldsymbol{\varphi} \circ \boldsymbol{L}^{-1}$ is Lipschitz continuous with Lipschitz constant less than or equal to $1+\varepsilon$. It follows by your homework, the area formula for $\boldsymbol{L}$, and (41),

$$\begin{aligned}\mathcal{H}^k_o(\boldsymbol{\varphi}(E)) &= \mathcal{H}^k_o((\boldsymbol{\varphi} \circ \boldsymbol{L}^{-1})(\boldsymbol{L}(E))) \leq (1+\varepsilon)^k \mathcal{H}^k_o(\boldsymbol{L}(E)) \\ &= (1+\varepsilon)^k \int_E |||J_{\boldsymbol{L}}|||\, d\boldsymbol{y} \leq (1+\varepsilon)^{k+1} \int_E |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y}.\end{aligned} \tag{43}$$

Similarly, by (42) and the fact that $\boldsymbol{L}$ is injective it follows that $\boldsymbol{\varphi}$ is injective, and so taking $\boldsymbol{y}_1 = \boldsymbol{\varphi}^{-1}(\boldsymbol{x}_1)$ and $\boldsymbol{y}_1 = \boldsymbol{\varphi}^{-1}(\boldsymbol{x}_2)$ we get

$$\|(\boldsymbol{L} \circ \boldsymbol{\varphi}^{-1})(\boldsymbol{x}_1) - (\boldsymbol{L} \circ \boldsymbol{\varphi}^{-1})(\boldsymbol{x}_2)\|_N \leq (1-\varepsilon)^{-1}\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_N$$

for all $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in \boldsymbol{\varphi}(B_k(\boldsymbol{y}_0, \delta))$. Thus $\boldsymbol{L} \circ \boldsymbol{\varphi}^{-1}$ is Lipschitz continuous with Lipschitz constant less than or equal to $(1-\varepsilon)^{-1}$. It follows by your homework, the area formula for $\boldsymbol{L}$, and (41),

$$\begin{aligned}(1+\varepsilon)^{-1} \int_E |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y} &\leq \int_E |||J_{\boldsymbol{L}}|||\, d\boldsymbol{y} = \mathcal{H}^k_o(\boldsymbol{L}(E)) \\ &= \mathcal{H}^k_o((\boldsymbol{L} \circ \boldsymbol{\varphi}^{-1})(\boldsymbol{\varphi}(E))) \leq (1-\varepsilon)^{-k} \mathcal{H}^k_o(\boldsymbol{\varphi}(E)),\end{aligned}$$

which gives the other inequality since $(1-\varepsilon) \leq (1+\varepsilon)^{-1}$.

We leave as an exercise to prove that $\boldsymbol{\varphi}(E)$ is $\mathcal{H}_o^k$-measurable. ■

We turn to the proof of the area formula

Proof. Fix $\varepsilon > 0$ and cover V with countably many balls $B_i \subseteq V$ such that for every $E \subseteq B_i$ the previous lemma apply. Given a Lebesgue measurable set $E \subseteq V$, define inductively, $E_1 := E \cap B_1$, $E_i := (E \cap B_i) \setminus \bigcup_{j=1}^{i-1} B_j$. Then the sets E_i are disjoint and their union is E. By the previous lemma applied to each E_i we get

$$(1-\varepsilon)^{k+1} \int_{E_i} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y} \le \mathcal{H}^k(\boldsymbol{\varphi}(E_i)) \le (1+\varepsilon)^{k+1} \int_{E_i} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y}.$$

Summing over i and using the fact that $\boldsymbol{\varphi}$ is injective in E gives

$$(1-\varepsilon)^{k+1} \int_{E} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y} \le \mathcal{H}^k(\boldsymbol{\varphi}(E)) \le (1+\varepsilon)^{k+1} \int_{E} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y}.$$

We now let $\varepsilon \to 0^+$. ■

Example 102 *Let $I \subseteq \mathbb{R}$ be an open interval and let $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ be a function of class C^1. Assume that $\boldsymbol{\varphi}$ is injective and that $\boldsymbol{\varphi}'(t) \neq \mathbf{0}$ for every $t \in I$. Then the set $M = \boldsymbol{\varphi}(I)$ is a 1-dimensional manifold of class C^1 (why?). Since*

$$\det\left(\boldsymbol{\varphi}'(t))^T \boldsymbol{\varphi}'(t)\right) = \|\boldsymbol{\varphi}'(t)\|^2,$$

we have that the length of M is given by

$$\mathcal{H}^1(M) = \int_I \|\boldsymbol{\varphi}'(t)\|\, dt.$$

Moreover, for every Lebesgue measurable set $E \subseteq I$,

$$\mathcal{H}^1(\boldsymbol{\varphi}(E)) = \int_E \|\boldsymbol{\varphi}'(t)\|\, dt.$$

Example 103 *Given an open set $V \subseteq \mathbb{R}^N$ and a function $f : V \to \mathbb{R}$ of class C^1, consider the graph of f,*

$$\operatorname{Gr} f := \{(\boldsymbol{x}, t) \in V \times \mathbb{R} : \ t = f(\boldsymbol{y})\} \subseteq \mathbb{R}^{N+1}.$$

We claim that $\operatorname{Gr} f$ is an N-dimensional manifold of class C^1. Indeed, a chart is given by the function $\varphi : V \to \mathbb{R}^{N+1}$ defined as $\varphi(\boldsymbol{x}) := (\boldsymbol{x}, f(\boldsymbol{x}))$. Then,

$$J_\varphi(\boldsymbol{x}) = \begin{pmatrix} I_N \\ \nabla f(\boldsymbol{x}) \end{pmatrix},$$

which has rank N. Note that φ is one-to-one and that $\varphi(V) = \operatorname{Gr} f$. Hence, there exists $\varphi^{-1} : \operatorname{Gr} f \to V$. Moreover, φ^{-1} is continuous, since the projection

$$\begin{aligned} \Pi : \mathbb{R}^{N+1} &\to \mathbb{R}^N \\ (\boldsymbol{x}, t) &\mapsto \boldsymbol{x} \end{aligned}$$

is of class C^∞ *and* φ^{-1} *is given by the restriction of* Π *to* $\operatorname{Gr} f$*. Hence,* $\operatorname{Gr} f$ *is an* N*-dimensional manifold of class* C^1*. Moreover,*

$$\sqrt{\det(J_\varphi(\boldsymbol{x}))^T J_\varphi(\boldsymbol{x})} = \sqrt{1+\sum_{i=1}^N \left(\frac{\partial f}{\partial x_i}(x)\right)^2} = \sqrt{1+\|\nabla f(x)\|_N^2}$$

Hence, the surface area of $\operatorname{Gr} f$ *is*

$$\mathcal{H}^N(\operatorname{Gr} f) = \int_V \sqrt{1+\|\nabla f(\boldsymbol{x})\|_N^2}\, d\boldsymbol{x},$$

and for every Lebesgue measurable set $E \subseteq V$*,*

$$\mathcal{H}^N(\varphi(E)) = \int_E \sqrt{1+\|\nabla f(\boldsymbol{x})\|_N^2}\, d\boldsymbol{x}.$$

Friday, November 4, 2022

Theorem 104 (Area formula, general case) *Let* $1 \le k \le N$*, let* $V \subseteq \mathbb{R}^k$ *be an open set and let* $\boldsymbol{\varphi} : V \to \mathbb{R}^N$ *be a function of class* C^1*. Let* $E \subseteq V$ *be a Lebesgue measurable set and assume that* $\boldsymbol{\varphi}$ *is injective in* E*. Then*

$$\mathcal{H}^k(\boldsymbol{\varphi}(E)) = \int_E |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\, d\boldsymbol{y}.$$

Theorem 105 (Cauchy–Binet formula) *Let* $1 \le k \le N$*, let* A *be a* $N \times k$ *matrix, and let* B *be a* $k \times N$ *matrix. Then*

$$\det BA = \sum_{\boldsymbol{\alpha}\in\Lambda_{N,k}} \det(a_{\alpha_i,j})_{i,j=1}^k \det(b_{i,\alpha_j})_{i,j=1}^k, \tag{44}$$

where

$$\Lambda_{N,k} := \{\boldsymbol{\alpha} \in \mathbb{N}^k : 1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_k \le N\}.$$

In particular,

$$\det A^t A = \sum_{\boldsymbol{\alpha}\in\Lambda_{N,k}} (\det(a_{\alpha_i,j})_{i,j=1}^k)^2.$$

Proof. Not done in class. We only give a sketch of the proof. Using the fact that for square matrices the determinant of the product of two matrices is given by the product of the determinants of the two matrices, we have

$$\begin{aligned}\det(I+AB) &= \det\begin{pmatrix} I & A \\ 0 & I\end{pmatrix}\begin{pmatrix} I & -A \\ B & I\end{pmatrix} = \det\begin{pmatrix} I & A \\ 0 & I\end{pmatrix}\det\begin{pmatrix} I & -A \\ B & I\end{pmatrix} \\ &= \det\begin{pmatrix} I & -A \\ B & I\end{pmatrix}\begin{pmatrix} I & A \\ 0 & I\end{pmatrix} = \det(I+BA),\end{aligned}$$

where I and 0 are identity matrices and zero matrices of whatever dimension is needed to make sense of the previous expressions. The identity

$$\det(I_N + AB) = \det(I_k + BA)$$

is called the *Sylvester determinant identity.* If we now let $t \in \mathbb{R}$ and rescale everything, we obtain

$$\det(tI_N + AB) = t^{N-k} \det(tI_k + BA).$$

Since both the left-end and right-end sides are polynomials of degree N in t, by equating the coefficients of the t^{N-k} terms we get (44). ■

Remark 106 *The last formula shows that to compute* $\det A^t A$ *one should consider all the* $k \times k$ *submatrices of* A*, compute their determinant and take the sum of their squares.*

Example 107 *Consider a 2-dimensional parametrized surface of class* C^1 *in* $\mathbb{R}^3$ *parametrized by* $\boldsymbol{\varphi} : V \to \mathbb{R}^3$*, where* $V \subseteq \mathbb{R}^2$*. Then*

$$J_{\boldsymbol{\varphi}}(\boldsymbol{y}) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1} & \frac{\partial \varphi_1}{\partial y_2} \\ \frac{\partial \varphi_2}{\partial y_1} & \frac{\partial \varphi_2}{\partial y_2} \\ \frac{\partial \varphi_3}{\partial y_1} & \frac{\partial \varphi_3}{\partial y_2} \end{pmatrix}$$

and so

$$\det(J_{\boldsymbol{\varphi}}(\boldsymbol{y}))^t J_{\boldsymbol{\varphi}}(\boldsymbol{y}) = \det{}^2 \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1} & \frac{\partial \varphi_1}{\partial y_2} \\ \frac{\partial \varphi_2}{\partial y_1} & \frac{\partial \varphi_2}{\partial y_2} \end{pmatrix} + \det{}^2 \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1} & \frac{\partial \varphi_1}{\partial y_2} \\ \frac{\partial \varphi_3}{\partial y_1} & \frac{\partial \varphi_3}{\partial y_2} \end{pmatrix} + \det{}^2 \begin{pmatrix} \frac{\partial \varphi_2}{\partial y_1} & \frac{\partial \varphi_2}{\partial y_2} \\ \frac{\partial \varphi_3}{\partial y_1} & \frac{\partial \varphi_3}{\partial y_2} \end{pmatrix}.$$

We can now prove the area formula in the general case.

Proof. Step 1: Let $\Sigma := \{\boldsymbol{y} \in V : J_{\boldsymbol{\varphi}}(\boldsymbol{y}) \text{ has rank less than } k\}$. We claim that $\mathcal{H}^k(\boldsymbol{\varphi}(\Sigma)) = 0$. To see this, assume first that V is bounded and that $J_{\boldsymbol{\varphi}}$ is bounded in V, say, $\|J_{\boldsymbol{\varphi}}(\boldsymbol{y})\| \le M$ for all $\boldsymbol{y} \in V$. For $\varepsilon > 0$ consider the function $\boldsymbol{\varphi}_\varepsilon : V \to \mathbb{R}^N \times \mathbb{R}^k$ given by $\boldsymbol{\varphi}_\varepsilon(\boldsymbol{y}) := (\boldsymbol{\varphi}(\boldsymbol{y}), \varepsilon \boldsymbol{y})$. Then

$$J_{\boldsymbol{\varphi}_\varepsilon}(\boldsymbol{y}) = \begin{pmatrix} J_{\boldsymbol{\varphi}}(\boldsymbol{y}) \\ \varepsilon I_k \end{pmatrix},$$

and so $J_{\boldsymbol{\varphi}_\varepsilon}(\boldsymbol{y})$ has rank k for every $\boldsymbol{y} \in V$. Then by the Cauchy–Binet formula (see Theorem 105),

$$\begin{aligned} |||J_{\boldsymbol{\varphi}_\varepsilon}(\boldsymbol{y})|||^2 &= \sum_{\alpha \in \Lambda_{N+k,k}} \left(\det \frac{\partial (\boldsymbol{\varphi}_\varepsilon)_\alpha}{\partial \boldsymbol{y}} (\boldsymbol{y}) \right)^2 \\ &\le |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||^2 + c(1 + M^2)\varepsilon^2 \end{aligned}$$

for some constant $c > 0$. If particular, if $\boldsymbol{y} \in \Sigma$, then

$$|||J_{\boldsymbol{\varphi}_\varepsilon}(\boldsymbol{y})|||^2 \le c\varepsilon^2, \tag{45}$$

where as usual the constant c changes from line to line. Since $\boldsymbol{\varphi} = \Pi \circ \boldsymbol{\varphi}_\varepsilon$, where $\Pi : \mathbb{R}^N \times \mathbb{R}^d \to \mathbb{R}^N$ is the projection operator given by $\Pi(\boldsymbol{x}, \boldsymbol{y}) := \boldsymbol{x}$ and since Π

is Lipschitz continuous with Lipschitz constant one, it follows from Proposition **??**, the area formula, and (45),

$$\begin{aligned}\mathcal{H}^k(\boldsymbol{\varphi}(\Sigma)) &= \mathcal{H}^k(\Pi(\boldsymbol{\varphi}_\varepsilon(\Sigma))) \le 1\mathcal{H}^k(\boldsymbol{\varphi}_\varepsilon(\Sigma)) \\ &= \int_\Sigma |||J_{\boldsymbol{\varphi}_\varepsilon}(\boldsymbol{y})|||\,d\boldsymbol{y} \le c\varepsilon\mathcal{L}^k(V).\end{aligned}$$

Letting $\varepsilon \to 0$ gives $\mathcal{H}^k(\boldsymbol{\varphi}(\Sigma)) = 0$.

The general case in which V and $J_{\boldsymbol{\varphi}}$ are not bounded follows by writing V as an increasing sequence of open bounded sets V_n with $\overline{V_n} \subseteq V_{n+1} \subseteq V$ for all n and by applying what we just did in each set V_n.

Step 2: Since $V \setminus \Sigma$ is open, we can apply the special case of the area formula (Theorem 98) to obtain that

$$\mathcal{H}^k(\boldsymbol{\varphi}(E \setminus \Sigma)) = \int_{E\setminus\Sigma} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y}.$$

On the other hand, by the previous lemma, $\mathcal{H}^k(\boldsymbol{\varphi}(\Sigma)) = 0$, and so

$$\mathcal{H}^k(\boldsymbol{\varphi}(E)) = \mathcal{H}^k(\boldsymbol{\varphi}(E\setminus\Sigma)\cup\boldsymbol{\varphi}(E\cap\Sigma)) = \mathcal{H}^k(\boldsymbol{\varphi}(E\setminus\Sigma)) = \int_{E\setminus\Sigma} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y} = \int_E |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y},$$

where in the last equality we used the fact that $|||J_{\boldsymbol{\varphi}}(\boldsymbol{y})||| = 0$ for all $\boldsymbol{y} \in \Sigma$. ■

As a consequence of the area formula we have the following change of variables formula for surface integrals.

Theorem 108 *Let $1 \le k \le N$, let $V \subseteq \mathbb{R}^k$ be an open set and let $\boldsymbol{\varphi} : V \to \mathbb{R}^N$ be a function of class C^1. Let $E \subseteq \boldsymbol{\varphi}(V)$ be a Borel set and let $f : E \to \mathbb{R}$ be a Borel function, which is either $\mathcal{H}^k$ integrable or has a sign. Assume that $\boldsymbol{\varphi}$ is injective in $\boldsymbol{\varphi}^{-1}(E)$. Then*

$$\int_E f(\boldsymbol{x})\,d\mathcal{H}^k(\boldsymbol{x}) = \int_{\boldsymbol{\varphi}^{-1}(E)} f(\boldsymbol{\varphi}(\boldsymbol{y}))|||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y}. \tag{46}$$

Proof. Assume first that $f = \chi_G$, where $G \subseteq E$ is a Borel set. Then $\boldsymbol{\varphi}^{-1}(G)$ is a Borel set and so by Theorem 104,

$$\begin{aligned}\int_E f(\boldsymbol{x})\,d\mathcal{H}^k(\boldsymbol{x}) &= \mathcal{H}^k(G) = \mathcal{H}^k(\boldsymbol{\varphi}(\boldsymbol{\varphi}^{-1}(G))) \\ &= \int_{\boldsymbol{\varphi}^{-1}(G)} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y} = \int_{\boldsymbol{\varphi}^{-1}(E)} \chi_G(\boldsymbol{\varphi}(\boldsymbol{y}))|||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y}.\end{aligned}$$

Next take f to be a simple function, $f = \sum_{i=1}^n c_i\chi_{G_i}$, where the Borel sets G_i are disjoint. Then by what we just proved and the linearity of integrals

$$\begin{aligned}\int_E f(\boldsymbol{x})\,d\mathcal{H}^k(\boldsymbol{x}) &= \sum_{i=1}^n c_i\mathcal{H}^k(G_i) = \sum_{i=1}^n c_i \int_{\boldsymbol{\varphi}^{-1}(G_i)} |||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y} \\ &= \int_{\boldsymbol{\varphi}^{-1}(E)} f(\boldsymbol{\varphi}(\boldsymbol{y}))|||J_{\boldsymbol{\varphi}}(\boldsymbol{y})|||\,d\boldsymbol{y}.\end{aligned}$$

For a nonnegative Borel function f, construct an increasing sequence of Borel simple functions converging pointwise to f and apply the Lebesgue monotone convergence theorem on both sides.

Finally, if the Borel function $f : E \to \mathbb{R}$ is $\mathcal{H}^k$ integrable, then as usual we can write $f = f^+ - f^-$, apply (46) to f^+ and f^-, and use the linearity of integrals to deduce (46) for f. ■

Exercise 109 *Prove that the previous theorem continues to hold if we assume that $\boldsymbol{\varphi}$ is injective in $\boldsymbol{\varphi}^{-1}(E) \setminus E_0$, where $\mathcal{L}^k(E_0) = 0$.*

Since $\mathcal{H}^N = \mathcal{L}^N$ (see Theorem 94) in the case $k = N$ we obtain the classical change of variables for Lebesgue integration.

Corollary 110 (Change of variables) *Let $\Omega \subseteq \mathbb{R}^N$ be an open set and let $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^N$ be a function of class C^1. Let $E \subseteq \boldsymbol{\varphi}(\Omega)$ be a Lebesgue measurable set and let $f : E \to \mathbb{R}$ be a Lebesgue measurable function, which is either Lebesgue integrable or has a sign. Assume that $\boldsymbol{\varphi}$ is injective in $\boldsymbol{\varphi}^{-1}(E)$. Then*

$$\int_E f(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\boldsymbol{\varphi}^{-1}(E)} f(\boldsymbol{\varphi}(\boldsymbol{y}))|\det J_{\boldsymbol{\varphi}}(\boldsymbol{y})|\, d\boldsymbol{y}.$$

Monday, November 7, 2022

We now consider some important examples.

Example 111 *Let M be a 1-dimensional parametrized manifold of class C^1 and $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ be a parametrization. Then*

$$\det(\boldsymbol{\varphi}'(t))^T \boldsymbol{\varphi}'(t) = \|\boldsymbol{\varphi}'(t)\|.$$

If $E \subseteq \boldsymbol{\varphi}(I)$ is a Borel set and $f : E \to \mathbb{R}$ a Borel function, which is either $\mathcal{H}^1$-integrable or has a sign, then

$$\int_E f(\boldsymbol{x})\, d\mathcal{H}^1(\boldsymbol{x}) = \int_{\boldsymbol{\varphi}^{-1}(E)} f(\boldsymbol{\varphi}(\boldsymbol{y}))\|\boldsymbol{\varphi}'(t)\|\, dt.$$

Example 112 *Given an open set $V \subseteq \mathbb{R}^k$ and a function $f : V \to \mathbb{R}$ of class C^1, consider the graph of f,*

$$\operatorname{Gr} f := \{(\boldsymbol{y}, t) \in V \times \mathbb{R} : \, t = f(\boldsymbol{y})\}.$$

We have seen that $\operatorname{Gr} f$ is an k-dimensional surface of class C^1 and that a chart is given by the function $\boldsymbol{\varphi} : V \to \mathbb{R}^{k+1}$ given by $\boldsymbol{\varphi}(\boldsymbol{y}) := (\boldsymbol{y}, f(\boldsymbol{y}))$. Moreover,

$$J_{\boldsymbol{\varphi}}(\boldsymbol{y}) = \begin{pmatrix} I_k \\ \nabla f(\boldsymbol{y}) \end{pmatrix},$$

which has rank k. Hence,

$$\sqrt{\det(J_{\boldsymbol{\varphi}}(\boldsymbol{y}))^T J_{\boldsymbol{\varphi}}(\boldsymbol{y})} = \sqrt{1 + \sum_{i=1}^k \left(\frac{\partial f}{\partial y_i}(\boldsymbol{y})\right)^2} = \sqrt{1 + \|\nabla f(\boldsymbol{y})\|_k^2}$$

If $E \subseteq \boldsymbol{\varphi}(V)$ is a Borel set and $g : E \to \mathbb{R}$ a Borel function, which is either $\mathcal{H}^k$-integrable or has a sign, then

$$\int_E g \, d\mathcal{H}^k = \int_{\boldsymbol{\varphi}^{-1}(E)} g\left(\boldsymbol{y}, f\left(\boldsymbol{y}\right)\right) \sqrt{1 + \left\|\nabla f\left(\boldsymbol{y}\right)\right\|_k^2} \, d\boldsymbol{y}. \tag{47}$$

In Theorem 108 we have seen how to compute

$$\int_E f \, d\mathcal{H}^k$$

in the case in which $E \subseteq M$, where M is a k-th dimensional parametrized manifold of class C^1. In many examples, we have a more general situation, where

$$E \subseteq E_0 \cup \bigcup_{n=1}^{\infty} M_n, \tag{48}$$

where $\mathcal{H}^k(E_0) = 0$ and M_n are k-th dimensional parametrized manifolds of class C^1 parametrized by $\boldsymbol{\varphi}_n : V_n \to \mathbb{R}^N$. Define $E_1 := E \cap M_1$, $E_n := E \cap M_n \setminus \bigcup_{k=1}^{n-1} E_k$. Then the sets E_n are disjoint, and so, if either $f \geq 0$ or f is $\mathcal{H}^k$ integrable, we can write

$$\int_E f \, d\mathcal{H}^k = \sum_{n=1}^{\infty} \int_{E_n} f \, d\mathcal{H}^k.$$

Now we can apply Theorem 108 in E_n to write

$$\int_E f \, d\mathcal{H}^k = \sum_{n=1}^{\infty} \int_{E_n} f \, d\mathcal{H}^k = \sum_{n=1}^{\infty} \int_{\boldsymbol{\varphi}_n^{-1}(E_n)} f(\boldsymbol{\varphi}_n(\boldsymbol{y})) |||J_{\boldsymbol{\varphi}_n}(\boldsymbol{y})||| \, d\boldsymbol{y}.$$

10 Manifolds

We now manifolds that cannot be parametrized by a single chart.

Definition 113 *Given $1 \leq k \leq N$, a nonempty set $M \subseteq \mathbb{R}^N$ is called a k-dimensional* differential surface *or* manifold *if for every $\boldsymbol{x}_0 \in M$ there exist an open set U containing $\boldsymbol{x}_0$ and a differentiable function $\boldsymbol{\varphi} : V \to \mathbb{R}^N$, where $V \subseteq \mathbb{R}^k$ is an open set such that*

(i) $\boldsymbol{\varphi} : V \to M \cap U$ is a homeomorphism, that is, it is invertible and continuous together with its inverse $\boldsymbol{\varphi}^{-1} : M \cap U \to V$,

(ii) $D\boldsymbol{\varphi}(\boldsymbol{y})$ has rank k for all $\boldsymbol{y} \in V$.

The function $\boldsymbol{\varphi}$ is called a local chart *or a* system of local coordinates *or a* local parametrization *around $\boldsymbol{x}_0$. We say that M is of class C^m, $m \in \mathbb{N}$, (respectively, C^∞) if all local charts are of class C^m (respectively, C^∞).*

Roughly speaking a set $M \subset \mathbb{R}^N$ is a k-dimensional differential surface if for every point $\boldsymbol{x}_0 \in M$ we can "cut" a piece of M around $\boldsymbol{x}_0$ and deform it/flatten it in a smooth way to get, say, a ball of $\mathbb{R}^k$. Another way to say this is that locally M looks like $\mathbb{R}^k$. Thus, a sphere in $\mathbb{R}^3$ is a 2-dimensional surface because locally it looks like $\mathbb{R}^2$, while a cone is not because near the tip it does not look like $\mathbb{R}^2$. A simple way to construct k-dimensional differential surface is to start with a set of $\mathbb{R}^k$ and then deform it in a smooth way.

Example 114 *Consider the hyperbola*

$$M := \left\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 1\right\}.$$

To cover M we need at least two local charts, precisely, we can take the open sets

$$V := \left\{(x, y) \in \mathbb{R}^2 : x > 0\right\}, \quad W := \left\{(x, y) \in \mathbb{R}^2 : x < 0\right\},$$

and the functions $\boldsymbol{\varphi} : \mathbb{R} \to M \cap V$ and $\boldsymbol{\psi} : \mathbb{R} \to M \cap W$ defined by

$$\boldsymbol{\varphi}(t) := \left(\sqrt{1+t^2}, t\right), \quad \boldsymbol{\psi}(t) := \left(-\sqrt{1+t^2}, t\right), \quad t \in \mathbb{R}.$$

Note that both $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ are of class C^∞ (the argument inside the square roots is never zero). Moreover, $\boldsymbol{\varphi}'(t) = \left(\frac{t}{\sqrt{1+t^2}}, 1\right)$ and $\boldsymbol{\psi}'(t) := \left(-\frac{t}{\sqrt{1+t^2}}, 1\right)$, and so the rank of $\boldsymbol{\varphi}'(t)$ and of $\boldsymbol{\psi}'(t)$ is one. Finally, $\boldsymbol{\varphi}^{-1} : M \cap V \to \mathbb{R}$ and $\boldsymbol{\psi}^{-1} : M \cap W \to \mathbb{R}$ are given by

$$\boldsymbol{\varphi}^{-1}(x, y) = y, \quad \boldsymbol{\psi}^{-1}(x, y) = y,$$

which are continuous. Thus, M is a 1-dimensional surface of class C^∞.

Remark 115 *Given a k-dimensional surface M of class C^m, for every $\boldsymbol{x} \in M$ there exist and open set $U_{\boldsymbol{x}}$ and a local chart $\boldsymbol{\varphi}_{\boldsymbol{x}} : V_{\boldsymbol{x}} \to \mathbb{R}^N$ such that $\boldsymbol{x} \in \boldsymbol{\varphi}_{\boldsymbol{x}}(V_{\boldsymbol{x}}) \subseteq U_{\boldsymbol{x}}$ and $M \cap U_{\boldsymbol{x}} = \boldsymbol{\varphi}_{\boldsymbol{x}}(V_{\boldsymbol{x}})$. Hence, $M \subseteq \bigcup_{\boldsymbol{x} \in M} U_{\boldsymbol{x}}$. But then we can find countably many U_n and local charts $\boldsymbol{\varphi}_n$ such that $\bigcup_n U_n = \bigcup_{\boldsymbol{x} \in M} U_{\boldsymbol{x}}$ so that*

$$M \subseteq \bigcup_n \boldsymbol{\varphi}_n(V_n) \subseteq \bigcup_n U_n.$$

Hence, if $E \subseteq M$ we are extactly in the situation (48).

Exercise 116 *Given the set*

$$M := \left\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, \, x, y, z > 0\right\},$$

prove that it is a 2-dimensional surface and find its surface area.

Exercise 117 *Given the set*

$$M := \left\{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2, \, x^2 + y^2 < 1\right\},$$

prove that it is a 2-dimensional surface and find its surface area.

The next two theorems give an equivalent definition of surfaces, which are very useful for examples. We begin by showing that a manifold can be written locally as the graph of a function.

Theorem 118 *Given $1 \le k < N$, a nonempty set $M \subseteq \mathbb{R}^N$, and $m \in \mathbb{N}$, the following are equivalent*

(i) M is a k-dimensional surface of class C^m.

(ii) For every $\boldsymbol{x}_0 \in M$ there exist an open set $U \subseteq \mathbb{R}^N$ containing $\boldsymbol{x}_0$, an open set $V \subseteq \mathbb{R}^k$, and a function $\boldsymbol{f} : V \to \mathbb{R}^{N-k}$ of class C^m, such that, by relabelling the coordinates, if necessary,

$$M \cap U = \{(\boldsymbol{y}, \boldsymbol{f}(\boldsymbol{y})) : \boldsymbol{y} \in V\}.$$

Proof. Step 1: We prove that (i) implies (ii). Given $\boldsymbol{x}_0 \in M$, let U, V, and $\boldsymbol{\varphi}$ be as in Definition 113. Let $\boldsymbol{y}_0 \in V$ be such that $\boldsymbol{\varphi}(\boldsymbol{y}_0) = \boldsymbol{x}_0$. Since $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$ has rank k, there is an $k \times k$ submatrix of $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$, which has determinant different from zero. By changing the coordinates axes of $\mathbb{R}^N$, if necessary, without loss of generality, we may assume for simplicity assume that

$$\det \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_1}{\partial y_k}(\boldsymbol{y}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_k}{\partial y_k}(\boldsymbol{y}_0) \end{pmatrix} \neq 0.$$

Let $\mathbf{w} := (x_{k+1}, \dots, x_N)$ so that $\boldsymbol{x} = (\mathbf{z}, \mathbf{w})$. Let $\boldsymbol{g} : V \to \mathbb{R}^k$ be defined by

$$\boldsymbol{g}(\boldsymbol{y}) := (\varphi_1(\boldsymbol{y}), \dots, \varphi_k(\boldsymbol{y})). \tag{49}$$

Then

$$\det J_{\boldsymbol{g}}(\boldsymbol{y}_0) = \det \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_1}{\partial y_k}(\boldsymbol{y}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_k}{\partial y_k}(\boldsymbol{y}_0) \end{pmatrix} \neq 0,$$

and so by the inverse function theorem there exists $r > 0$ such that $B_k(\boldsymbol{y}_0, r) \subseteq V$, $\boldsymbol{f}(B_k(\boldsymbol{y}_0, r))$ is open, and $\boldsymbol{g} : B_k(\boldsymbol{y}_0, r) \to \boldsymbol{g}(B_k(\boldsymbol{y}_0, r))$ is invertible, with inverse $\boldsymbol{g}^{-1} : \boldsymbol{f}(B_k(\boldsymbol{y}_0, r)) \to B_k(\boldsymbol{y}_0, r)$ of class C^m. Hence, we have shown that we can write $\boldsymbol{y}$ as a function of $\mathbf{z}$, $\boldsymbol{y} = \boldsymbol{g}^{-1}(\mathbf{z})$.

Since $\boldsymbol{\varphi}$ is a homeomorphism, the set $\boldsymbol{\varphi}(B_k(\boldsymbol{y}_0, r))$ is relatively open in M, that is, it can be written as

$$\boldsymbol{\varphi}(B_k(\boldsymbol{y}_0, r)) = M \cap U_1$$

for some open set $U_1 \subseteq \mathbb{R}^N$. Then

$$\begin{aligned} M \cap U_1 &= \{\boldsymbol{\varphi}(\boldsymbol{y}) : \boldsymbol{y} \in B_k(\boldsymbol{y}_0, r)\}. \\ &= \left\{\left(\mathbf{z}, \varphi_{k+1}\left(\boldsymbol{g}^{-1}(\mathbf{z})\right), \dots, \varphi_N\left(\boldsymbol{g}^{-1}(\mathbf{z})\right)\right) : \mathbf{z} \in \boldsymbol{g}(B_k(\boldsymbol{y}_0, r))\right\}. \end{aligned}$$

This shows that $M \cap U_1$ is given by the graph of the function $\mathbf{z} \in \boldsymbol{g}(B_k(\boldsymbol{y}_0, r)) \mapsto \left(\varphi_{k+1}\left(\boldsymbol{g}^{-1}(\mathbf{z})\right), \ldots, \varphi_N\left(\boldsymbol{g}^{-1}(\mathbf{z})\right)\right)$.

Step 2: We prove that (ii) implies (i). Given $\boldsymbol{x}_0 \in M$, let U, V and $\boldsymbol{f} : V \to \mathbb{R}^{N-k}$ of class C^m be as such that, by relabelling the coordinates, if necessary,

$$M \cap U = \{(\boldsymbol{y}, \boldsymbol{f}(\boldsymbol{y})) : \boldsymbol{y} \in V\}.$$

Define $\boldsymbol{\varphi} : V \to \mathbb{R}^N$ by

$$\boldsymbol{\varphi}(\boldsymbol{y}) := (\boldsymbol{y}, \boldsymbol{f}(\boldsymbol{y})).$$

Then $\boldsymbol{\varphi}$ is of class C^m, injective, $\boldsymbol{\varphi}^{-1} : \boldsymbol{\varphi}(V) \to \mathbb{R}^k$ is continuous, since $\boldsymbol{\varphi}^{-1}(\boldsymbol{y}, \boldsymbol{w}) = \boldsymbol{y}$, and

$$J_{\boldsymbol{\varphi}}(\boldsymbol{y}) = \begin{pmatrix} I_k \\ J_{\boldsymbol{f}}(\boldsymbol{y}) \end{pmatrix}$$

which has rank k. ■

Next we show that a manifold can be written locally as the set of zeros of a function.

Proposition 119 *Given $1 \leq k < N$, a nonempty set $M \subseteq \mathbb{R}^N$, and $m \in \mathbb{N}$, then the following are equivalent:*

(i) M is a k-dimensional surface of class C^m.

(ii) For every $\boldsymbol{x}_0 \in M$ there exist an open set $U \subseteq \mathbb{R}^N$ containing $\boldsymbol{x}_0$ and a function $\boldsymbol{g} : U \to \mathbb{R}^{N-k}$ of class C^m, such that

$$M \cap U = \{\boldsymbol{x} \in U : \boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}\}$$

and $J_{\boldsymbol{g}}(\boldsymbol{x})$ has rank $N-k$ for all $\boldsymbol{x} \in M \cap U$.

Proof. Step 1: We prove that (i) implies (ii). By the previous theorem, for every $\boldsymbol{x}_0 \in M$ there exist an open set $U \subseteq \mathbb{R}^N$ containing $\boldsymbol{x}_0$, an open set $V \subseteq \mathbb{R}^k$, and a function $\boldsymbol{f} : V \to \mathbb{R}^{N-k}$ of class C^m, such that, by relabelling the coordinates, if necessary,

$$M \cap U = \{(\boldsymbol{y}, \boldsymbol{f}(\boldsymbol{y})) : \boldsymbol{y} \in V\}.$$

Consider the function $\boldsymbol{g} : U \to \mathbb{R}^{N-k}$ of class C^m defined by

$$\boldsymbol{g}(\boldsymbol{x}) := (x_{k+1} - f_1(x_1, \ldots, x_k), \ldots, x_N - f_{N-k}(x_1, \ldots, x_k)).$$

Then

$$M \cap U = \{\boldsymbol{x} \in U : \boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}\}.$$

Moreover, $J_{\boldsymbol{g}}(\boldsymbol{x})$ contains the submatrix I_{N-k}, since for $i, j \geq k+1$,

$$\frac{\partial g_i}{\partial x_j}(\boldsymbol{x}) = \frac{\partial}{\partial x_j}(x_i - f_{i-k}(x_1, \ldots, x_k)) = \delta_{i,j} - 0.$$

Hence, $J_{\boldsymbol{g}}(\boldsymbol{x})$ has rank $N-k$ for all $\boldsymbol{x} \in U_1$. ■

Wednesday, November 9, 2022

Proof. Step 2: We prove that (ii) implies (i). Since $J_{\boldsymbol{g}}(\boldsymbol{x}_0)$ has rank $N-k$, there is an $(N-k)\times(N-k)$ submatrix of $J_{\boldsymbol{g}}(\boldsymbol{x}_0)$, which has determinant different from zero. By relabeling the coordinates, if necessary, we can assume for simplicity that

$$\det\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial g_1}{\partial x_{N-k}}(\boldsymbol{x}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_{N-k}}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial g_{N-K}}{\partial x_{N-k}}(\boldsymbol{x}_0) \end{pmatrix} \neq 0.$$

Let $\mathbf{z} := (x_1, \dots, x_{N-k})$ and $\boldsymbol{y} := (x_{N-k+1}, \dots, x_N)$, so that $\boldsymbol{x} = (\mathbf{z}, \boldsymbol{y})$, $\boldsymbol{x}_0 = (\mathbf{z}_0, \boldsymbol{y}_0)$, and $\det \frac{\partial \boldsymbol{g}}{\partial \mathbf{z}}(\boldsymbol{x}_0) \neq 0$. Consider the function $\boldsymbol{f} : U \to \mathbb{R}^N$ defined by

$$\boldsymbol{f}(\boldsymbol{x}) := (\boldsymbol{g}(\boldsymbol{x}), \boldsymbol{y}).$$

Then

$$\det J_{\boldsymbol{f}}(\boldsymbol{x}_0) = \det\begin{pmatrix} \frac{\partial \boldsymbol{g}}{\partial \mathbf{z}}(\boldsymbol{x}_0) & \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{y}}(\boldsymbol{x}_0) \\ 0_{k\times(N-k)} & I_k \end{pmatrix} = \det \frac{\partial \boldsymbol{g}}{\partial \mathbf{z}}(\boldsymbol{x}_0) \neq 0,$$

and so by the inverse function theorem there exists $r > 0$ such that $B(\boldsymbol{x}_0, r) \subseteq U$, $\boldsymbol{f}(B(\boldsymbol{x}_0, r))$ is open, and $\boldsymbol{f} : B(\boldsymbol{x}_0, r) \to \boldsymbol{f}(B(\boldsymbol{x}_0, r))$ is invertible, with inverse $\boldsymbol{f}^{-1} : \boldsymbol{f}(B(\boldsymbol{x}_0, r)) \to B(\boldsymbol{x}_0, r)$ of class C^m. Since $\boldsymbol{f}(B(\boldsymbol{x}_0, r))$ is open and contains $\boldsymbol{f}(\boldsymbol{x}_0) = (\mathbf{0}, \boldsymbol{y}_0)$, we may find balls $B_{N-k}(\mathbf{0}, r_0)$ and $B_k(\boldsymbol{y}_0, r_0)$ such that $B_{N-k}(\mathbf{0}, r_0) \times B_k(\boldsymbol{y}_0, r_0) \subseteq \boldsymbol{f}(B(\boldsymbol{x}_0, r))$. Then $U_1 := \boldsymbol{f}^{-1}(B_{N-k}(\mathbf{0}, r_0) \times B_k(\boldsymbol{y}_0, r_0))$ is open. Moreover, if $\boldsymbol{x} \in M \cap U_1$, then $\boldsymbol{f}(\boldsymbol{x}) = (\mathbf{0}, \boldsymbol{y})$. Hence, the function $\boldsymbol{\varphi} : B_k(\boldsymbol{y}_0, r_0) \to M \cap U_1$ defined by

$$\boldsymbol{\varphi}(\boldsymbol{y}) := \boldsymbol{f}^{-1}(\mathbf{0}, \boldsymbol{y})$$

is a homeomorphism. Since $\boldsymbol{f}(\boldsymbol{x}) = (\boldsymbol{g}(\boldsymbol{x}), \boldsymbol{y})$, we have that $\boldsymbol{\varphi}(\boldsymbol{y}) = \boldsymbol{f}^{-1}(\mathbf{0}, \boldsymbol{y})$ takes the form $\boldsymbol{\varphi}(\boldsymbol{y}) = \boldsymbol{f}^{-1}(\mathbf{0}, \boldsymbol{y}) = (\mathbf{h}(\mathbf{0}, \boldsymbol{y}), \boldsymbol{y})$, and so

$$J_{\boldsymbol{\varphi}}(\boldsymbol{y}) = \begin{pmatrix} J_{\boldsymbol{h}}(\boldsymbol{y}) \\ I_k \end{pmatrix},$$

which shows that $J_{\boldsymbol{\varphi}}(\boldsymbol{y})$ has rank k for all $\boldsymbol{y} \in B_k(\boldsymbol{y}_0, r_0)$. ■

Next we define tangent vectors and normal vectors.

Definition 120 *Let $1 \leq k < N$, and let M be a k-dimensional differential surface. Given $\boldsymbol{x}_0 \in M$, a vector $\boldsymbol{t} \in \mathbb{R}^N$ is called a* tangent vector *to M at the point $\boldsymbol{x}_0$ if there exists a function $\mathbf{h} : (-\delta, \delta) \to \mathbb{R}^N$ differentiable at $t = 0$ such that $\mathbf{h}((-\delta, \delta)) \subseteq M$, $\mathbf{h}(0) = \boldsymbol{x}_0$ and $\mathbf{h}'(0) = \boldsymbol{t}$. The set of all tangent vectors to M at $\boldsymbol{x}_0$ is called the* tangent space *to M at $\boldsymbol{x}_0$ and is denoted $T_M(\boldsymbol{x}_0)$.*

Friday, November 11, 2022

Theorem 121 *Let $1 \leq k < N$, and let M be a k-dimensional differential surface of class C^m, $m \in \mathbb{N}$. Given $\boldsymbol{x}_0 \in M$, let $\boldsymbol{\varphi} : V \to \mathbb{R}^N$ be a local chart such that $\boldsymbol{\varphi}(\boldsymbol{y}_0) = \boldsymbol{x}_0$ for some $\boldsymbol{y}_0 \in V$. Then the vectors $\frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0)$, $\ldots$, $\frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0)$ form a basis for the tangent space $T_M(\boldsymbol{x}_0)$ to M at $\boldsymbol{x}_0$.*

Proof. Step 1: We prove that

$$T_M(\boldsymbol{x}_0) \subseteq \operatorname{span}\left\{ \frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0), \ldots, \frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0) \right\}.$$

Let $\boldsymbol{t} \in T_M(\boldsymbol{x}_0)$ and let $\mathbf{h} : (-\delta, \delta) \to \mathbb{R}^N$ be differentiable at $t = 0$ with $\mathbf{h}((-\delta, \delta)) \subseteq M$, $\mathbf{h}(0) = \boldsymbol{x}_0$ and $\mathbf{h}'(0) = \boldsymbol{t}$. Let $\boldsymbol{\varphi} : V \to \mathbb{R}^N$ be a local chart, with $\boldsymbol{x}_0 \in \boldsymbol{\varphi}(V) = M \cap U$ and let $\boldsymbol{y}_0 \in V$ be such that $\boldsymbol{\varphi}(\boldsymbol{y}_0) = \boldsymbol{x}_0$. Since $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$ has rank k, there is an $k \times k$ submatrix of $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$, which has determinant different from zero. By changing the coordinates axes of $\mathbb{R}^N$, if necessary, without loss of generality, we may assume for simplicity assume that

$$\det \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_1}{\partial y_k}(\boldsymbol{y}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) \end{pmatrix} \neq 0.$$

Consider the function $\boldsymbol{f} : V \times \mathbb{R}^{N-k} \to \mathbb{R}^N$ be defined by

$$\boldsymbol{f}(\boldsymbol{y}, \mathbf{z}) := (\varphi_1(\boldsymbol{y}), \ldots, \varphi_k(\boldsymbol{y}), \varphi_{k+1}(\boldsymbol{y}) + z_1, \ldots, \varphi_N(\boldsymbol{y}) + z_{N-k}).$$

Then

$$\begin{aligned} \det J_{\boldsymbol{f}}(\boldsymbol{y}_0, \mathbf{0}) &= \det \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_1}{\partial y_k}(\boldsymbol{y}_0) & \\ \vdots & \cdots & \vdots & 0_{k \times (N-k)} \\ \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) & \\ \frac{\partial \varphi_{k+1}}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_{k+1}}{\partial y_k}(\boldsymbol{y}_0) & \\ \vdots & \cdots & \vdots & I_{N-k} \\ \frac{\partial \varphi_N}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_N}{\partial y_1}(\boldsymbol{y}_0) & \end{pmatrix} \\ &= \det \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_1}{\partial y_k}(\boldsymbol{y}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) & \cdots & \frac{\partial \varphi_k}{\partial y_1}(\boldsymbol{y}_0) \end{pmatrix} \neq 0, \end{aligned}$$

and so by the inverse function theorem there exists $r > 0$ such that $B_N((\boldsymbol{y}_0, \mathbf{0}), r) \subset V \times \mathbb{R}^{N-k}$, $\boldsymbol{f}(B_N((\boldsymbol{y}_0, \mathbf{0}), r))$ is open, and $\boldsymbol{f} : B_N((\boldsymbol{y}_0, \mathbf{0}), r) \to \boldsymbol{f}(B_N((\boldsymbol{y}_0, \mathbf{0}), r))$ is invertible, with inverse $\boldsymbol{f}^{-1} : \boldsymbol{f}(B_N((\boldsymbol{y}_0, \mathbf{0}), r)) \to B_N((\boldsymbol{y}_0, \mathbf{0}), r)$ of class C^m. Moreover,

$$\boldsymbol{f}(\boldsymbol{y}, \mathbf{0}) = \boldsymbol{\varphi}(\boldsymbol{y}) \in M \quad \text{for all } \boldsymbol{y} \in B_k(\boldsymbol{y}_0, r).$$

In turn for $\boldsymbol{y} \in B_k(\boldsymbol{y}_0, r)$,

$$(\boldsymbol{y}, \mathbf{0}) = \boldsymbol{f}^{-1}(\boldsymbol{f}(\boldsymbol{y}, \mathbf{0})) = \boldsymbol{f}^{-1}(\boldsymbol{\varphi}(\boldsymbol{y})),$$

which shows that for all $i = i, \ldots, k$, the first k components of $\boldsymbol{f}^{-1}$ coincide with $\boldsymbol{\varphi}^{-1}$ on points $\boldsymbol{\varphi}(\boldsymbol{y})$. In particular, if $\boldsymbol{x} \in \boldsymbol{f}(B_N((\boldsymbol{y}_0, \mathbf{0}), r)) \cap M$, then $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{y}, \mathbf{0}) = \boldsymbol{\varphi}(\boldsymbol{y})$ for some $\boldsymbol{y} \in B_k(\boldsymbol{y}_0, r)$ and so

$$\left(\boldsymbol{\varphi}^{-1}\right)_i(\boldsymbol{x}) = \left(\boldsymbol{f}^{-1}\right)_i(\boldsymbol{x})$$

for all $i = i, \ldots, k$. Thus, $\boldsymbol{\varphi}^{-1}$ is differentiable in $\boldsymbol{f}(B_N((\boldsymbol{y}_0, \mathbf{0}), r)) \cap M$. It follows by the chain rule that the function $\boldsymbol{\varphi}^{-1} \circ \mathbf{h} : (-\delta, \delta) \to \mathbb{R}^k$ is differentiable at 0. Writing

$$\mathbf{h} = \boldsymbol{\varphi} \circ \left(\boldsymbol{\varphi}^{-1} \circ \mathbf{h}\right),$$

it follows by the chain rule that

$$\boldsymbol{t} = \mathbf{h}'(0) = J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)\left(\boldsymbol{\varphi}^{-1} \circ \mathbf{h}\right)'(0),$$

which shows that $\boldsymbol{t} \in \operatorname{span}\left\{\frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0), \ldots, \frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0)\right\}$. This shows that

$$T_M(\boldsymbol{x}_0) \subseteq \operatorname{span}\left\{\frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0), \ldots, \frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0)\right\}.$$

Step 2: We prove that

$$\operatorname{span}\left\{\frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0), \ldots, \frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0)\right\} \subseteq T_M(\boldsymbol{x}_0).$$

Since V is open, there exists $B_k(\boldsymbol{y}_0, r) \subseteq V$. Let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ be the standard orthonormal basis of $\mathbb{R}^k$. Given a vector $\mathbf{w} \in \mathbb{R}^k$, let

$$\delta := \frac{r}{1 + \|\mathbf{w}\|_k} > 0$$

and consider the function $\mathbf{h} : (-\delta, \delta) \to \mathbb{R}^N$ defined by

$$\mathbf{h}(t) := \boldsymbol{\varphi}(\boldsymbol{y}_0 + t\mathbf{w}).$$

Then $\mathbf{h}((-\delta, \delta)) \subseteq \boldsymbol{\varphi}(B_k(\boldsymbol{y}_0, r)) \subseteq M$, $\mathbf{h}(0) = \boldsymbol{\varphi}(\boldsymbol{y}_0) = \boldsymbol{x}_0$. If $\mathbf{w} = \mathbf{0}$, then $\mathbf{h}$ is constant and so $\mathbf{h}'(0) = \mathbf{0}$. This shows that $\mathbf{0}$ is a tangent vector to M at $\boldsymbol{x}_0$. If $\mathbf{w} \neq \mathbf{0}$, we have

$$\frac{\mathbf{h}(t) - \mathbf{h}(0)}{t} = \frac{\boldsymbol{\varphi}(\boldsymbol{y}_0 + t\mathbf{w}) - \boldsymbol{\varphi}(\boldsymbol{y}_0)}{t} \to \frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{w}}(\boldsymbol{y}_0).$$

This shows that $\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{w}}(\boldsymbol{y}_0)$ is a tangent vector to M at $\boldsymbol{x}_0$. Since $\boldsymbol{\varphi}$ is differentiable, by a theorem from a semester ago (applied to each component),

$$\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{w}}(\boldsymbol{y}_0) = \sum_{i=1}^{k} w_i \frac{\partial \boldsymbol{\varphi}}{\partial y_i}(\boldsymbol{y}_0).$$

This shows that each linear combination of the vectors $\frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0), \ldots, \frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0)$ is a tangent vector, that is that

$$\operatorname{span}\left\{\frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0), \ldots, \frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0)\right\} \subseteq T_M(\boldsymbol{x}_0).$$

Note that since $J_{\boldsymbol{\varphi}}(\boldsymbol{y}_0)$ has rank k, the vectors $\frac{\partial \boldsymbol{\varphi}}{\partial y_1}(\boldsymbol{y}_0), \ldots, \frac{\partial \boldsymbol{\varphi}}{\partial y_k}(\boldsymbol{y}_0)$ are linearly independent. ■

Theorem 122 *Let $1 \leq k < N$, and let M be a k-dimensional surface of class C^m, $m \in \mathbb{N}$, of the form given in Proposition 119. Given $\boldsymbol{x}_0 \in M$, let $\boldsymbol{g} : B(\boldsymbol{x}_0, r) \to \mathbb{R}^{N-k}$ be the function given in Proposition 119 corresponding to the point $\boldsymbol{x}_0$. Then*

$$T_M(\boldsymbol{x}_0) = \ker d\boldsymbol{g}_{\boldsymbol{x}_0} = \left\{\boldsymbol{x} \in \mathbb{R}^N : J_{\boldsymbol{g}}(\boldsymbol{x}_0)\,\boldsymbol{x} = \mathbf{0}\right\}.$$

Proof. Let $\boldsymbol{t} \in T_M(\boldsymbol{x}_0)$ and let $\mathbf{h} : (-\delta, \delta) \to \mathbb{R}^N$ be differentiable at $t = 0$ with $\mathbf{h}((-\delta, \delta)) \subseteq M$, $\mathbf{h}(0) = \boldsymbol{x}_0$ and $\mathbf{h}'(0) = \boldsymbol{t}$. Taking δ smaller, if necessary, we have that

$$\boldsymbol{g}(\mathbf{h}(t)) = \mathbf{0}$$

for all $t \in (-\delta, \delta)$. It follows by the chain rule that

$$\mathbf{0} = J_{\boldsymbol{g}}(\boldsymbol{x}_0)\,\mathbf{h}'(0) = J_{\boldsymbol{g}}(\boldsymbol{x}_0)\,\boldsymbol{t},$$

which shows that $\boldsymbol{t} \in \ker d\boldsymbol{g}_{\boldsymbol{x}_0}$. Hence, $T_M(\boldsymbol{x}_0) \subseteq \ker d\boldsymbol{g}_{\boldsymbol{x}_0}$. On the other hand, since $J_{\boldsymbol{g}}(\boldsymbol{x}_0)$ has rank $N - k$, the dimension of $\ker d\boldsymbol{g}_{\boldsymbol{x}_0}$ is given by

$$N - \operatorname{rank} J_{\boldsymbol{g}}(\boldsymbol{x}_0) = N - (N - k) = k.$$

But $T_M(\boldsymbol{x}_0)$ has also dimension k by the previous theorem. Hence, $T_M(\boldsymbol{x}_0) = \ker d\boldsymbol{g}_{\boldsymbol{x}_0}$. ■

Definition 123 *Let $1 \leq k < N$, and let M be a k-dimensional differential surface. Given $\boldsymbol{x}_0 \in M$, a vector $\boldsymbol{\nu} \in \mathbb{R}^N$ is called a* normal vector to M at the point $\boldsymbol{x}_0$ *if*

$$\boldsymbol{\nu} \cdot \boldsymbol{t} = 0 \quad \text{for all } \boldsymbol{t} \in T_M(\boldsymbol{x}_0).$$

The set of all normal vectors to M at $\boldsymbol{x}_0$ is called the normal space *to M at $\boldsymbol{x}_0$ and is denoted $N_M(\boldsymbol{x}_0)$.*

Since $T_M(\boldsymbol{x}_0)$ is a subspace of dimension k, the normal space $N_M(\boldsymbol{x}_0)$ has dimension $N - K$. When $K = N - 1$, then $N_M(\boldsymbol{x}_0)$ has dimension 1, so $N_M(\boldsymbol{x}_0) = \{t\boldsymbol{\nu} : t \in \mathbb{R}\}$, where $\boldsymbol{\nu} \neq \mathbf{0}$. Taking $\|\boldsymbol{\nu}\| = 1$, at $\boldsymbol{x}_0$ there two choices of unit normal vectors, $\boldsymbol{\nu}$ and $-\boldsymbol{\nu}$.

Exercise 124 *Let M be a k-dimensional surface M of class C^m, $m \in \mathbb{N}$ and let $\boldsymbol{\varphi} : V \to M$, $\boldsymbol{\psi} : W \to M$ be two local charts such that $\boldsymbol{\varphi}(V) \cap \boldsymbol{\psi}(W) =: Z$ is nonempty. Prove that the function $\boldsymbol{\psi}^{-1} \circ \boldsymbol{\varphi} : \boldsymbol{\varphi}^{-1}(Z) \to \boldsymbol{\psi}^{-1}(Z)$ is a diffeomorphism of class C^m, that is, $\boldsymbol{\psi}^{-1} \circ \boldsymbol{\varphi}$ and its inverse are both of class C^m.*

The function $\boldsymbol{\psi}^{-1} \circ \boldsymbol{\varphi}$ is called a *change of parameters* or a *change of coordinates*. The previous exercise leads to the definition of abstract manifolds.

Definition 125 *Given $k \in \mathbb{N}$, a k-dimensional* differential surface *or* manifold *is a nonempty set M together with a family of injective functions $\boldsymbol{\varphi}_\alpha : V_\alpha \to M$, $\alpha \in \Lambda$, where $V_\alpha \subseteq \mathbb{R}^k$ is an open set, such that*

(i) $\bigcup_\alpha \boldsymbol{\varphi}_\alpha(V_\alpha) = M$,

(ii) If $\alpha, \beta \in \Lambda$ are such that $\boldsymbol{\varphi}_\alpha(V_\alpha) \cap \boldsymbol{\varphi}_\beta(V_\beta) =: Z_{\alpha,\beta}$ is nonempty then $\boldsymbol{\varphi}_\alpha^{-1}(Z_{\alpha,\beta})$ and $\boldsymbol{\varphi}_\beta^{-1}(Z_{\alpha,\beta})$ are open sets and the functions $\boldsymbol{\varphi}_\beta^{-1} \circ \boldsymbol{\varphi}_\alpha : \boldsymbol{\varphi}_\alpha^{-1}(Z_{\alpha,\beta}) \to \mathbb{R}^k$ and $\boldsymbol{\varphi}_\alpha^{-1} \circ \boldsymbol{\varphi}_\beta : \boldsymbol{\varphi}_\beta^{-1}(Z_{\alpha,\beta}) \to \mathbb{R}^k$ are differentiable.

The family $\{\boldsymbol{\varphi}_\alpha\}_{\alpha \in \Lambda}$ is called an atlas.

Remark 126 *A differential structure on a set induces a natural topology. We say that $U \subseteq M$ is* open *if $\varphi_\alpha^{-1}(U)$ is open for every $\alpha \in \Lambda$. With this topology, all the local charts φ_α are continuous and $\varphi_\alpha(V_\alpha)$ are open, so that all φ_α become homeomorphisms.*

Definition 127 *Given $k \in \mathbb{N}$, a k-dimensional differential M is called* orientable *if there exists an atlas $\{\boldsymbol{\varphi}_\alpha\}_{\alpha \in \Lambda}$ such that for every $\alpha, \beta \in \Lambda$ with $\boldsymbol{\varphi}_\alpha(V_\alpha) \cap \boldsymbol{\varphi}_\beta(V_\beta) =: Z_{\alpha,\beta}$ nonempty, $J_{\boldsymbol{\varphi}_\beta^{-1} \circ \boldsymbol{\varphi}_\alpha}$ has positive determinant in $\boldsymbol{\varphi}_\alpha^{-1}(Z_{\alpha,\beta})$. Otherwise M is called* non orientable.

It can be shown that an $N-1$-dimensional manifold M of class C^ℓ is orientable if and only if at every point $\boldsymbol{x} \in M$ one can choose a unit normal vector $\boldsymbol{\nu}(\boldsymbol{x}) \in N_M(\boldsymbol{x})$ in such a way that the map

$$\begin{aligned} M &\to \mathbb{R}^N \setminus \{\mathbf{0}\} \\ \boldsymbol{x} &\mapsto \boldsymbol{\nu}(\boldsymbol{x}) \end{aligned}$$

is continuous.

Monday, November 14, 2022

11 Mollifiers

Definition 128 *Given a metric space (X, d), a set $E \subseteq X$ and a function $f : E \to \mathbb{R}$, the* support *of f is the set*

$$\operatorname{supp} f := \overline{\{x \in E : f(x) \neq 0\}}.$$

Definition 129 *Given a metric space (X, d), the space $C_c(X)$ is the space of all continuous functions whose support if compact.*

Definition 130 *Given an open set $U \subseteq \mathbb{R}^N$ and $n \in \mathbb{N}$, the space $C_c^n(U)$ is the space of all functions in $C^n(U)$ whose support is compact set and contained in U. Similarly, $C_c^\infty(U)$ is the space of all functions in $C^\infty(U)$ whose support is compact and contained in U.*

Consider the function

$$\varphi(\boldsymbol{x}) := \begin{cases} c\exp\left(\frac{1}{\|\boldsymbol{x}\|^2-1}\right) & \text{if } \|\boldsymbol{x}\| < 1, \\ 0 & \text{if } \|\boldsymbol{x}\| \geq 1, \end{cases} \tag{50}$$

where the constant $c > 0$ is chosen so that

$$\int_{\mathbb{R}^N} \varphi(\boldsymbol{x})\, d\boldsymbol{x} = 1. \tag{51}$$

We leave as an exercise to prove that $\varphi \in C_c^\infty(\mathbb{R}^N)$. For every $\varepsilon > 0$ we define

$$\varphi_\varepsilon(\boldsymbol{x}) := \frac{1}{\varepsilon^N}\varphi\left(\frac{\boldsymbol{x}}{\varepsilon}\right), \quad \boldsymbol{x} \in \mathbb{R}^N.$$

The functions φ_ε are called *standard mollifiers.*

Remark 131 *Fix $\boldsymbol{x} \in \mathbb{R}^N$. Using the change of variables $\boldsymbol{z} = \frac{\boldsymbol{x}-\boldsymbol{y}}{\varepsilon}$ we have that*

$$\begin{aligned} \int_{B(\boldsymbol{x},\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y})\, d\boldsymbol{y} &= \frac{1}{\varepsilon^N}\int_{B(\boldsymbol{x},\varepsilon)} \varphi\left(\frac{\boldsymbol{x}-\boldsymbol{y}}{\varepsilon}\right) d\boldsymbol{y} \\ &= \frac{\varepsilon^N}{\varepsilon^N}\int_{B(\boldsymbol{0},1)} \varphi(\boldsymbol{z})\, d\boldsymbol{z} = 1. \end{aligned}$$

Given a Lebesgue measurable set $E \subseteq \mathbb{R}^N$ and a Lebesgue integrable function $f : E \to \mathbb{R}$, we define

$$f_\varepsilon(\boldsymbol{x}) := \int_E \varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y})\, f(\boldsymbol{y})\, d\boldsymbol{y} \tag{52}$$

for $\boldsymbol{x} \in \mathbb{R}^N$. Since φ_ε is bounded and continuous, and f is Lebesgue integrable, $f_\varepsilon(\boldsymbol{x})$ is well-defined. The function $f_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ is called a *mollification* of f.

Theorem 132 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set, $f : E \to \mathbb{R}$ be a Lebesgue integrable function. Then $f_\varepsilon \in C^\infty(\mathbb{R}^N)$ and for every multi-index $\boldsymbol{\alpha}$,*

$$\frac{\partial^{\boldsymbol{\alpha}} f_\varepsilon}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}) = \int_E \frac{\partial^{\boldsymbol{\alpha}} \varphi_\varepsilon}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}-\boldsymbol{y})\, f(\boldsymbol{y})\, d\boldsymbol{y}.$$

Proof. Let's prove that f_ε is of class C^1. Fix $\boldsymbol{x} \in \mathbb{R}^N$ and let $\boldsymbol{e}_i$, $i = 1, \ldots, N$, be an element of the canonical basis of $\mathbb{R}^N$. For every $h \in \mathbb{R} \setminus \{0\}$ consider

$$\begin{aligned} &\frac{f_\varepsilon(\boldsymbol{x}+h\boldsymbol{e}_i) - f_\varepsilon(\boldsymbol{x})}{h} - \int_E \frac{\partial \varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y} \\ &= \int_E \left(\frac{\varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y}+h\boldsymbol{e}_i) - \varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y})}{h} - \frac{\partial \varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y})\right) f(\boldsymbol{y})\, d\boldsymbol{y}. \end{aligned}$$

By the mean value theorem

$$\frac{\varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y}+h\boldsymbol{e}_i)-\varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y})}{h}=\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y}+\theta h\boldsymbol{e}_i)$$

for some $\theta\in(0,1)$. Since $\frac{\partial\varphi_\varepsilon}{\partial x_i}$ is continuous and it is zero outside $B(\mathbf{0},\varepsilon)$, by the Weierstrass theorem applied in the compact set $\overline{B(\mathbf{0},\varepsilon)}$ there exists a constant $M_\varepsilon>0$ such that

$$\left|\frac{\partial\varphi_\varepsilon}{\partial x_i}(z)\right|\le M_\varepsilon$$

for all $\boldsymbol{z}\in\mathbb{R}^N$. Then we have

$$\begin{aligned}&\left|\left(\frac{\varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y}+h\boldsymbol{e}_i)-\varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y})}{h}-\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y})\right)f(\boldsymbol{y})\right|\\&=\left|\left(\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y}+\theta h\boldsymbol{e}_i)-\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y})\right)f(\boldsymbol{y})\right|\le 2M_\varepsilon|f(\boldsymbol{y})|.\end{aligned}$$

Since f is Lebesgue integrable, we can apply the Lebesgue dominated convergence theorem to conclude that

$$\begin{aligned}&\lim_{h\to0}\left(\frac{f_\varepsilon(\boldsymbol{x}+h\boldsymbol{e}_i)-f_\varepsilon(\boldsymbol{x})}{h}-\int_E\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y})f(\boldsymbol{y})\,d\boldsymbol{y}\right)\\&=\lim_{h\to0}\int_E\left(\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y}+\theta h\boldsymbol{e}_i)-\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y})\right)f(\boldsymbol{y})\,d\boldsymbol{y}\\&=\int_E\lim_{h\to0}\left(\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y}+\theta h\boldsymbol{e}_i)-\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y})\right)f(\boldsymbol{y})\,d\boldsymbol{y}=0,\end{aligned}$$

where we used the fact that $\frac{\partial\varphi_\varepsilon}{\partial x_i}$ is continuous. This shows that

$$\frac{\partial f_\varepsilon}{\partial x_i}(\boldsymbol{x})=\int_E\frac{\partial\varphi_\varepsilon}{\partial x_i}(\boldsymbol{x}-\boldsymbol{y})\,f(\boldsymbol{y})\,d\boldsymbol{y}.$$

A similar but simppler argument shows that $\frac{\partial f_\varepsilon}{\partial x_i}$ is continuous.

Note that the only properties that we have used on the function φ_ε are that $\varphi_\varepsilon\in C_c^\infty(\mathbb{R}^N)$ with $\operatorname{supp}\varphi_\varepsilon\subseteq\overline{B(\mathbf{0},\varepsilon)}$. Hence, the same proof carries over if we replace φ_ε with $\psi_\varepsilon:=\frac{\partial\varphi_\varepsilon}{\partial x_i}$. Thus, by induction we may prove that for every multi-index $\boldsymbol{\alpha}$ there holds

$$\frac{\partial^{\boldsymbol{\alpha}}f_\varepsilon}{\partial\boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x})=\int_E\frac{\partial^{\boldsymbol{\alpha}}\varphi_\varepsilon}{\partial\boldsymbol{x}^{\boldsymbol{\alpha}}}(\boldsymbol{x}-\boldsymbol{y})\,f(\boldsymbol{y})\,d\boldsymbol{y}.$$

■

Exercise 133 *Prove that f_ε and all its partial derivatives are uniformly continuous.*

Theorem 134 *Let $E \subseteq \mathbb{R}^N$ be a Lebesgue measurable set, $f : E \to \mathbb{R}$ be a Lebesgue integrable function. Assume that f is continuous at some $\boldsymbol{x}_0 \in E^\circ$. Then $f_\varepsilon(\boldsymbol{x}_0) \to f(\boldsymbol{x}_0)$. Moreover, if f is continuous in E°. Then $f_\varepsilon \to f$ uniformly on compact sets of E°.*

Proof. Step 1: Since $\boldsymbol{x}_0 \in E^\circ$, there exists $B(\boldsymbol{x}_0, r) \subseteq E$. Take $0 < \varepsilon \leq r$. Then $B(\boldsymbol{x}_0, \varepsilon) \subseteq E$. Since $\varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y}) = 0$ for all $\boldsymbol{y}$ with $\|\boldsymbol{x}_0 - \boldsymbol{y}\| \geq \varepsilon$,

$$f_\varepsilon(\boldsymbol{x}_0) = \int_{B(\boldsymbol{x}_0,\varepsilon)\cap E} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y} = \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y}.$$

Using Remark 131 we can write

$$\begin{aligned} f_\varepsilon(\boldsymbol{x}_0) - f(\boldsymbol{x}_0) &= \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y}) f(\boldsymbol{y})\, d\boldsymbol{y} - 1 f(\boldsymbol{x}_0) \\ &= \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y}) \left[f(\boldsymbol{y}) - f(\boldsymbol{x}_0)\right] d\boldsymbol{y} \end{aligned}$$

and so

$$|f_\varepsilon(\boldsymbol{x}_0) - f(\boldsymbol{x}_0)| \leq \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y}) \left|f(\boldsymbol{y}) - f(\boldsymbol{x}_0)\right| d\boldsymbol{y}. \tag{53}$$

Since f is continuous at $\boldsymbol{x}_0$ given $\rho > 0$ there exists $\delta = \delta(\boldsymbol{x}_0, \rho) > 0$ such that

$$|f(\boldsymbol{y}) - f(\boldsymbol{x}_0)| \leq \rho$$

for all $\boldsymbol{y} \in R$ with $\|\boldsymbol{y} - \boldsymbol{x}_0\| < \delta$. Hence, taking $\varepsilon < \delta$ we have that

$$\begin{aligned} |f_\varepsilon(\boldsymbol{x}_0) - f(\boldsymbol{x}_0)| &\leq \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y}) \left|f(\boldsymbol{y}) - f(\boldsymbol{x}_0)\right| d\boldsymbol{y} \\ &\leq \rho \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y})\, d\boldsymbol{y} = \rho, \end{aligned}$$

which proves that $f_\varepsilon(\boldsymbol{x}_0) \to f(\boldsymbol{x}_0)$ as $\varepsilon \to 0^+$.

Step 2: Assume that f is continuous in E°. Let $K \subset E^\circ$ be a compact set. For any fixed

$$0 < \eta < \operatorname{dist}(K, \partial E)$$

let

$$K_\eta := \left\{\boldsymbol{x} \in \mathbb{R}^N : \operatorname{dist}(\boldsymbol{x}, K) \leq \eta\right\}.$$

so that $K_\eta \subset E^\circ$. Since K_η is compact and f is uniformly continuous on K_η, for every $\rho > 0$ there exists $\delta = \delta(\eta, K, \rho) > 0$ such that

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq \rho \tag{54}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in K_\eta$, with $\|\boldsymbol{x}-\boldsymbol{y}\| \leq \delta$. Let $0 < \varepsilon < \min\{\delta, \eta\}$. Then for all $\boldsymbol{x} \in K$, we have that $B(\boldsymbol{x}, \varepsilon) \subseteq K_\eta$ and so reasoning as in the first part of the proof

$$\begin{aligned} |f_\varepsilon(\boldsymbol{x}) - f(\boldsymbol{x})| &\leq \int_{B(\boldsymbol{x},\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y}) |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} \\ &\leq \rho \int_{B(\boldsymbol{x},\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}-\boldsymbol{y}) \, d\boldsymbol{y} = \rho, \end{aligned}$$

which shows that

$$\sup_{\boldsymbol{x}\in K} |f_\varepsilon(\boldsymbol{x}) - f(\boldsymbol{x})| \leq \rho$$

for all $0 < \varepsilon < \min\{\delta, \eta\}$. ■

Wednesday, November 16, 2022

Given a locally integrable function $f : \mathbb{R}^N \to \mathbb{R}$, the *(Hardy–Littlewood) maximal function* of f is defined by

$$\mathrm{M}(f)(\boldsymbol{x}) := \sup_{r>0} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f| \, d\boldsymbol{y}$$

for all $\boldsymbol{x} \in \mathbb{R}^N$.

Exercise 135 *Prove that for every $t > 0$ the set $\{\boldsymbol{x} \in \mathbb{R}^N : \mathrm{M}(f)(\boldsymbol{x}) > t\}$ is open and that $\mathrm{M}(f)$ is a Borel function.*

Theorem 136 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be integrable. Then for every $t > 0$,*

$$\mathcal{L}^N\left(\left\{\boldsymbol{x} \in \mathbb{R}^N : \mathrm{M}(f)(\boldsymbol{x}) > t\right\}\right) \leq \frac{3^N}{t} \int_{\mathbb{R}^N} |f| \, d\boldsymbol{y}. \tag{55}$$

Lemma 137 (Vitali's covering) *Let $\mathcal{F}$ be a finite family of be open balls in $\mathbb{R}^N$. Then there exists a subfamily $\mathcal{G}$ of disjoint balls such that*

$$\bigcup_{B\in\mathcal{F}} B \subseteq \bigcup_{B\in\mathcal{G}} 3B,$$

where $3B$ denotes the ball with the same center of B and three times its radius.

Proof. Let B_1 be the ball with largest radius. If all the other balls intersect B_1, we stop and take $\mathcal{G} = \{B_1\}$. Otherwise, let $\mathcal{F}_1$ be the subfamily of balls that do not intersect B_1. Let B_2 be the ball with largest radius in $\mathcal{F}_1$ and add B_2 to $\mathcal{G}$. Inductively, assume that $B_1, \ldots, B_n$ have been chosen. If every ball in $\mathcal{F}$ intersects one of the balls $B_1, \ldots, B_n$, we stop. Otherwise, let $\mathcal{F}_n$ be the subfamily of balls that do not intersect $B_1, \ldots, B_n$. Let B_{n+1} be the ball with largest radius in $\mathcal{F}_n$ and add B_{n+1} to $\mathcal{G}$. Since $\mathcal{F}$ has finitely many elements. this process stop.

Hence, we constructed a subfamily $\mathcal{G}$ of disjoint balls with the property that every ball in $\mathcal{F}$ intersects one ball in $\mathcal{G}$.

Let $\boldsymbol{x} \in \bigcup_{B \in \mathcal{F}} B$. Then there exists $B = B(\boldsymbol{x}_0, r_0) \in \mathcal{F}$ such that $\boldsymbol{x} \in B$. If B belongs to $\mathcal{G}$, we are done. Otherwise, $B(\boldsymbol{x}_0, r_0)$ had been discarded at some point. This means that there exists $B(\boldsymbol{x}_1, r_1) \in \mathcal{G}$ be such that $B(\boldsymbol{x}_0, r_0) \cap B(\boldsymbol{x}_1, r_1) \neq \emptyset$ and $r_1 \geq r_0$. Let $\boldsymbol{y} \in B(\boldsymbol{x}_0, r_0) \cap B(\boldsymbol{x}_1, r_1)$. Since $\boldsymbol{x} \in B$,

$$\|\boldsymbol{x} - \boldsymbol{x}_1\| \leq \|\boldsymbol{x} - \boldsymbol{x}_0\| + \|\boldsymbol{x}_0 - \boldsymbol{y}\| + \|\boldsymbol{y} - \boldsymbol{x}_1\| < r_0 + r_0 + r_1 \leq 3r_1.$$

Thus, $\boldsymbol{x} \in B(\boldsymbol{x}_1, 3r_1)$. ■

Remark 138 *In the previous lemma, we could have used closed balls.*

We prove the theorem.
Proof. Let

$$E_t := \left\{ \boldsymbol{x} \in \mathbb{R}^N : \mathrm{M}(f)(\boldsymbol{x}) > t \right\}$$

and let $K \subseteq E_t$ be a compact set. By the definition of $\mathrm{M}(f)$, for every $\boldsymbol{x} \in K$ we can find a ball $B(\boldsymbol{x}, r_{\boldsymbol{x}})$, with $r_{\boldsymbol{x}} > 0$, such that

$$\frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r_{\boldsymbol{x}}))} \int_{B(\boldsymbol{x}, r_{\boldsymbol{x}})} |f| \, d\boldsymbol{y} > t. \tag{56}$$

Since $K \subset \bigcup_{\boldsymbol{x} \in K} B(\boldsymbol{x}, r_{\boldsymbol{x}})$ (note that we are using the open balls), by compactness we can find a finite number of balls such that $K \subset \bigcup_{i=1}^n B(\boldsymbol{x}_i, r_{\boldsymbol{x}_i})$.

By the Vitali's covering lemma, we can find ℓ disjoint balls $B(\boldsymbol{y}_k, R_k)$ such that

$$K \subset \bigcup_{i=1}^n B(\boldsymbol{x}_i, r_{\boldsymbol{x}_i}) \subseteq \bigcup_{k=1}^{\ell} B(\boldsymbol{y}_k, 3R_k).$$

Hence, by (56) and the fact that the balls $B(\boldsymbol{y}_k, R_k)$ are disjoint,

$$\begin{aligned}
\mathcal{L}^N(K) &\leq \sum_{n=1}^{\ell} \mathcal{L}^N(B(\boldsymbol{y}_k, 3R_{\boldsymbol{k}})) = 3^N \sum_{n=1}^{\ell} \mathcal{L}^N(B(\boldsymbol{y}_k, R_{\boldsymbol{k}})) \leq 3^N \sum_{n=1}^{\ell} \frac{1}{t} \int_{B(\boldsymbol{y}_k, R_k)} |f| \, d\boldsymbol{y} \\
&= \frac{3^N}{t} \int_{\bigcup_k B(\boldsymbol{y}_k, R_k)} |f| \, d\boldsymbol{y} \leq \frac{3^N}{t} \int_{\mathbb{R}^N} |f| \, d\boldsymbol{y}.
\end{aligned}$$

By your homework E_t can be written as

$$E_t = \bigcup_{j=1}^{\infty} K_j \cup F,$$

where $K_j \subseteq K_{j+1}$ are compact and $\mathcal{L}^N(F) = 0$. By applying the previous inequality to K_j, we obtain that

$$\mathcal{L}^N(K_j) \leq \frac{3^N}{t} \int_{\mathbb{R}^N} |f| \, d\boldsymbol{y}.$$

Letting $j \to \infty$ we get

$$\mathcal{L}^N(E_t) = \lim_{j \to \infty} \mathcal{L}^N(K_j) \leq \frac{3^N}{t} \int_{\mathbb{R}^N} |f| \, d\boldsymbol{y}.$$

■

Exercise 139 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a Lebesgue integrable function. Prove that for every $\varepsilon > 0$ there exists $g \in C_c\left(\mathbb{R}^N\right)$ such that*

$$\int_{\mathbb{R}^N} |f - g| \, d\boldsymbol{y} \leq \varepsilon.$$

Theorem 140 (Lebesgue density theorem) *Let $f : \mathbb{R}^N \to \mathbb{R}$ be integrable. Then there exists a Borel set $E_0 \subset \mathbb{R}^N$, with $\mathcal{L}^N(E_0) = 0$, such that for every $\boldsymbol{x} \in \mathbb{R}^N \setminus E_0$,*

$$\lim_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} = 0. \tag{57}$$

In particular,

$$\lim_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} f(\boldsymbol{y}) \, d\boldsymbol{y} = f(\boldsymbol{x}) \tag{58}$$

for every $\boldsymbol{x} \in \mathbb{R}^N \setminus E_0$.

Proof. Given $\varepsilon > 0$, by the previous exercise, we may find a function $g \in C_c\left(\mathbb{R}^N\right)$ such that

$$\int_{\mathbb{R}^N} |f - g| \, d\boldsymbol{y} \leq \varepsilon.$$

Note that g depends on ε. Since g is continuous, for every $\boldsymbol{x} \in \mathbb{R}^N$ we have that

$$\lim_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |g(\boldsymbol{y}) - g(\boldsymbol{x})| \, d\boldsymbol{y} = 0. \tag{59}$$

Indeed, given $\eta > 0$ there exists $\delta = \delta(\eta) \in (0, 1)$ such that

$$|g(\boldsymbol{x}) - g(\boldsymbol{y})| \leq \eta$$

for all $\boldsymbol{y} \in \mathbb{R}^N$ with $\|\boldsymbol{x} - \boldsymbol{y}\| \leq \delta$. Hence, for $0 < r \leq \delta$,

$$\frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |g(\boldsymbol{y}) - g(\boldsymbol{x})| \, d\boldsymbol{y} \leq \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} \eta \, d\boldsymbol{y} = \eta.$$

For every $\boldsymbol{x} \in \mathbb{R}^N$, we have

$$\begin{aligned}
\frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} &\leq \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - g(\boldsymbol{y})| \, d\boldsymbol{y} \\
&\quad + \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |g(\boldsymbol{y}) - g(\boldsymbol{x})| \, d\boldsymbol{y} + |g(\boldsymbol{x}) - f(\boldsymbol{x})| \\
&\leq \mathrm{M}(f - g)(\boldsymbol{x}) + \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |g(\boldsymbol{y}) - g(\boldsymbol{x})| \, d\boldsymbol{y} + |g(\boldsymbol{x}) - f(\boldsymbol{x})|
\end{aligned}$$

Using (59), we have

$$\limsup_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} \leq \mathrm{M}(f - g)(\boldsymbol{x}) + 0 + |g(\boldsymbol{x}) - f(\boldsymbol{x})|.$$

For every $t > 0$, define

$$G_t := \left\{ \boldsymbol{x} \in \mathbb{R}^N : \limsup_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} > t \right\},$$
$$E_{t,\varepsilon} := \left\{ \boldsymbol{x} \in \mathbb{R}^N : \mathrm{M}(f - g)(\boldsymbol{x}) > t \right\},$$
$$F_{t,\varepsilon} := \left\{ \boldsymbol{x} \in \mathbb{R}^N : |g(\boldsymbol{x}) - f(\boldsymbol{x})| > t \right\}.$$

Then by the previous inequality, if $\boldsymbol{x} \in G_{2t}$, we have

$$2t < \limsup_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} \leq \mathrm{M}(f - g)(\boldsymbol{x}) + |g(\boldsymbol{x}) - f(\boldsymbol{x})|,$$

which implies that $\mathrm{M}(f - g)(\boldsymbol{x}) > t$ or $|g(\boldsymbol{x}) - f(\boldsymbol{x})| > t$. This shows that $G_{2t} \subseteq E_{t,\varepsilon} \cup F_{t,\varepsilon}$.

By (55),

$$\mathcal{L}^N(E_{t,\varepsilon}) \leq \frac{3^N}{t} \int_{\mathbb{R}^N} |f - g| \, d\boldsymbol{y} \leq \frac{3^N \varepsilon}{t},$$

while

$$\mathcal{L}^N(F_{t,\varepsilon}) \leq \frac{1}{t} \int_{\mathbb{R}^N} |f - g| \, d\boldsymbol{y} \leq \frac{\varepsilon}{t}.$$

Hence,

$$\mathcal{L}^N(G_{2t}) \leq \mathcal{L}^N(E_{t,\varepsilon}) + \mathcal{L}^N(F_{t,\varepsilon}) \leq \frac{\left(3^N + 1\right)\varepsilon}{t}.$$

Since G_{2t} does not depend on $\varepsilon > 0$, we can let $\varepsilon \to 0^+$ in the previous inequality to conclude that $\mathcal{L}^N(G_{2t}) = 0$ for all $t > 0$. Let

$$E_0 := \bigcup_{n=1}^{\infty} G_{\frac{1}{n}}.$$

Then $\mathcal{L}^N(E_0) = 0$ and if $\boldsymbol{x} \in \mathbb{R}^N \setminus E_0$, then

$$\limsup_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} \leq \frac{1}{n}$$

for every n, that is,

$$\limsup_{r \to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x}, r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})| \, d\boldsymbol{y} = 0,$$

which implies that (57) holds. ■

Exercise 141 *Prove that the theorem continues to hold if f is assumed to be locally integrable, that is, integrable on compact sets.*

A point $\boldsymbol{x} \in \mathbb{R}^N$ for which (57) holds is called *a Lebesgue point* of f.

Corollary 142 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be integrable and let f_ε be the mollification of f. Then $f_\varepsilon(\boldsymbol{x}) \to f(\boldsymbol{x})$ as $\varepsilon \to 0^+$ at every Lebesgue point $\boldsymbol{x} \in \mathbb{R}^N$ of f.*

Proof. Let $\boldsymbol{x}_0 \in \mathbb{R}^N$ be a Lebesgue point of f. By (53)

$$
\begin{aligned}
|f_\varepsilon(\boldsymbol{x}_0) - f(\boldsymbol{x}_0)| &\le \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi_\varepsilon(\boldsymbol{x}_0 - \boldsymbol{y})\, |f(\boldsymbol{y}) - f(\boldsymbol{x}_0)|\, d\boldsymbol{y} \\
&= \frac{1}{\varepsilon^N} \int_{B(\boldsymbol{x}_0,\varepsilon)} \varphi\left(\frac{(\boldsymbol{x}_0 - \boldsymbol{y})}{\varepsilon}\right) |f(\boldsymbol{y}) - f(\boldsymbol{x}_0)|\, d\boldsymbol{y} \\
&\le \frac{\sup_{\mathbb{R}^N} |\varphi|}{\varepsilon^N} \int_{B(\boldsymbol{x}_0,\varepsilon)} |f(\boldsymbol{y}) - f(\boldsymbol{x}_0)|\, d\boldsymbol{y} \\
&= \frac{\alpha_N \sup_{\mathbb{R}^N} |\varphi|}{\mathcal{L}^N(B(\boldsymbol{x}_0,\varepsilon))} \int_{B(\boldsymbol{x}_0,\varepsilon)} |f(\boldsymbol{y}) - f(\boldsymbol{x}_0)|\, d\boldsymbol{y} \to 0
\end{aligned}
$$

as $\varepsilon \to 0$. Here, α_N is the measure of the unit ball. ■

Friday, November 16, 2022

Remark 143 *If $\boldsymbol{x} \in \mathbb{R}^N \setminus E_0$, then given a family of Lebesgue measurable sets $\{E_{\boldsymbol{x},r}\}_{r>0}$ such that $E_{\boldsymbol{x},r} \subseteq B(\boldsymbol{x}, r)$ and*

$$
\mathcal{L}^N(E_{\boldsymbol{x},r}) \ge \alpha \mathcal{L}^N(B(\boldsymbol{x}, r))
$$

for some constant $\alpha > 0$ independent of $r > 0$, we have that

$$
\begin{aligned}
&\limsup_{r\to 0^+} \frac{1}{\mathcal{L}^N(E_{\boldsymbol{x},r})} \int_{E_{\boldsymbol{x},r}} |f(\boldsymbol{y}) - f(\boldsymbol{x})|\, d\boldsymbol{y} \\
&\qquad \le \limsup_{r\to 0^+} \frac{1}{\mathcal{L}^N(E_{\boldsymbol{x},r})} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})|\, d\boldsymbol{y} \\
&\qquad \le \frac{1}{\alpha} \lim_{r\to 0^+} \frac{1}{\mathcal{L}^N(B(\boldsymbol{x},r))} \int_{B(\boldsymbol{x},r)} |f(\boldsymbol{y}) - f(\boldsymbol{x})|\, d\boldsymbol{y} = 0.
\end{aligned}
$$

Note that the sets $E_{\boldsymbol{x},r}$ need not contain $\boldsymbol{x}$.

Lebesgue's density theorem allows us to give a different proof of Theorem 84.

Corollary 144 *Let $g : [a, b] \to \mathbb{R}$ be Lebesgue integrable and*

$$
f(x) = \int_a^x g(t)\, dt, \quad x \in [a, b].
$$

Then for $\mathcal{L}^1$-a.e. $x \in [a, b]$ the function f is differentiable and $f'(x) = g(x)$.

Proof. Extend g to be zero outside $[a, b]$. By Lebesgue's density theorem, for $\mathcal{L}^1$-a.e. $x \in \mathbb{R}$,

$$
\lim_{r\to 0^+} \frac{1}{2r} \int_{x-r}^{x+r} |g(y) - g(x)|\, dy = 0.
$$

In view of the previous remark, for every Lebesgue point $x \in \mathbb{R}$ of g,

$$\lim_{r\to 0^+} \frac{1}{r}\int_x^{x+r} |g(y) - g(x)|\, dy = 0,$$
$$\lim_{r\to 0^+} \frac{1}{r}\int_{x-r}^{x} |g(y) - g(x)|\, dy = 0.$$

In turn,

$$\lim_{r\to 0^+} \frac{f(x+r) - f(x)}{r} = \lim_{r\to 0^+} \frac{1}{r}\int_x^{x+r} g(y)\, dy = g(x),$$
$$\lim_{r\to 0^+} \frac{f(x) - f(x-r)}{r} = \lim_{r\to 0^+} \int_{x-r}^{x} g(y)\, dy = g(x).$$

Hence, f is differentiable at x and $f'(x) = g(x)$. ■

An important application of the theory of mollifiers is the existence of smooth partitions of unity.

Theorem 145 (Smooth partition of unity) *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\{V_\alpha\}_{\alpha\in\Lambda}$ be an open cover of U. Then there exists a sequence $\{\psi_n\}_n$ of nonnegative functions in in $C_c^\infty\left(\mathbb{R}^N\right)$ such that*

(i) each ψ_n has support in some $V_\alpha \cap U$;

(ii) $\sum_{n=1}^{\infty} \psi_n(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in U$;

(iii) for every compact set $K \subset U$ there exists an integer $\ell \in \mathbb{N}$ and an open set V, with $K \subset V \subseteq U$, such that

$$\sum_{n=1}^{\ell} \psi_n(\boldsymbol{x}) = 1$$

for all $\boldsymbol{x} \in V$.

Proof. Let S be a countable dense set in U, for example, $S := \{\boldsymbol{x} \in \mathbb{Q}^N \cap U\}$, and consider the countable family $\mathcal{F}$ of closed balls

$$\mathcal{F} := \{\overline{B(\boldsymbol{x}, r)} : r \in (0,1) \cap \mathbb{Q},\ \boldsymbol{x} \in S,\ \overline{B(\boldsymbol{x}, r)} \subset V_\alpha \cap U \text{ for some } \alpha \in \Lambda\}.$$

Since $\mathcal{F}$ is countable we may write $\mathcal{F} = \{\overline{B(\boldsymbol{x}_n, r_n)} : n \in \mathbb{N}\}$. We claim that

$$U = \bigcup_{n=1}^{\infty} B\left(\boldsymbol{x}_n, \tfrac{r_n}{2}\right). \tag{60}$$

Indeed, given $\boldsymbol{x} \in U$, since $\{V_\alpha\}_{\alpha\in\Lambda}$ is an open cover of U, we can find V_β such that $\boldsymbol{x} \in V_\beta$. Since $V_\beta \cap U$ is open, there exists $0 < r < 1$ such that $B(\boldsymbol{x}, r) \subseteq V_\beta \cap U$. By the density of $\mathbb{Q}$ in $\mathbb{R}$ there exist $\boldsymbol{y} \in \mathbb{Q}^N$ such that

$\|\boldsymbol{x}-\boldsymbol{y}\| < \frac{r}{8}$ and $q \in \mathbb{Q}$ such that $\frac{6r}{8} < q < \frac{7r}{8}$. Then $\overline{B(\boldsymbol{y}, q)} \subset B(\boldsymbol{x}, r)$, since if $\boldsymbol{z} \in \overline{B(\boldsymbol{y}, q)}$, then

$$\|\boldsymbol{z}-\boldsymbol{x}\| \leq \|\boldsymbol{z}-\boldsymbol{y}\| + \|\boldsymbol{x}-\boldsymbol{y}\| < q + \frac{r}{8} < \frac{7r}{8} + \frac{r}{8} = r.$$

Hence, $\overline{B(\boldsymbol{y}, q)} \in \mathcal{F}$. Moreover, $\|\boldsymbol{x}-\boldsymbol{y}\| < \frac{r}{8} < \frac{q}{2}$ and so $\boldsymbol{x} \in B\left(\boldsymbol{y}, \frac{q}{2}\right)$. This shows that $U \subseteq \bigcup_{n=1}^{\infty} B\left(\boldsymbol{x}_n, \frac{r_n}{2}\right)$. The other inclusion follows from the fact that each ball in $\mathcal{F}$ is contained in U.

For each $n \in \mathbb{N}$ consider

$$\phi_n := \varphi_{\frac{r_n}{4}} * \chi_{B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right)},$$

where $\varphi_{\frac{r_n}{4}}$ are standard mollifiers (with $\varepsilon := \frac{r_n}{4}$). By Theorem 132 $\phi_n \in C^{\infty}\left(\mathbb{R}^N\right)$. Moreover, if $\boldsymbol{x} \in B\left(\boldsymbol{x}_n, \frac{r_n}{2}\right)$, then

$$\begin{aligned}
\phi_n(\boldsymbol{x}) &= \int_{\mathbb{R}^N} \varphi_{\frac{r_n}{4}}(\boldsymbol{x}-\boldsymbol{y}) \chi_{B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right)}(\boldsymbol{y}) \, d\boldsymbol{y} \\
&= \int_{B\left(\boldsymbol{x}, \frac{r_n}{4}\right)} \varphi_{\frac{r_n}{4}}(\boldsymbol{x}-\boldsymbol{y}) \chi_{B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right)}(\boldsymbol{y}) \, d\boldsymbol{y} \\
&= \int_{B\left(\boldsymbol{x}, \frac{r_n}{4}\right)} \varphi_{\frac{r_n}{4}}(\boldsymbol{x}-\boldsymbol{y}) \, d\boldsymbol{y} = 1,
\end{aligned}$$

where we have used (51) and the fact that if $\boldsymbol{x} \in B\left(\boldsymbol{x}_n, \frac{r_n}{2}\right)$, then $B\left(\boldsymbol{x}, \frac{r_n}{4}\right) \subset B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right)$. Since $0 \leq \chi_{B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right)} \leq 1$ a similar calculation shows that $0 \leq \phi_n \leq 1$. On the other hand, if $\boldsymbol{x} \notin \overline{B(\boldsymbol{x}_n, r_n)}$, then

$$\begin{aligned}
\phi_n(\boldsymbol{x}) &= \int_{\mathbb{R}^N} \varphi_{\frac{r_n}{4}}(\boldsymbol{x}-\boldsymbol{y}) \chi_{B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right)}(\boldsymbol{y}) \, d\boldsymbol{y} \\
&= \int_{B\left(\boldsymbol{x}, \frac{r_n}{4}\right)} \varphi_{\frac{r_n}{4}}(\boldsymbol{x}-\boldsymbol{y}) \chi_{B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right)}(\boldsymbol{y}) \, d\boldsymbol{y} = 0,
\end{aligned}$$

where we have used the fact that if $\boldsymbol{x} \notin \overline{B(\boldsymbol{x}_n, r_n)}$, then $B\left(\boldsymbol{x}, \frac{r_n}{4}\right) \cap B\left(\boldsymbol{x}_n, \frac{3}{4} r_n\right) = \emptyset$. In particular, $\phi_n \in C_c^{\infty}\left(\mathbb{R}^N\right)$ and $\operatorname{supp} \phi_n \subset \overline{B(\boldsymbol{x}_n, r_n)}$. Note that in view of the definition of $\mathcal{F}$, $\operatorname{supp} \phi_n \subset V_\alpha \cap U$ for some $\alpha \in \Lambda$. ■

Monday, November 22, 2022

Proof. Define $\psi_1 := \phi_1$ and

$$\psi_n := (1-\phi_1) \cdots (1-\phi_{n-1}) \phi_n \tag{61}$$

for $n \geq 2$, $n \in \mathbb{N}$. Since $0 \leq \phi_k \leq 1$ and and $\operatorname{supp} \phi_k \subset \overline{B(\boldsymbol{x}_k, r_k)}$ for all $k \in \mathbb{N}$ we have that $0 \leq \psi_n \leq 1$ and $\operatorname{supp} \psi_n \subset \overline{B(\boldsymbol{x}_n, r_n)}$. This gives (i). To prove (ii) we prove by induction that

$$\psi_1 + \cdots + \psi_n = 1 - (1-\phi_1) \cdots (1-\phi_n) \tag{62}$$

for all $n \in \mathbb{N}$. The relation (62) is true for $n = 1$, since $\psi_1 := \phi_1$. Assume that (62) holds for n, then by (61)

$$\begin{aligned}\psi_1 + \cdots + \psi_n + \psi_{n+1} &= 1 - (1-\phi_1)\cdots(1-\phi_n) + \psi_{n+1} \\ &= 1 - (1-\phi_1)\cdots(1-\phi_n) + (1-\phi_1)\cdots(1-\phi_n)\,\phi_{n+1} \\ &= 1 - (1-\phi_1)\cdots(1-\phi_{n+1})\,.\end{aligned}$$

Hence (62) holds for all $n \in \mathbb{N}$.

Since $\phi_k = 1$ in $B\left(\boldsymbol{x}_k, \frac{r_k}{2}\right)$ for all $k \in \mathbb{N}$ it follows that from (62) that

$$\psi_1(\boldsymbol{x}) + \cdots + \psi_n(\boldsymbol{x}) = 1 \quad \text{for all } \boldsymbol{x} \in \bigcup_{k=1}^{n} B\left(\boldsymbol{x}_k, \tfrac{r_k}{2}\right). \tag{63}$$

Thus, in view of (60) property (ii) holds.

Finally, if $K \subset U$ is compact, again by (60), we may find $\ell \in \mathbb{N}$ so large that

$$\bigcup_{k=1}^{\ell} B\left(\boldsymbol{x}_k, \tfrac{r_k}{2}\right) \supset K$$

and so (iii) follows by (63). ■

12 Divergence Theorem

Given $i \in \{1, \ldots, N\}$ and $\boldsymbol{x} \in \mathbb{R}^N$ let $\boldsymbol{x}_i \in \mathbb{R}^{N-1}$ be the vector obtained from $\boldsymbol{x}$ by removing the i-th component x_i of $\boldsymbol{x}$. With an abuse of notation we write $\boldsymbol{x} = (\boldsymbol{x}_i, x_i) \in \mathbb{R}^{N-1} \times \mathbb{R}$. When $i = N$ we use the usual notation $\boldsymbol{x} = (\boldsymbol{x}', x_N) \in \mathbb{R}^{N-1} \times \mathbb{R}$.

Definition 146 *Given an open set $U \subseteq \mathbb{R}^N$ we say that its boundary ∂U is of* class C^m, $m \in \mathbb{N}$ *if for every $\boldsymbol{x}_0 \in \partial U$ there exist $i \in \{1, \ldots, N\}$, $r > 0$, and a function $h : \mathbb{R}^{N-1} \to \mathbb{R}$ of class C^m such that, writing $\boldsymbol{x} = (\boldsymbol{x}_i, x_i)$, we have either*

$$U \cap B(\boldsymbol{x}_0, r) := \{\boldsymbol{x} \in B(\boldsymbol{x}_0, r) : \, h(\boldsymbol{x}_i) < x_i\}$$

or

$$U \cap B(\boldsymbol{x}_0, r) := \{\boldsymbol{x} \in B(\boldsymbol{x}_0, r) : \, h(\boldsymbol{x}_i) > x_i\}\,.$$

Note that i, h, and r depend on $\boldsymbol{x}_0$. If ∂U is of class C^m for $m \in \mathbb{N}$, then ∂U is an $(N-1)$-dimensional surface of class C^m. Also, instead of balls we can use cubes. Note that $T_{\partial U}(\boldsymbol{x})$ is given by the $N-1$ dimensional vector space given by $\ker \nabla g$, where $g(\boldsymbol{x}) = x_i - h(\boldsymbol{x}_i)$.

Definition 147 *Given an open set $U \subseteq \mathbb{R}^N$ with boundary of class C^m, $m \in \mathbb{N}$, a unit normal vector $\boldsymbol{\nu}$ to ∂U at $\boldsymbol{x}_0$ is called a* unit outward normal *to U at $\boldsymbol{x}_0$ if there exists $\delta > 0$ such that $\boldsymbol{x}_0 - t\boldsymbol{\nu} \in U$ and $\boldsymbol{x}_0 + t\boldsymbol{\nu} \in \mathbb{R}^N \setminus U$ for all $0 < t < \delta$.*

We are ready to prove the divergence theorem.

Theorem 148 (Divergence Theorem) *Let $U \subset \mathbb{R}^N$ be an open, bounded set with boundary of class C^1 and let $\boldsymbol{f} : \overline{U} \to \mathbb{R}^N$ be such that $\boldsymbol{f}$ is continuous in $\overline{U}$ and there exist the partial derivatives of $\boldsymbol{f}$ in $\mathbb{R}^N$ at all $\boldsymbol{x} \in U$ and they are continuous and bounded. Then*

$$\int_U \operatorname{div} \boldsymbol{f}\,(\boldsymbol{x})\;d\boldsymbol{x} = \int_{\partial U} \boldsymbol{f} \cdot \boldsymbol{\nu}\,d\mathcal{H}^{N-1},$$

where

$$\operatorname{div} \boldsymbol{f} := \sum_{i=1}^{N} \frac{\partial f_i}{\partial x_i}.$$

Remark 149 *In physics $\int_{\partial U} \boldsymbol{f}\,(\boldsymbol{x}) \cdot \boldsymbol{\nu}\,(\boldsymbol{x})\;d\mathcal{H}^{N-1}\,(\boldsymbol{x})$ represents the* outward flux *of a vector field $\boldsymbol{f}$ across the boundary of a region U.*

Corollary 150 (Integration by Parts) *Let $U \subset \mathbb{R}^N$ be an open, bounded, set with boundary of class C^1 and let $f : \overline{U} \to \mathbb{R}$ and $g : \overline{U} \to \mathbb{R}$ be such that f and g are continuous in $\overline{U}$ and there exist the partial derivatives of f and g in $\mathbb{R}$ at all $\boldsymbol{x} \in U$ and they are continuous and bounded. Then for every $i = 1, \ldots, N$,*

$$\int_U f\,(\boldsymbol{x})\,\frac{\partial g}{\partial x_i}\,(\boldsymbol{x})\;d\boldsymbol{x} = -\int_U g\,(\boldsymbol{x})\,\frac{\partial f}{\partial x_i}\,(\boldsymbol{x})\;d\boldsymbol{x} + \int_{\partial U} f g \nu_i\,d\mathcal{H}^{N-1}.$$

Proof. Fix $i \in \{1, \ldots, N\}$. We apply the divergence theorem to the function $\boldsymbol{f} : \overline{U} \to \mathbb{R}^N$ defined by

$$f_j\,(\boldsymbol{x}) := \begin{cases} f\,(\boldsymbol{x})\,g\,(\boldsymbol{x}) & \text{if } j = i, \\ 0 & \text{if } j \neq i. \end{cases}$$

Then

$$\operatorname{div} \boldsymbol{f} = \sum_{j=1}^{N} \frac{\partial f_j}{\partial x_j} = \frac{\partial\,(fg)}{\partial x_i} = f\frac{\partial g}{\partial x_i} + g\frac{\partial f}{\partial x_i},$$

and so

$$\int_U \left(f\frac{\partial g}{\partial x_i} + g\frac{\partial f}{\partial x_i}\right)\,d\boldsymbol{x} = \int_U \operatorname{div} \boldsymbol{f}\;d\boldsymbol{x} = \int_{\partial U} \boldsymbol{f} \cdot \boldsymbol{\nu}\,d\mathcal{H}^{N-1} = \int_{\partial U} f g \nu_i\,d\mathcal{H}^{N-1}.$$

■

If $E \subseteq \mathbb{R}^N$ and $\boldsymbol{f} : E \to \mathbb{R}^N$ is differentiable, then $\boldsymbol{f}$ is called a *divergence-free field* or *solenoidal field* if

$$\operatorname{div} \boldsymbol{f} = 0.$$

Thus for a smooth solenoidal field, the outward flux across the boundary of a regular set U is zero. Examples of solenoidal fields are the magnetic field in Maxwell's equations, the velocity of an incompressible fluid, the vorticity.

Monday, November 28, 2022

Exercise 151 *Calculate the outward flux of the function*

$$\boldsymbol{f}\,(x,y,z) := (0, yz, x)$$

across the boundary of the region

$$U := \left\{(x,y,z) \in \mathbb{R}^3 :\, x^2+y^2 < z^2,\, x^2+y^2+z^2 < 2y,\, z > 0\right\}.$$

Definition 152 *Given an open set $U \subseteq \mathbb{R}^N$ and an integer $m \in \mathbb{N}$, we say that a function $f : \overline{U} \to \mathbb{R}$ is of class $C^m\left(\overline{U}\right)$ if f can be extended to a function of class $C^m(V)$, where V is an open set containing $\overline{U}$.*

Lemma 153 *Let $R := (a_1,b_1) \times \cdots \times (a_N,b_N)$ and $\boldsymbol{f} : \overline{R} \to \mathbb{R}^N$ be such that $\boldsymbol{f}$ is continuous in $\overline{R}$ and there exist the partial derivatives of $\boldsymbol{f}$ at all $\boldsymbol{x} \in R$ and they are continuous and bounded. Then*

$$\int_R \operatorname{div} \boldsymbol{f}\,(\boldsymbol{x})\,d\boldsymbol{x} = \int_{\partial R} \boldsymbol{f} \cdot \boldsymbol{\nu}\, d\mathcal{H}^{N-1},$$

Proof. Given $k \in \{1,\ldots,N\}$ write $R_k := \Pi_{i\neq k}(a_i,b_i) \subset \mathbb{R}^{N-1}$ and $\boldsymbol{x} = (\boldsymbol{x}_k, x_k)$, where $\boldsymbol{x}_k \in \mathbb{R}^{N-1}$ is the vector obtained by removing the k-th component from $\boldsymbol{x}$. Then by Fubini's theorem,

$$\begin{aligned}
\int_R \frac{\partial f_k}{\partial x_k}(\boldsymbol{x})\,d\boldsymbol{x} &= \int_{R_k}\left(\int_{a_k}^{b_k}\frac{\partial f_k}{\partial x_k}(\boldsymbol{x}_k,x_k)\,dx_k\right)d\boldsymbol{x}_k \\
&= \int_{R_k}\left(f_k(\boldsymbol{x}_k,b_k) - f_k(\boldsymbol{x}_k,a_k)\right)d\boldsymbol{x}_k \\
&= \int_{R_k}\boldsymbol{f}\,(\boldsymbol{x}_k,b_k)\cdot\boldsymbol{e}_k\,d\boldsymbol{x}_k + \int_{R_k}\boldsymbol{f}\,(\boldsymbol{x}_k,a_k)\cdot(-\boldsymbol{e}_k)\,d\boldsymbol{x}_k \\
&= \int_{R_k\times\{b_k\}}\boldsymbol{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^{N-1} + \int_{R_k\times\{a_k\}}\boldsymbol{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^{N-1},
\end{aligned}$$

where in the third equality we have used the fundamental theorem of calculus applied to the function of one variable $x_k \in [a_k,b_k] \mapsto f_k(\boldsymbol{x}_k,x_k)$ with $\boldsymbol{x}_k$ fixed, and in the last equality we have used formula (47). Summing the resulting identities gives the desired result. ■

Lemma 154 *Let*

$$U := \{(\boldsymbol{x}',x_N) \in R_N \times \mathbb{R} :\, h(\boldsymbol{x}') < x_N < b_N\},$$

where $R_N = (a_1,b_1) \times \cdots \times (a_{N-1},b_{N-1})$, $h : \overline{R_N} \to \mathbb{R}$ is of class C^1 and $\max_{\overline{R_N}} h < b_N$ *and let $\boldsymbol{f} : \overline{U} \to \mathbb{R}^N$ be such that $\boldsymbol{f}$ is continuous in $\overline{U}$ and there exist the partial derivatives of $\boldsymbol{f}$ at all $\boldsymbol{x} \in U$ and they are continuous and bounded. Assume that there exists an open set V containing $\partial U \setminus \operatorname{graph} h$ such that $\boldsymbol{f} = \boldsymbol{0}$ in $\overline{U} \cap V$. Then*

$$\int_U \operatorname{div} \boldsymbol{f}\,(\boldsymbol{x})\,d\boldsymbol{x} = \int_{\operatorname{Gr} h} \boldsymbol{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^{N-1},$$

Proof. Since $\max_{\overline{R_N}} h < b_N$ and there exists an open set V containing $\partial U \setminus \operatorname{graph} h$ such that $\boldsymbol{f} = \mathbf{0}$ in $\overline{U} \cap V$, we have that $\boldsymbol{f} = \mathbf{0}$ in $\overline{R_N} \times (b_N - \delta, b_N)$. Hence, if we define $\boldsymbol{f}(\mathbf{x}) = \mathbf{0}$ for $x_N > b_N$, we have that $\boldsymbol{f}$ is continuous in the closed set

$$\left\{(\boldsymbol{x}', x_N) \in \overline{R_N} \times \mathbb{R} : \, h(\boldsymbol{x}') \leq x_N < \infty\right\} \tag{64}$$

and the partial derivatives are continuous and bounded in the open set

$$\{(\boldsymbol{x}', x_N) \in R_N \times \mathbb{R} : \, -\infty < x_N < h(\boldsymbol{x}')\}. \tag{65}$$

Since a chart for $\operatorname{graph} h$ a chart is given by $\boldsymbol{\varphi}(\boldsymbol{x}') = (\boldsymbol{x}', h(\boldsymbol{x}'))$, by Theorem 121, a basis for the tangent space is given by $\frac{\partial \boldsymbol{\varphi}}{\partial x_1}(\boldsymbol{x}')$, $\ldots$, $\frac{\partial \boldsymbol{\varphi}}{\partial x_{N-1}}(\boldsymbol{x}')$. Note that

$$\frac{\partial \boldsymbol{\varphi}}{\partial x_i}(\boldsymbol{x}') = \left(\boldsymbol{e}_i', \frac{\partial h}{\partial x_i}(\boldsymbol{x}')\right)$$

and so a vector orthogonal to all $\frac{\partial \boldsymbol{\varphi}}{\partial x_i}(\boldsymbol{x}')$, $i = 1, \ldots, N-1$, is $(\nabla h(\boldsymbol{x}'), -1)$. Hence, the unit normal is

$$\pm \frac{(\nabla h(\boldsymbol{x}'), -1)}{\sqrt{1 + \|\nabla h(\boldsymbol{x}')\|_{N-1}^2}}.$$

Assume that the domain is of the the type (64). Then,

$$\begin{aligned}\int_{\partial U} \boldsymbol{f} \cdot \boldsymbol{\nu} \, d\mathcal{H}^{N-1} &= \int_{\operatorname{Gr} h} \boldsymbol{f} \cdot \boldsymbol{\nu} \, d\mathcal{H}^{N-1} = \int_{R_N} \boldsymbol{f}\,(\boldsymbol{x}', h(\boldsymbol{x}')) \cdot (\nabla h(\boldsymbol{x}'), -1) \frac{\sqrt{1 + \|\nabla h(\boldsymbol{x}')\|_{N-1}^2}}{\sqrt{1 + \|\nabla h(\boldsymbol{x}')\|_{N-1}^2}} \, d\boldsymbol{x}' \\ &= -\int_{R_N} f_N(\boldsymbol{x}', h(\boldsymbol{x}')) \, d\boldsymbol{x}' + \sum_{i=1}^{N-1} \int_{R_N} f_i(\boldsymbol{x}', h(\boldsymbol{x}')) \frac{\partial h}{\partial x_i}(\boldsymbol{x}') \, d\boldsymbol{x}'.\end{aligned}$$

■

Wednesday, November 30, 2022

Proof. The change of variables

$$y_N := x_N - h(\boldsymbol{x}'), \quad \boldsymbol{y}' := \boldsymbol{x}'$$

maps the set U into the set

$$W := \{(\boldsymbol{y}', y_N) \in R_N \times \mathbb{R} : \, 0 < y_N < b_N - h(\boldsymbol{y}')\}.$$

Let $R := R_N \times (0, c_N)$, where $c_N > b_N + \max_{\overline{R_N}} h$, and consider the function

$$\boldsymbol{g}(\boldsymbol{y}) := \boldsymbol{f}\,(\boldsymbol{y}', y_N + h(\boldsymbol{y}'))$$

By the chain rule, if $\boldsymbol{y} \in W$,

$$\begin{aligned}
\operatorname{div} \boldsymbol{g}(\boldsymbol{y}) &= \frac{\partial f_N}{\partial x_N}(\boldsymbol{y}', y_N + h(\boldsymbol{y}')) \\
&\quad + \sum_{i=1}^{N-1} \left(\frac{\partial f_i}{\partial x_N}(\boldsymbol{y}', y_N + h(\boldsymbol{y}')) \frac{\partial h}{\partial y_i}(\boldsymbol{y}') + \frac{\partial f_i}{\partial x_i}(\boldsymbol{y}', y_N + h(\boldsymbol{y}')) \right) \\
&= \operatorname{div} \boldsymbol{f}(\boldsymbol{y}', y_N + h(\boldsymbol{y}')) + \sum_{i=1}^{N-1} \frac{\partial f_i}{\partial x_N}(\boldsymbol{y}', y_N + h(\boldsymbol{y}')) \frac{\partial h}{\partial y_i}(\boldsymbol{y}').
\end{aligned}$$

By Step 1 applied to $\boldsymbol{g}$ in the set R,

$$\int_W \operatorname{div} \boldsymbol{g}(\boldsymbol{y})\, d\boldsymbol{y} = \int_R \operatorname{div} \boldsymbol{g}(\boldsymbol{y})\, d\boldsymbol{y} = \int_{\partial R} \boldsymbol{g}(\boldsymbol{y}) \cdot \boldsymbol{\nu}(\boldsymbol{y})\, d\mathcal{H}^{N-1}(\boldsymbol{y}) = -\int_{R_N} g_N(\boldsymbol{y}', 0)\, d\boldsymbol{y}'. \tag{66}$$

Hence, we can rewrite (66) as

$$\begin{aligned}
\int_W \operatorname{div} \boldsymbol{f}(\boldsymbol{y}', y_N + h(\boldsymbol{y}'))\, d\boldsymbol{y} = &-\int_{R_N} f_N(\boldsymbol{y}', h(\boldsymbol{y}'))\, d\boldsymbol{y}' \\
&- \sum_{i=1}^{N-1} \int_W \frac{\partial f_i}{\partial x_N}(\boldsymbol{y}', y_N + h(\boldsymbol{y}')) \frac{\partial h}{\partial y_i}(\boldsymbol{y}')\, d\boldsymbol{y}.
\end{aligned}$$

Consider the change of variables

$$\begin{aligned}
\mathbf{k} : \mathbb{R}^N &\to \mathbb{R}^N, \\
\boldsymbol{y} &\mapsto (\boldsymbol{y}', y_N + h(\boldsymbol{y}')).
\end{aligned}$$

Note that $\mathbf{k}$ is invertible, with inverse given by

$$\begin{aligned}
\mathbf{k}^{-1} : \mathbb{R}^N &\to \mathbb{R}^N \\
\boldsymbol{x} &\mapsto (\boldsymbol{x}', x_N - h(\boldsymbol{x}')).
\end{aligned}$$

Moreover, we have

$$J_{\mathbf{k}}(\boldsymbol{y}) = \begin{pmatrix} & & & 0 \\ & I_{N-1} & & \vdots \\ & & & 0 \\ -\frac{\partial h}{\partial y_2}(\boldsymbol{y}') & \cdots & -\frac{\partial h}{\partial y_N}(\boldsymbol{y}') & 1 \end{pmatrix},$$

which implies that $\det J_{\mathbf{k}}(\boldsymbol{y}) = 1$. Hence, by changing variables (see Theorem **??**),

$$\int_U \operatorname{div} \boldsymbol{f}(\boldsymbol{x})\, d\boldsymbol{x} = -\sum_{i=1}^{N-1} \int_U \frac{\partial f_i}{\partial x_N}(\boldsymbol{x}) \frac{\partial h}{\partial x_i}(\boldsymbol{x}')\, d\boldsymbol{x} - \int_{R_N} f_N(\boldsymbol{x}', h(\boldsymbol{x}'))\, d\boldsymbol{x}'.$$

By the fundamental theorem of calculus,

$$\begin{aligned}\int_U \frac{\partial f_i}{\partial x_N}(\boldsymbol{x})\frac{\partial h}{\partial x_i}(\boldsymbol{x}')\,d\boldsymbol{x} &= \int_{R_N}\left(\int_{h(\boldsymbol{x}')}^{b_N}\frac{\partial f_i}{\partial x_N}(\boldsymbol{x}',x_N)\,dx_N\right)\frac{\partial h}{\partial x_i}(\boldsymbol{x}')\,d\boldsymbol{x}'\\ &= \int_{R_N}(f_i(\boldsymbol{x}',b_N)-f_i(\boldsymbol{x}',h(\boldsymbol{x}')))\frac{\partial h}{\partial x_i}(\boldsymbol{x}')\,d\boldsymbol{x}'\\ &= \int_{R_N}(0-f_i(\boldsymbol{x}',h(\boldsymbol{x}')))\frac{\partial h}{\partial x_i}(\boldsymbol{x}')\,d\boldsymbol{x},\end{aligned}$$

and so

$$\begin{aligned}&\int_U \operatorname{div}\boldsymbol{f}(\boldsymbol{x})\,d\boldsymbol{x} = -\int_{R_N} f_N(\boldsymbol{x}',h(\boldsymbol{x}'))\,d\boldsymbol{x}' + \sum_{i=1}^{N-1}\int_{R_N} f_i(\boldsymbol{x}',h(\boldsymbol{x}'))\frac{\partial h}{\partial x_i}(\boldsymbol{x}')\,d\boldsymbol{x}'\\ &= \int_{\partial U}\boldsymbol{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^{N-1}(\boldsymbol{x}),\end{aligned}$$

where in the last equality we have used formula (47). This concludes the proof in this case. ■

Friday, December 2, 2022

We turn to the proof of the divergence theorem

Proof. By Definition 146 for every $\boldsymbol{x}_0 \in \partial U$ there exist $i \in \{1,\dots,N\}$, $R_{\boldsymbol{x}_0} > 0$, and a function $h_{\boldsymbol{x}_0} : \mathbb{R}^{N-1} \to \mathbb{R}$ of class C^1 such that

$$U \cap B(\boldsymbol{x}_0, R_{\boldsymbol{x}_0}) := \{\boldsymbol{x} \in B(\boldsymbol{x}_0, R_{\boldsymbol{x}_0}) : \ h_{\boldsymbol{x}_0}(\boldsymbol{x}_i) < x_i\} \tag{67}$$

or

$$U \cap B(\boldsymbol{x}_0, R_{\boldsymbol{x}_0}) := \{\boldsymbol{x} \in B(\boldsymbol{x}_0, R_{\boldsymbol{x}_0}) : \ h_{\boldsymbol{x}_0}(\boldsymbol{x}_i) > x_i\} \tag{68}$$

By continuity, find $0 < r_{\boldsymbol{x}_0} < \frac{R_{\boldsymbol{x}_0}}{\sqrt{N-1}}$ such that

$$|h_{\boldsymbol{x}_0}(\boldsymbol{x}_i) - h_{\boldsymbol{x}_0}(\boldsymbol{x}_{0,i})| < \frac{R_{\boldsymbol{x}_0}}{4} \tag{69}$$

for all $\boldsymbol{x}_i \in B_{N-1}(\boldsymbol{x}_{0,i}, r_{\boldsymbol{x}_0})$. On the other hand, if $\boldsymbol{x} \in U$, which is open, then there exists $B(\boldsymbol{x}, r_{\boldsymbol{x}}) \subseteq U$.

Since $\overline{U} \subseteq \bigcup_{\boldsymbol{x}\in\overline{U}} B(\boldsymbol{x}, r_{\boldsymbol{x}}/2)$, by compactness, we can find finitely many balls that cover $\overline{U}$. Let $\psi_1, \dots, \psi_n$ be a partition of unity subordinated to the family of open balls $B(\boldsymbol{x}^1, r_1/2), \dots, B(\boldsymbol{x}^n, r_n/2)$. For every $k = 1, \dots, n$, if $B(\boldsymbol{x}^k, r_k) \subseteq U$, we consider the rectangle T of side-length $r_{\boldsymbol{x}_k}$ centered at $\boldsymbol{x}^k$. Since $T \subset B(\boldsymbol{x}^k, r_{\boldsymbol{x}_k})$ and the function $\psi_k \boldsymbol{f}$ has compact support contained in $B(\boldsymbol{x}^k, r_k/2)$, it is zero in $\overline{T} \setminus B(\boldsymbol{x}^k, r_k/2)$. Thus we can apply Step 1 to $\psi_k \boldsymbol{f}$ in T to conclude that

$$\int_U \operatorname{div}(\psi_k \boldsymbol{f})\,d\boldsymbol{x} = \int_T \operatorname{div}(\psi_k \boldsymbol{f})\,d\boldsymbol{x} = \int_{\partial T}\psi_k \boldsymbol{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^{N-1} = 0 \tag{70}$$

since $\psi_k \boldsymbol{f} = \boldsymbol{0}$ on ∂T.

On the other hand, if $U \cap B\left(\boldsymbol{x}^k, R_k\right)$ is of the form (67) or (68), using the notation $\boldsymbol{x} = (\boldsymbol{x}_i, x_i)$ and $\boldsymbol{x}^k = (\boldsymbol{x}_i^k, h_k\left(\boldsymbol{x}_i^k\right))$, define the rectangle

$$T = \left(\boldsymbol{x}_i^k + \left(-\frac{r_k}{2}, \frac{r_k}{2}\right)^{N-1}\right) \times \left(h_k\left(\boldsymbol{x}_i^k\right) - \frac{R_k}{2}, h_k\left(\boldsymbol{x}_i^k\right) + \frac{R_k}{2}\right).$$

Note that $T \subseteq B\left(\boldsymbol{x}^k, R_k\right)$ since if $\boldsymbol{x} \in T$, then

$$\|\boldsymbol{x}-\boldsymbol{x}^k\|^2 = \sum_{j\neq i}(x_j - x_j^k)^2 + (x_i - h_k\left(\boldsymbol{x}_i^k\right))^2 < (N-1)\frac{r_k^2}{4} + \frac{R_k^2}{4} < \frac{R_k^2}{4} + \frac{R_k^2}{4} < R_k^2.$$

Thus, by (67),

$$U \cap T := \left\{\boldsymbol{x} : \boldsymbol{x}_i \in \boldsymbol{x}_i^k + \left(-\frac{r_k}{2}, \frac{r_k}{2}\right)^{N-1}, h_k\left(\boldsymbol{x}_i\right) < x_i < h_k\left(\boldsymbol{x}_i^k\right) + \frac{R_k}{2}\right\}$$

or

$$U \cap T := \left\{\boldsymbol{x} : \boldsymbol{x}_i \in \boldsymbol{x}_i^k + \left(-\frac{r_k}{2}, \frac{r_k}{2}\right)^{N-1}, h_k\left(\boldsymbol{x}_i\right) - \frac{R_k}{2} < x_i < h_k\left(\boldsymbol{x}_i^k\right)\right\}$$

Moreover, if $\boldsymbol{x}_i \in \boldsymbol{x}_i^k + \left(-\frac{r_k}{2}, \frac{r_k}{2}\right)^{N-1}$, then by (69), $h_k\left(\boldsymbol{x}_i\right) \leq h_k\left(\boldsymbol{x}_i^k\right) + \frac{R_k}{4} < h_k\left(\boldsymbol{x}_i^k\right) + \frac{R_k}{2}$. Since the function $\psi_k \boldsymbol{f}$ has compact support contained in $B\left(\boldsymbol{x}^k, r_k/2\right)$, it is zero on an open set that contains in $\partial T \setminus \operatorname{graph} h_k$. Thus, we can apply Lemma 154 to conclude that

$$\begin{aligned}\int_U \operatorname{div}\left(\psi_k \boldsymbol{f}\right) d\boldsymbol{x} &= \int_T \operatorname{div}\left(\psi_k \boldsymbol{f}\right) d\boldsymbol{x} = \int_{\boldsymbol{T}_{\boldsymbol{x}_k}(\operatorname{graph} h_{\boldsymbol{x}_k})} \psi_k \boldsymbol{f} \cdot \boldsymbol{\nu}\, d\mathcal{H}^{N-1} \\ &= \int_{\partial U} \psi_k \boldsymbol{f} \cdot \boldsymbol{\nu}\, d\mathcal{H}^{N-1},\end{aligned} \tag{71}$$

where we have used the fact that $\psi_k \boldsymbol{f}$ is zero outside $B\left(\boldsymbol{x}_k, r_{\boldsymbol{x}_k}/2\right)$. Summing (70) and (71) over k and using the fact that $\sum_{k=1}^n \psi_k = 1$ in $\overline{U}$, we have

$$\begin{aligned}\int_U \operatorname{div} \boldsymbol{f}\, d\boldsymbol{x} &= \int_U \operatorname{div}\left(\sum_{k=1}^n \psi_k \boldsymbol{f}\right) d\boldsymbol{x} = \sum_{k=1}^n \int_U \operatorname{div}\left(\psi_k \boldsymbol{f}\right) d\boldsymbol{x} \\ &= \sum_{k=1}^n \int_{\partial U} \psi_k \boldsymbol{f} \cdot \boldsymbol{\nu}\, d\mathcal{H}^{N-1} = \int_{\partial U} \sum_{k=1}^n \psi_k \boldsymbol{f} \cdot \boldsymbol{\nu}\, d\mathcal{H}^{N-1} = \int_{\partial U} \boldsymbol{f} \cdot \boldsymbol{\nu}\, d\mathcal{H}^{N-1},\end{aligned}$$

which is what we wanted. ■

Monday, December 5, 2022

13 Conservative and Irrotational Vector Fields

Definition 155 *Given two intervals $I, J \subseteq \mathbb{R}$, and two functions $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ and $\boldsymbol{\psi} : J \to \mathbb{R}^N$ of class C^k, $k \in \mathbb{N}_0$, we say that they are* equivalent *if there exists a bijective function $h : I \to J$ with h and h^{-1} of class C^k such that*

$$\boldsymbol{\varphi}(t) = \boldsymbol{\psi}(h(t))$$

for all $t \in I$. *We write* $\boldsymbol{\varphi} \sim \boldsymbol{\psi}$ *and we call* $\boldsymbol{\varphi}$ *and* $\boldsymbol{\psi}$ parametric representations of class C^k *and the function* h *a* parameter change of class C^k. *A* curve $\boldsymbol{\gamma}$ *of class* C^k *is an equivalence class of parametric representations of class* C^k, *that is,* $[\boldsymbol{\varphi}] := \{\boldsymbol{\psi} : \boldsymbol{\psi} \sim \boldsymbol{f}\}$. *The set* $\Sigma = \boldsymbol{\varphi}(I)$ *is called the* range of the curve.

Definition 156 *A curve* $\boldsymbol{\gamma}$ *of class* C^k *is* closed *if it has a parametric representation* $\boldsymbol{\varphi} : [a, b] \to \mathbb{R}^N$ *with* $\boldsymbol{\varphi}(a) = \boldsymbol{\varphi}(b)$.

Similarly we can define C^∞ curves, Lipschitz curves, analytic curves, and so on.

Remark 157 *Note that given a curve* $\boldsymbol{\gamma}$ *of class* C^k *with parametric representation* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$, *the function* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ *is not in general a local chart for a one-dimensional manifold, since we are not assuming that* $\boldsymbol{\varphi}$ *is injective or that* $\boldsymbol{\varphi}'(t) \neq \mathbf{0}$ *for every* $t \in I$. *In particular, a curve could self intersects but a one-dimensional manifold cannot.*

Next we introduce the notion of an oriented curve.

Definition 158 *Given a curve* $\boldsymbol{\gamma}$ *in* $\mathbb{R}^N$ *of class* C^k, $k \in \mathbb{N}_0$, *with parametric representations* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ *and* $\boldsymbol{\psi} : J \to \mathbb{R}^N$, *we say that* $\boldsymbol{\varphi}$ *and* $\boldsymbol{\psi}$ *have the* same orientation *if the parameter change* $h : I \to J$ *is increasing and* opposite orientation *if the parameter change* $h : I \to J$ *is decreasing. If* $\boldsymbol{\varphi}$ *and* $\boldsymbol{\psi}$ *have the same orientation, we write* $\boldsymbol{\varphi} \overset{*}{\sim} \boldsymbol{\psi}$.

Exercise 159 *Prove that* $\overset{*}{\sim}$ *is an equivalence relation.*

Definition 160 *An* oriented curve $\boldsymbol{\gamma}$ *in* $\mathbb{R}^N$ *of class* C^k, $k \in \mathbb{N}_0$, *is an equivalence class of parametric representations with the same orientation.*

Note that any curve $\boldsymbol{\gamma}$ in $\mathbb{R}^N$ gives rise to two oriented curves. Indeed, it is enough to fix a parametric representation $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ and considering the equivalence class $\boldsymbol{\gamma}^+$ of parametric representations with the same orientation of $\boldsymbol{\varphi}$ and the equivalence class $\boldsymbol{\gamma}^-$ of parametric representations with the opposite orientation of φ.

Definition 161 *Given a Lipschitz continuous oriented curve* $\boldsymbol{\gamma}$ *in* $\mathbb{R}^N$ *and a function* $\boldsymbol{g} : E \to \mathbb{R}^N$, *where* E *contains the range of* $\boldsymbol{\gamma}$, *we define the* curve *(or* line*)* integral *of* $\boldsymbol{g}$ *along the curve* $\boldsymbol{\gamma}$ *as the number*

$$\int_{\boldsymbol{\gamma}} \boldsymbol{g} := \int_I \boldsymbol{g}\left(\boldsymbol{\varphi}(t)\right) \cdot \boldsymbol{\varphi}'(t)\, dt.$$

provided the function $t \in I \mapsto \boldsymbol{g}\left(\boldsymbol{\varphi}(t)\right) \cdot \boldsymbol{\varphi}'(t)$ *is Lebesgue integrable for every parametric representation* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ *of* $\boldsymbol{\gamma}$ *and the value of the integral does not change with the representation.*

Exercise 162 *Let $\boldsymbol{\gamma}$ be an oriented Lipschitz continuous curve in $\mathbb{R}^N$ with parametric representations $\boldsymbol{\varphi} : [a,b] \to \mathbb{R}^N$ and $\boldsymbol{\psi} : [c,d] \to \mathbb{R}^N$. Given a continuous function $\boldsymbol{g} : E \to \mathbb{R}^N$, where E contains the range of $\boldsymbol{\gamma}$, prove that*

$$\int_a^b \boldsymbol{g}(\boldsymbol{\varphi}(t)) \cdot \boldsymbol{\varphi}'(t)\, dt = \int_c^d \boldsymbol{g}(\boldsymbol{\psi}(\tau)) \cdot \boldsymbol{\psi}'(\tau)\, d\tau.$$

Also, a result analogous to Proposition **??** continues to hold for this type of line integral.

Proposition 163 *Let $\boldsymbol{\gamma}$ be an oriented Lipschitz continuous curve and $\boldsymbol{f}, \boldsymbol{g} : E \to \mathbb{R}^N$, where E contains the range of $\boldsymbol{\gamma}$. Then*

(i) if $\int_{\boldsymbol{\gamma}} \boldsymbol{f}$ and $\int_{\boldsymbol{\gamma}} \boldsymbol{g}$ are well defined, then for all $a, b \in \mathbb{R}$,

$$\int_{\boldsymbol{\gamma}} (a\boldsymbol{f} + b\boldsymbol{g}) = a \int_{\boldsymbol{\gamma}} \boldsymbol{f} + b \int_{\boldsymbol{\gamma}} \boldsymbol{g},$$

(ii) If $\int_{\boldsymbol{\gamma}} \boldsymbol{f}$ is well defined and $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ is a parametric representation of $\boldsymbol{\gamma}$, then $\left| \int_{\boldsymbol{\gamma}} \boldsymbol{f} \right| \le \operatorname{Var}_I \boldsymbol{\varphi} \sup_{\Sigma} \|\boldsymbol{f}\|$, where Σ is the range of $\boldsymbol{\gamma}$,

(iii) If $\int_{\boldsymbol{\gamma}} \boldsymbol{f}$ is well defined, $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ is a parametric representation of $\boldsymbol{\gamma}$, $c \in I^\circ$, and $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ are the oriented curves of parametric representations $\boldsymbol{\varphi}_1 : I \cap (-\infty, c] \to \mathbb{R}^N$ and $\boldsymbol{\varphi}_2 : I \cap [c, \infty) \to \mathbb{R}^N$, then

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} = \int_{\boldsymbol{\gamma}_1} \boldsymbol{f} + \int_{\boldsymbol{\gamma}_2} \boldsymbol{f}.$$

Definition 164 *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\boldsymbol{g} : U \to \mathbb{R}^N$. We say that $\boldsymbol{g}$ is* conservative vector field *if there exists a differentiable function $f : U \to \mathbb{R}$ such that*

$$\nabla f(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x})$$

for all $\boldsymbol{x} \in U$. The function f is called a scalar potential *for $\boldsymbol{g}$.*

Wednesday, December 7, 2022

Theorem 165 (Fundamental Theorem of Calculus for Curves) *Let $U \subseteq \mathbb{R}^N$ be an open set, let $f \in C^1(U)$, let $\boldsymbol{x}, \boldsymbol{y} \in U$ and let $\boldsymbol{\gamma}$ a Lipschitz oriented curve with parametric representation $\boldsymbol{\varphi} : [a,b] \to \mathbb{R}^N$ such that $\boldsymbol{\varphi}(b) = \boldsymbol{x}$, $\boldsymbol{\varphi}(a) = \boldsymbol{y}$, and $\boldsymbol{\varphi}([a,b]) \subset U$. Then*

$$\int_{\boldsymbol{\gamma}} \nabla f = f(\boldsymbol{x}) - f(\boldsymbol{y}).$$

Proof. Define $p(t) := f(\boldsymbol{\varphi}(t))$ and observe that p is Lipschitz with

$$p'(t) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\boldsymbol{\varphi}(t))\,\varphi_i'(t)$$

for $\mathcal{L}^1$ a.e. $t \in [a,b]$. Hence,

$$\int_{\boldsymbol{\gamma}} \nabla f = \int_a^b \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\boldsymbol{\varphi}(t))\,\varphi_i'(t)\,dt = \int_a^b p'(t)\,dt = p(b) - p(a) = f(\boldsymbol{x}) - f(\boldsymbol{y}),$$

where we have used the fundamental theorem of calculus for Lebesgue integral. ■

The previous theorem shows that if a conservative vector field is continuous, then its integral along a curve joining two points depends only on the value at the two points and not on the particular curve. If U is patwise connected, then this condition turns out to be equivalent to the vector field being conservative.

Definition 166 *A set $E \subseteq \mathbb{R}^N$ is* pathwise connected *if for every $\boldsymbol{x}, \boldsymbol{y} \in E$ there exists a continuous curve with range in E joining $\boldsymbol{x}$ with $\boldsymbol{y}$, that is, $\gamma = [\boldsymbol{\varphi}]$, and $\boldsymbol{\varphi} : [a,b] \to \mathbb{R}^N$ is such that $\boldsymbol{\varphi}(b) = \boldsymbol{x}$, $\boldsymbol{\varphi}(a) = \boldsymbol{y}$.*

Exercise 167 *Prove that if $U \subseteq \mathbb{R}^N$ is open and pathwise connected, then for every $\boldsymbol{x}, \boldsymbol{y} \in E$ there exists a polygonal path with range in U joining $\boldsymbol{x}$ with $\boldsymbol{y}$.*

Theorem 168 *Let $U \subseteq \mathbb{R}^N$ be an open pathwise connected set and let $\boldsymbol{g} : U \to \mathbb{R}^N$ be a continuous function. Then the following conditions are equivalent.*

(i) $\boldsymbol{g}$ is a conservative vector field,

(ii) for every $\boldsymbol{x}$, $\boldsymbol{y} \in U$ and for every two Lipschitz oriented curves $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ with parametric representations $\boldsymbol{\varphi}_1 : [a,b] \to \mathbb{R}^N$ and $\boldsymbol{\varphi}_2 : [c,d] \to \mathbb{R}^N$, respectively, such that $\boldsymbol{\varphi}_1(b) = \boldsymbol{\varphi}_2(d) = \boldsymbol{x}$, $\boldsymbol{\varphi}_1(a) = \boldsymbol{\varphi}_2(c) = \boldsymbol{y}$, and $\boldsymbol{\varphi}_1([a,b]), \boldsymbol{\varphi}_2([c,d]) \subset U$,

$$\int_{\boldsymbol{\gamma}_1} \boldsymbol{g} = \int_{\boldsymbol{\gamma}_2} \boldsymbol{g}.$$

(iii) for every Lipschitz closed oriented curve $\boldsymbol{\gamma}$ with range contained in U,

$$\int_{\boldsymbol{\gamma}} \boldsymbol{g} = \boldsymbol{0}.$$

Proof. We prove that (i) implies (ii). Assume that $\boldsymbol{g}$ is a conservative vector field with scalar potential $f : U \to \mathbb{R}$, let $\boldsymbol{x}$, $\boldsymbol{y} \in U$ and let $\boldsymbol{\varphi}_1 : [a,b] \to \mathbb{R}^N$ and $\boldsymbol{\varphi}_2 : [c,d] \to \mathbb{R}^N$ be as in (ii). Then by the previous theorem

$$\int_{\boldsymbol{\gamma}_1} \boldsymbol{g} = \int_{\boldsymbol{\gamma}_1} \nabla f = f(\boldsymbol{x}) - f(\boldsymbol{y}) = \int_{\boldsymbol{\gamma}_2} \nabla f = \int_{\boldsymbol{\gamma}_2} \boldsymbol{g}.$$

Conversely assume that (ii) holds. We need to find a scalar potential for $\boldsymbol{g}$. Fix a point $\boldsymbol{x}_0 \in U$ and for every $\boldsymbol{x} \in U$ define

$$f(\boldsymbol{x}) := \int_{\boldsymbol{\gamma}} \boldsymbol{g},$$

where $\boldsymbol{\gamma}$ a Lipschitz continuous oriented curve with parametric representation $\boldsymbol{\varphi} : [a,b] \to \mathbb{R}^N$ such that $\boldsymbol{\varphi}(b) = \boldsymbol{x}$, $\boldsymbol{\varphi}(a) = \boldsymbol{x}_0$, and $\boldsymbol{\varphi}([a,b]) \subset U$. We claim that there exist

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}) = g_i(\boldsymbol{x}).$$

Since U is open and $\boldsymbol{x} \in U$, there exists $B(\boldsymbol{x}, r) \subseteq U$. Fix $|h| < r$, then the segment joining the point $\boldsymbol{x} + h\boldsymbol{e}_i$ with $\boldsymbol{x}$ is contained in $B(\boldsymbol{x}, r)$. Define the curve $\boldsymbol{\psi} : [a, b+1] \to \mathbb{R}^N$ as follows

$$\boldsymbol{\psi}(t) := \begin{cases} \boldsymbol{\varphi}(t) & \text{if } t \in [a,b], \\ \boldsymbol{x} + (t-b)\,h\boldsymbol{e}_i & \text{if } t \in [b, b+1]. \end{cases}$$

Using (ii), we have that

$$\begin{aligned} f(\boldsymbol{x} + h\boldsymbol{e}_i) = \int_{\boldsymbol{\psi}} \boldsymbol{g} &= f(\boldsymbol{x}) + \int_b^{b+1} \sum_{j=1}^N g_j(\boldsymbol{x} + (t-b)h\boldsymbol{e}_i)\, h\delta_{ij}\, dt \\ &= f(\boldsymbol{x}) + \int_b^{b+1} g_i(\boldsymbol{x} + (t-b)\,h\boldsymbol{e}_i)\, h\, dt = \\ &= f(\boldsymbol{x}) + \int_0^h g_i(\boldsymbol{x} + s\boldsymbol{e}_i)\, ds, \end{aligned}$$

where in the last equality we have used the change of variable $s = (t-b)\,h$. It follows by the mean value theorem that

$$\frac{f(\boldsymbol{x} + h\boldsymbol{e}_i) - f(\boldsymbol{x})}{h} = \frac{1}{h}\int_0^h g_i(\boldsymbol{x} + s\boldsymbol{e}_i)\, ds = g_i(\boldsymbol{x} + s_h\boldsymbol{e}_i),$$

where s_h is between 0 and h. As $h \to 0$, we have that $s_h \to 0$ and so $\boldsymbol{x} + s_h\boldsymbol{e}_i \to \boldsymbol{x}$. Using the continuity of g_i, we have that there exists

$$\lim_{h\to 0} \frac{f(\boldsymbol{x} + h\boldsymbol{e}_i) - f(\boldsymbol{x})}{h} = \lim_{h\to 0} g_i(\boldsymbol{x} + s_h\boldsymbol{e}_i) = g_i(\boldsymbol{x}),$$

which proves the claim.

The equivalence between (ii) and (iii) is left as an exercise. ■

Remark 169 *The previous theorem is used to prove that a vector field is not conservative. Indeed, if $U \subseteq \mathbb{R}^N$ is an open pathwise connected set and $\boldsymbol{g} : U \to \mathbb{R}^N$ is a continuous function, if you can construct a Lipschitz closed oriented curve $\boldsymbol{\gamma}$ with range contained in U such that*

$$\int_{\boldsymbol{\gamma}} \boldsymbol{g} \neq \boldsymbol{0},$$

then $\boldsymbol{g}$ cannot be conservative.

Friday, December 9, 2022

Next we give a simple necessary condition for a field $\boldsymbol{g}$ to be conservative.

Definition 170 *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\boldsymbol{g} : U \to \mathbb{R}^N$ be differentiable. We say that $\boldsymbol{g}$ is an* irrotational vector field *or* a curl-free vector field *if*

$$\frac{\partial g_i}{\partial x_j}(\boldsymbol{x}) = \frac{\partial g_j}{\partial x_i}(\boldsymbol{x})$$

for all $i, j = 1, \dots, N$ and all $\boldsymbol{x} \in U$.

Theorem 171 *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\boldsymbol{g} : U \to \mathbb{R}^N$ be a conservative vector field of class C^1. Then $\boldsymbol{g}$ is irrotational.*

Proof. Since $\boldsymbol{g}$ is a conservative vector field, there exists a a scalar potential $f : U \to \mathbb{R}$ with $\nabla f = \boldsymbol{g}$ in U. But since $\boldsymbol{g}$ is of class C^1, we have that f is of class C^2. Hence, we are in a position to apply the Schwartz theorem to conclude that

$$\frac{\partial g_i}{\partial x_j}(\boldsymbol{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}) = \frac{\partial g_j}{\partial x_i}(\boldsymbol{x})$$

for all $i, j = 1, \dots, N$ and all $\boldsymbol{x} \in U$. ■

The next example shows that there exist irrotational vector fields that are not conservative.

Example 172 *Let $U := \mathbb{R}^2 \setminus \{(0,0)\}$ and consider the function*

$$\boldsymbol{g}(x, y) := \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right).$$

Then $\boldsymbol{g}$ is irrotational but not conservative. Indeed,

$$\begin{aligned} \frac{\partial}{\partial y}\left(-\frac{y}{x^2 + y^2} \right) &= -\frac{x^2 - y^2}{(x^2 + y^2)^2} \\ \frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2} \right) &= -\frac{x^2 - y^2}{(x^2 + y^2)^2} \end{aligned}$$

but, taking the oriented curve $\boldsymbol{\gamma}$ parametrized by $\boldsymbol{\varphi}(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$, we get

$$\begin{aligned} \int_{\boldsymbol{\gamma}} \boldsymbol{g} &= \int_0^{2\pi} \boldsymbol{g}(\cos t, \sin t) \cdot (-\sin t, \cos t)\, dt \\ &= \int_0^{2\pi} \left(-\frac{\sin t}{\cos^2 t + \sin^2 t}, \frac{\cos t}{\cos^2 t + \sin^2 t} \right) \cdot (-\sin t, \cos t)\, dt = 2\pi \neq 0. \end{aligned}$$

Hence, by Theorem 168(iii), $\boldsymbol{g}$ cannot be conservative.

The problem here is the fact that the domain has a hole.

Definition 173 *A set $E \subseteq \mathbb{R}^N$ is* starshaped *with respect to a point $\boldsymbol{x}_0 \in \mathbb{R}^N$ if for every $\boldsymbol{x} \in E$, the segment joining $\boldsymbol{x}$ and $\boldsymbol{x}_0$ is contained in E.*

Theorem 174 (Poincaré's Lemma) *Let $U \subseteq \mathbb{R}^N$ be an open set starshaped with respect to a point $\boldsymbol{x}_0$ and let $\boldsymbol{g} : U \to \mathbb{R}^N$ be an irrotational vector field of class C^1. Then $\boldsymbol{g}$ is a conservative vector field.*

Proof. For every $\boldsymbol{x} \in U$ define

$$f(\boldsymbol{x}) := \int_{\boldsymbol{\gamma}} \boldsymbol{g},$$

where $\boldsymbol{\gamma}$ is the curve given by the parametric representation $\boldsymbol{\varphi} : [0,1] \to \mathbb{R}^N$ is defined by

$$\boldsymbol{\varphi}(t) := \boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0).$$

Note that

$$f(\boldsymbol{x}) = \int_0^1 \sum_{j=1}^N g_j(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))(x_j - x_{0j})\, dt.$$

Since $\boldsymbol{g}$ is of class C^1 we can differentiate under the integral sign to get

$$\begin{aligned}
\frac{\partial f}{\partial x_i}(\boldsymbol{x}) &= \int_0^1 \frac{\partial}{\partial x_i}\left(\sum_{j=1}^N g_j(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))(x_j - x_{0j})\right) dt \\
&= \int_0^1 \left(\sum_{j=1}^N \frac{\partial g_j}{\partial x_i}(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\, t(x_j - x_{0j}) + g_i(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\, 1\right) dt \\
&= \int_0^1 \left(\sum_{j=1}^N \frac{\partial g_i}{\partial x_j}(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\, t(x_j - x_{0j}) + g_i(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\, 1\right) dt,
\end{aligned}$$

where we have used the fact that $\boldsymbol{g}$ is an irrotational vector field. Define

$$h(t) := t g_i(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)).$$

By the chain rule,

$$h'(t) = \sum_{j=1}^N \frac{\partial g_i}{\partial x_j}(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\, t(x_j - x_{0j}) + g_i(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)).$$

Hence, by the fundamental theorem of calculus,

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}) = \int_0^1 h'(t)\, dt = h(1) - h(0) = 1 g_i(\boldsymbol{x}) - 0,$$

which completes the proof. ■

Definition 175 *Given a set* $E \subseteq \mathbb{R}^N$, $\boldsymbol{x}$, $\boldsymbol{y} \in E$, *and two continuous oriented curves* $\boldsymbol{\gamma}_1$ *and* $\boldsymbol{\gamma}_2$ *with range in* E *and parametric representations* $\boldsymbol{\varphi}_1 : [a, b] \to \mathbb{R}^N$ *and* $\boldsymbol{\varphi}_2 : [a, b] \to \mathbb{R}^N$, *respectively, such that* $\boldsymbol{\varphi}_1(a) = \boldsymbol{\varphi}_2(a) = \boldsymbol{x}$ *and* $\boldsymbol{\varphi}_1(b) = \boldsymbol{\varphi}_2(b) = \boldsymbol{y}$, *we say that* $\boldsymbol{\gamma}_1$ *and* $\boldsymbol{\gamma}_2$ *are* path homotopic *in* E *if there exists a continuous function* $\boldsymbol{h} : [0, 1] \times [a, b] \to \mathbb{R}^N$ *such that* $\boldsymbol{h}([0, 1] \times [a, b]) \subseteq E$,

$$\begin{aligned}&\boldsymbol{h}(0, t) = \boldsymbol{\varphi}_1(t) \text{ for all } t \in [a, b], \quad \boldsymbol{h}(1, t) = \boldsymbol{\varphi}_2(t) \text{ for all } t \in [a, b],\\ &\boldsymbol{h}(s, a) = \boldsymbol{x},\ \boldsymbol{h}(s, b) = \boldsymbol{y} \text{ for all } s \in [0, 1].\end{aligned}$$

The function $\boldsymbol{h}$ *is called a* path-homotopy in E *or* fixed endpoint homotopy *between the two curves.*

Roughly speaking, two curves are path homotopic in E if it is possible to deform the first continuously until it becomes the second *without leaving* the set E.

Definition 176 *Given a set* $E \subseteq \mathbb{R}^N$ *and* $\boldsymbol{x} \in E$, *a continuous oriented closed curve* $\boldsymbol{\gamma}_1$ *with range in* E *and parametric representation* $\boldsymbol{\varphi}_1 : [a, b] \to \mathbb{R}^N$ *such that* $\boldsymbol{\varphi}_1(a) = \boldsymbol{\varphi}_1(b) = \boldsymbol{x}$, *we say that* $\boldsymbol{\gamma}_1$ *is* null homotopic *in* E *if it is path homotopic in* E *to the continuous oriented curves* $\boldsymbol{\gamma}_2$ *parametrized by the constant function* $\boldsymbol{\varphi}_2(t) := \boldsymbol{x}$.

Definition 177 *A set* $E \subseteq \mathbb{R}^N$ *is* simply connected *if it is pathwise connected and if every continuous closed curve with range in* E *is null homotopic in* E.

Example 178 *A star-shaped set is simply connected. Indeed, let* $E \subseteq \mathbb{R}^N$ *be star-shaped with respect to some point* $\boldsymbol{x}_0 \in E$ *and consider a continuous closed curve* $\boldsymbol{\gamma}$ *with parametric representation* $\boldsymbol{\varphi} : [a, b] \to \mathbb{R}^N$ *such that* $\boldsymbol{\varphi}([a, b]) \subseteq E$. *Then the function*

$$\boldsymbol{h}(s, t) := s\boldsymbol{\varphi}(t) + (1 - s)\boldsymbol{x}_0$$

is an homotopy between $\boldsymbol{\gamma}$ *and the point* $\boldsymbol{x}_0$.

Theorem 179 *Let* $U \subseteq \mathbb{R}^N$ *be an open set, let* $\boldsymbol{\gamma}_1$ *and* $\boldsymbol{\gamma}_2$ *be two oriented closed Lipschitz continuous curves which are path homotopic in* U *and let* $\boldsymbol{g} : U \to \mathbb{R}^N$ *be of class* C^1 *and irrotational. Then*

$$\int_{\boldsymbol{\gamma}_1} \boldsymbol{g} = \int_{\boldsymbol{\gamma}_2} \boldsymbol{g}.$$

In particular, if U *is simply connected, then*

$$\int_{\boldsymbol{\gamma}} \boldsymbol{g} = 0$$

for every Lipschitz continuous closed oriented curve $\boldsymbol{\gamma}$ *with range in* U.