Georgii Potoshin, 2025

A Few Notions from Geometric Measure Theory

Preface

I have spent half a year reading about Geometric Measure Theory. Here, I have laid down my most interesting observations and few foundational theorems, ordered to make each new step as natural as possible. Geometric Measure Theory is indeed a geometric theory, because it is driven by geometric ideas. A primary interest of this domain is to study how measures can be used to investigate geometric properties of objects. With measures, we can describe shape convergence and many local geometric properties of objects without any restrictions on smoothness.

1 Conventions, acknowledged results and definition from analysis and measure theory

Convention: To any product, we associate morphisms denoted by π_i, where i is an index of the product component we are projecting to, or is the component itself. For example, for $A \times B$ we have $\pi_A : A \times B \to A$, for A^n we have $\pi_i : A^n \to A$ with $i \in [\![1, n]\!]$, for $\prod_{i \in I} A_i$ we have $\pi_{i_0} : \prod_{i \in I} A_i \to A_{i_0}$ for $i_0 \in I$. Similarly, we introduce a notion i for morphisms associated to coproducts. A projection on the tangent space is always associated to $T_x M \times (T_x M)^\perp$.

Convention: Let $\{w_i\}_i$ be a family of some kind of objects. We write a row of those objects as (w_i) and column as (w^i). We will use matrix multiplication notion to shorten equations, even when dealing with things that might not be scalars. For example, a basis change formula from basis (w_i) to a basis (u_i) with a matrix P is denoted by $(u_i) = (w_i)P$. If we have two function f, g defined on X and we have $a, b \in X$ then we may write

$$\begin{pmatrix} f \\ g \end{pmatrix} (a \ b) = \begin{pmatrix} f(a) & f(b) \\ g(a) & g(b) \end{pmatrix} \qquad f \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} f(a) \\ f(b) \end{pmatrix}$$

Often this way of writing equations compactifies not only notions but also proofs, as you will see in Section 7, by using substitutions. Although, I switch later in that section to Einstein's convention, because it eases permutations of sums and regrouping of coordinates.

1.1 Measure theory

Geometric Measure Theory is based upon Radon measure theory, where the notion of outer measure is obligatory. We consider that any set can be measured.

Definition: *An **outer measure** on X is a set function on X with values in $[0, \infty]$ with*

- $\mu(\emptyset) = 0$

- $E \subseteq \bigcup_{h \in \mathbb{N}} E_h \quad \Rightarrow \quad \mu(E) \le \sum_{h \in \mathbb{N}} \mu(E_h)$

Carathéodory's theorem: *If μ is an outer measure on X and $\mathcal{M}(\mu)$ is the family of those $E \subseteq X$ such that*

$$\mu(F) = \mu(E \cap F) + \mu(F \setminus E), \qquad \forall F \subseteq X$$

then $\mathcal{M}(\mu)$ is a σ-algebra and μ is a measure on $\mathcal{M}(\mu)$.

A proof can be found on pages 8-9 of [Maggi, 2012].

Definition: *μ is a **Borel measure** on a topological space X if it is an outer measure on X such that $\mathcal{B}(X) \subseteq \mathcal{M}(\mu)$.*

Definition: *A measure μ is said to be **absolutely continuous with respect to** measure λ if for any set A, $\lambda(A) = 0$ implies $\mu(A) = 0$ and we write it $\mu << \lambda$.*

Definition: *We say that a Borel measure μ is **regular** if for every $F \subseteq X$ there exists a Borel set $E \in \mathcal{B}(X)$ such that*

$$F \subseteq E, \qquad\qquad \mu(F) = \mu(E)$$

Definition: *An outer measure μ on X is **locally finite** if $\mu(K) < \infty$ for every compact set $K \subseteq X$.*

Definition: *An outer measure μ is a **Radon measure** on a topological space if it is locally finite, Borel regular measure on X.*

Property of Radon measures on $\mathbb{R}^n$: *If μ is a Radon measure on $\mathbb{R}^n$, then*

$$\mu(E) = \inf\{\mu(A) \mid E \subseteq A, A \text{ is open}\} = \sup\{\mu(K) \mid K \subseteq E, K \text{ is compact}\}$$

Those are only top-level results. While this theory incorporates many more profound statements, I've avoided them here to maintain the holistic picture I aim to present and not to have just another copy of an already existing book. This necessarily means sacrificing some depth, as I have not included few necessary proofs and results.

1.2 Analysis

Geometric measure theory relies on several profound and not-trivial results established within metric spaces or $\mathbb{R}^n$. This foundational properties, which I adapted with minor modifications or gave direct generalisations, are primarily sourced from "Measure theory and fine properties of functions".

For a ball $B = B(x, r)$ of center x and radius r we shall note $^\epsilon B = B(x, (1 + \epsilon)r)$ for every $\epsilon > 0$. I chose the prefix notation to avoid confusion with set power and Minkowski product.

Vitali's covering lemma: *Let $\mathcal{F}$ be any collection of non-degenerated closed balls in a metric space X with*

$$\sup\{diam\, B \mid B \in \mathcal{F}\} < \infty$$

Then for every $\epsilon > 1$ there exists a countable family $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\bigcup_{B \in \mathcal{F}} B \subseteq \bigcup_{B \in \mathcal{G}} {}^{2\epsilon}B$$

Proof: Set $D = \sup\{\mathrm{diam}\, B \mid B \in \mathcal{F}\}$. Set

$$\mathcal{F}_j = \left\{ B \in \mathcal{F} \mid \frac{D}{\epsilon^j} < \mathrm{diam}\, B \leq \frac{D}{\epsilon^{j-1}} \right\}, \quad j = 1, 2, \ldots$$

We define $\mathcal{G}_j \subseteq \mathcal{F}_j$ as follows

- Let $\mathcal{G}_1$ be any maximal disjoint collection of balls in $\mathcal{F}_1$.

- Assuming $\mathcal{G}_1, \ldots, \mathcal{G}_{k-1}$ have been selected, we chose $\mathcal{G}_k$ to be any maximal disjoint subcollection of
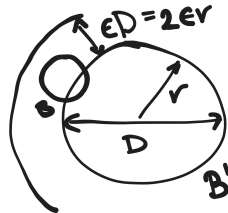
$$\{B \in \mathcal{F}_k \mid B \cap B' = \emptyset \text{ for all } B' \in \bigcup_{j=1}^{k-1} \mathcal{G}_j\}$$

They exist by Zorn's Lemma. Finally, define $\mathcal{G} = \bigcup_{j \in \mathbb{N}^*} \mathcal{G}_j$ a collection of disjoint balls and $\mathcal{G} \subseteq \mathcal{F}$.

Proving that for each ball $B \in \mathcal{F}$, there exists a ball $B' \in \mathcal{G}$ such that $B \cap B' \neq \emptyset$ and $B \subseteq {}^\epsilon B'$. Fix $B \in \mathcal{F}$, there exists and index j such that $B \in \mathcal{F}_j$ and by maximality of $\mathcal{G}_k$ there exists a ball $B' \in \bigcup_{k=1}^{j} \mathcal{G}_k$ with $B \cap B' \neq \emptyset$. But $\mathrm{diam}\, B' > \frac{D}{\epsilon^j}$ and $\mathrm{diam}\, B \leq \frac{D}{\epsilon^{j-1}}$; so that

$$\mathrm{diam}\, B \leq \frac{D}{\epsilon^{j-1}} < \epsilon\, \mathrm{diam}\, B'$$

Thus $B \subseteq {}^{2\epsilon}B'$.

Remark: This is a generalised version of the proof from the book "Measure theory and fine properties of functions" where it is done for the smallest integral case $\epsilon = 2$. The generalised proof shows the reason why the final dilatation is $5 = 1 + 2\epsilon$, but actually it is true for dilatation > 3 and the smallest such integer is 4. An interesting question is whether there is continuity and can we take a limit and get this result also for 3.

There are also more substantial covering results, crucial for extension theorems, measure differentiation and characterisation of rectifiable sets. Their proofs usually occupies a dozen of pages and are deeply analytic, primarily involving the extensive rewrite of inequalities and introduction of different constants. I initially intended to revise them for greater clarity, but after several failed attempts, when I had to modify constants and propagate the change back over multiple pages, I decided it was best to simply reference them here. Full proofs are available in [Lawrence C. Evans, 2015] and in [Alberti, 2017].

Besicovitch's Covering Theorem: *There exists a constant N_n, depending only on the dimension n with the following property:*

If $\mathcal{F}$ is any collection of non-degenerated closed balls in $\mathbb{R}^n$ with

$$\sup\{\operatorname{diam} B \mid B \in \mathcal{F}\} < \infty$$

and A is the set of centers of balls in $\mathcal{F}$, then there N_n countable collections $\mathcal{G}_1, \ldots, \mathcal{G}_{N_n}$ of disjoint balls in $\mathcal{F}$ such that

$$A \subseteq \bigcup_{i=1}^{N_n} \bigcup_{B \in \mathcal{G}_i} B$$

Definition: *A cover of A by a family B is called **fine** if for any $x \in A$ we can find a covering set from B of arbitrary small diameter.*

Filling open sets with balls theorem: *Let μ be a Borel measure on $\mathbb{R}^n$, and $\mathcal{F}$ any collection of non-degenerated closed balls. Let A denote the set of centers of the balls in $\mathcal{F}$. Assume*

$$\mu(A) < \infty$$

and $\mathcal{F}$ is a fine cover of A. Then for each open set $U \subseteq \mathbb{R}^n$, there exists a countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\bigcup_{B \in \mathcal{G}} B \subseteq U$$

and

$$\mu\left((A \cap U) \setminus \bigcup_{B \in \mathcal{G}} B\right) = 0.$$

Whitney's extension theorem: *Let $C \subseteq \mathbb{R}^n$ be a closed set and $f : C \to \mathbb{R}, d : C \to \mathbb{R}^{n*}$ be continuous functions. We shall use notions*

$$R(y, x) = \frac{f(y) - f(x) - d(x)(y - x)}{|x - y|}, \quad \forall x, y \in C, x \neq y$$

$$\rho_k(\delta) = \sup\{|R(x, y)| \mid 0 < |x - y| \leq \delta, x, y \in K\}$$

if we suppose that for every compact $K \subseteq C$

$$\rho_K(\delta) \to 0 \text{ as } \delta \to 0$$

Then there exists a function $\overline{f} \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ and $D\overline{f}|_C = d$.

Definition: *For a function $f : X \to Y$ between metric spaces we can define its **Lipschitz constant** $\operatorname{Lip}(f) = \inf\{L \in \mathbb{R} \mid d(f(x), f(y)) \leq Ld(x, y) \forall x, y \in X\}$*

Lipschitz function extension theorem: *Let X be a metric space, $A \subseteq X$ and $f : A \to \mathbb{R}$. Then there exists a Lipschitz function $\overline{f} : X \mapsto \mathbb{R}$ such that $\operatorname{Lip}(f) = \operatorname{Lip}(\overline{f})$ and $\overline{f}|_A = f$.*

Let's set $L = \operatorname{Lip}(f)$. Then we define

$$\overline{f}(x) = \inf_{y \in A}(f(y) + Ld(x, y))$$

By the definition, for all $x \in A$, $\overline{f}(x) \leq f(x)$ as in particular we can chose $y = x$. Furthermore, for all $a, b \in A$ and $x \in X$, we have an inequality for a Lipschitz function $f(b) - f(a) \leq Ld(b, a) \leq Ld(b, x) + Ld(a, x)$ and thus we have

$$f(a) + Ld(a, x) \geq f(b) - Ld(b, x)$$

and if we apply an infinum over a, we have $\overline{f}(x) \geq f(b) - L(b, x)$ and if $x \in A$ we can chose $b = x$ and we have an inequality in the other direction and thus the equality $\overline{f}(x) = f(x)$.

Now we check the Lipschitz constant

Consequence: *Let X be a metric space, $A \subseteq X$ and $f : A \to \mathbb{R}^n$. Then there exists a Lipschitz function $\overline{f} : X \mapsto \mathbb{R}^m$ such that $\overline{f}|_A = f$*

Let's set $\overline{f} = (\overline{f}_i)_i$ extension by coordinate functions.

Remark: I considered extending the theorem to vector-valued functions, but I could only prove it for the maximum norm.

1.3 Differentiation of Radon measure and Radon-Nikodym Theorem

Definition: *Let μ and ν be Radon measures on $\mathbb{R}^n$. Then we can define upper and lower derivatives of ν by μ by*

$$\overline{D}_\mu \nu(x) = \begin{cases} \limsup_{r \to 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0 \end{cases}$$

$$\underline{D}_\mu \nu(x) = \begin{cases} \liminf_{r \to 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0 \end{cases}$$

If $\overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x) < +\infty$ then we say that ν is differentiable *with respect to μ at x and we write*

$$D_\mu \nu(x) = \overline{D}_\mu \nu(x) = \underline{D}_\mu \nu(x)$$

Definition: *Let μ be Radon measure and ν be a vector measure on $\mathbb{R}^n$. Then we define a derivatives as*

$$D_\mu \nu(x) = \begin{cases} \lim_{r \to 0} \frac{\nu(B(x,r))}{\mu(B(x,r))} & \text{if } \mu(B(x,r)) > 0 \text{ for all } r > 0 \\ +\infty & \text{if } \mu(B(x,r)) = 0 \text{ for some } r > 0 \end{cases}$$

While vector-valued measures are defined later, it suffices to know that they are set functions with values in vector spaces that satisfy a particular version of sigma-additivity.

Differentiating measures theorem: *Let μ and ν be Radon measures on $\mathbb{R}^n$. Then*

- *$D_\mu \nu$ exists and is finite μ-a.e.*

- *$D_\mu \nu$ is μ-measurable*

Remark: This is a brilliant version of the proof of Radon-Nikodym Theorem by von Neumann, that will suffice for representing vector measures.

Radon-Nikodym Theorem: *Let μ and ν be finite measures on $(\Omega, \mathcal{F})$ then there exists a non-negative measurable function f and a μ-null set B such that*

$$\nu(A) = \int_A f \, d\mu + \nu(A \cap B)$$

for each $A \in \mathcal{F}$.

Proof: Let $\pi = \mu + \nu$ and consider $T(f) = \int f d\nu$. Then for every $f \in L^2(\pi)$ we have $f \in L^1(\nu)$ and since measures are finite $f \in L^1(\pi)$ and thus in $L^1(\nu)$, then

$$|T(f)| = |\int f \, d\nu| \leq \|f\|_{L^2(\nu)} \|1\|_{L^2(\nu)} \leq C\|f\|_{L^2(\pi)}$$

where C is a constant. Thus T is a continuous operator on $L^2(\pi)$ and by Reisz representation theorem for Hilbert spaces we find a function $h \in L^2(\pi)$ such that

$$T(f) = \int f \, d\nu = \int f h \, d\pi$$

Now consider the following sets

$$N = \{h < 0\}, \quad M = \{0 \le h < 1\}, \quad B = \{h \ge 1\}$$

Then

$$0 \ge \int_N h \, d\pi = \int \chi_N h \, d\pi = \nu(N) \ge 0$$

Thus we have $\nu(N) = 0$ and since $h < 0$ on N and $\int h \, d\mu = 0$ we have $\mu(N) = 0$ also.

Now let's study B. We have

$$\nu(B) = T(\chi_B) = \int_B h \, d\mu + \int_B h \, d\nu \ge \nu(B) + \mu(B)$$

thus $\mu(B) = 0$.

For the last let we set $M_n = \{0 \le h \le 1 - 1/n\}$ and from representation of T with h we have $\int (1 - h)f \, d\nu = \int hf \, d\mu$ and why apply it to

$$\nu(M_n) = \int \frac{\chi_{M_n}}{1-h}(1-h) \, d\nu = \int h \frac{\chi_{M_n}}{1-h} \, d\mu$$

Let $f = \frac{h}{1-h}$ then by applying monotone convergence and recalling that $\mu(B) = \mu(N) = 0$ we have

$$\nu(M \cap A) = \int_A f \, d\mu$$

Finally for all $A \in \mathcal{F}$ we have

$$\nu(A) = \nu(A \cap N) + \nu(A \cap M) + \nu(A \cap B) = \int_A f \, d\mu + \nu(A \cap B)$$

Corollary: If $\mu >> \nu$, then $\nu(A) = \int_A f \, d\mu$

It's evident since $\mu(B) = 0 \Rightarrow \nu(B) = 0$ then.

Remark: If μ and ν are Radon measures, then $f = D_\mu \nu$.

Corollary: This gives us a representation of a vector valued measure μ as $\mu = f \, |\mu|$. And we can define the integration with respect to such measure as

$$\int g \, d\mu = \int g \cdot f \, d|\mu|$$

Beyond representation of measure (or continuous linear forms), we can view this process from another angle. If we associate to a linear continuous form to a shape, then function f describes certain geometric properties of that shape. We will explore one such example in the last section.

2 Hausdorff measure

The Hausdorff measure generalizes the notion of measure for lower-dimensional objects in a higher-dimensional space or even in an arbitrary metric space. The idea is essentially similar to the construction of Lebesgue measure, and the general case is called the Carathéodory construction. The difference is that we measure sets by their diameter, attributing to each an s-dimensional "volume" based on its diameter raised to the power of s. For this construction, we define a cover of E by sets of diameter less than δ as a δ-cover of E. And we consider only countable covers. We note that

$$\mathcal{H}_\delta^s(E) = \inf_C \sum_{I \in C} \omega_s \left(\frac{\text{diam}(I)}{2} \right)^s$$

where $s \in \mathbb{R}_{\geq 0}$ is a dimension, $\omega_s \in \mathbb{R}$ is a coefficient, preferably continuous or smooth as a function of s, and C is a δ-cover of E. We may assume that

$$\omega_s = \frac{\pi^{s/2}}{\Gamma(1 + s/2)}$$

We define the Hausdorff measure as a limit of the previous value. It exists because $\mathcal{H}_\delta^s(E)$ is an increasing function of δ. We note

$$\mathcal{H}^s(E) = \lim_{\delta \to 0^+} \mathcal{H}_\delta^s(E)$$

I shall introduce the notion of s-variation of a cover S as

$$\mathrm{Var}^s(S) = \sum_{I \in S} \omega_s \left(\frac{\mathrm{diam}(I)}{2} \right)^s$$

Proposition: For a natural $n \geq 0$, ω_n is a volume of a unit n-dimensional ball.

2.1   Properties of Hausdorff measure

Proposition: *Hausdorff measure is a Borel measure.*

Proposition: $\mathcal{H}^0$ is the counting measure.

Proposition: *In the definition of Hausdorff measure we can consider only closed or open sets.*

Proposition: *Hausdorff measure of dimension $m \in N$ coincide on m-dimensional affine subspaces with their Lebesgue measure.*

Proposition: *The n-dimetional Hausdorff measure traced to a n-dimensional $\mathcal{C}^1$-submanifold of R^m induces the area measure on this submanifold and coincides with the integral measure via parametrisation on it.*

Proposition: *A restriction of $\mathcal{H}^s$ on a locally $\mathcal{H}^s$-finite set is a Radon measure.*

Remark: Proofs to those proposition can be found in the book [Maggi, 2012].

Definition: *Let E be a Borel subset of a metric space X. The **upper d-dimensional density** (with respect to $\mathcal{H}^d$) of E at the point x is defined by setting*

$$\Theta_d^*(E, x) = \limsup_{r \to 0} \frac{\mathcal{H}^d(E \cap B_c(x, r))}{\omega_d r^d}$$

Theorem: *Let E be a Borel subset of a metric space X, and assume that E is locally $\mathcal{H}^d$-finite. Then the following properties hold true:*

1. *The upper density of external points is almost everywhere zero, $\Theta_d^*(E, x) = 0$ for $\mathcal{H}^d$-almost every $x \in E^c$.*

2. *The upper density is bounded from below, $\frac{1}{2^d} \leq \Theta_d^*(E, x)$ for $\mathcal{H}^d$-almost every $x \in E^c$.*

3. *If $X \cong \mathbb{R}^n$, then the upper density is bounded from above by 1, $\Theta_d^*(E, x) \leq 1$ for $\mathcal{H}^d$-almost every $x \in E$.*

4. *If X is a generic metric space, then the upper density is bounded from above by a slightly different constant, $\Theta_d^*(E, x) \leq 3^d$ for $\mathcal{H}^d$-almost every $x \in E$.*

Proof: Proofs of propositions 2-3 can be found in [Alberti, 2017] on page 48.

1. The core of this proof relies on the regularity of $\mu := \mathcal{H}^d \llcorner E$ and Vitali's Covering Lemma. We define $E_t := \{x \in E^c \mid \Theta_d^*(E, x) > t\}$ and aim to prove that it is $\mathcal{H}^d$-null for $t > 0$. From the definition of E_t we know that E is highly concentrated around the points of E_t. This effectively allows us to reduce the study of E_t to the study of appropriate covers. Let $A \supseteq E_t$ be an open neighborhood of E_t. We then consider the family of closed balls $\mathcal{F} := \{\overline{B(x, r)} \subset A \mid x \in E_m \ \& \ \mathcal{H}^d(\overline{B(x, r)} \cap E) > t\omega_d r^d\}$. Within this family, we can consider balls of diameter < 1. Then for **Vitali's Covering Lemma** we can choose a dilation coefficient $k \in (3, +\infty)$ and we fix ball dilatation procedure with

coefficient k by $B \mapsto \hat{B}$. Lemma provides a disjoint subfamily $\mathcal{G}$ such that the delated by $\hat{}$ balls cover E_t. We then have the following inequality

$$\mu(A) \geq \sum_{B \in \mathcal{G}} \mu(B) \geq t\omega_d \sum_{B \in \mathcal{G}} r(B)^d = \frac{t}{k^d}\omega_d \sum_{B \in \mathcal{G}} r(\hat{B})^d \geq \frac{t}{k^d}\mathcal{H}_1^d(E_n)$$

Since $\mu(A)$ can be chosen arbitrary small, it follows that $\mathcal{H}_1^d(E_t) = 0$ and consequently $\mathcal{H}^d(E_t) = 0$.

2. [Alberti, 2017]

3. [Alberti, 2017]

4. If we follow a similar construction to the proof of the first proposition for a set $E_t := \{x \in E \mid \Theta_d^*(E, x) > t\}$, then we arrive at

$$\mathcal{H}^d(E_m) \geq \frac{t}{k^d}\mathcal{H}^d(E_m)$$

and if $t > k^d$ we get $\mathcal{H}^d(E_m) = 0$, and as it is true for all $k > 3$, we have $\Theta_d^*(E, x) \leq 3^d$. It's worth noticing that usually there are no generalisation in Vitali's lemma and one usually proves for 5 instead of 3.

2.2   Hausdorff dimension

To a set S we can associate a number $s = \inf\{a \geq 0 \mid \mathcal{H}^a(S) = 0\}$. It's called its Hausdorff dimension.

Proposition: *If $E \subseteq \mathbb{R}$ then $\dim(E) \in [0, n]$. Moreover $\mathcal{H}^s(E) = \infty$ for every $s < \dim(E)$ and $\mathcal{H}^S(E) \in (0, \infty)$ implies $s = \dim(E)$.*

**Proposition:** *If A is an open set in $\mathbb{R}^n$, then $\dim(A) = n$.*

Proposition: *For a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^m$ we have the following inequality*

$$\mathcal{H}^s(f[E]) \leq Lip(f)^s \mathcal{H}^s(E)$$

for every $s > 0$ and $E \subseteq \mathbb{R}^n$ and $dim(E) < dim(f[E])$.

2.3   Dimension of Cantor sets

Here we compute the dimension of generalized set. Let $n \in \mathbb{N}$ and $m \in \mathbb{N}^*$ so that $2m < n$. Then we can define C_k ($k \in \mathbb{N}$) define recursively by agreeing that $C_0 = \{[0, 1]\}$ and we obtain C_{k+1} from C_k by cutting out the open middle part from each segment of C_k and living side parts of length m/n of original interval. We will note $C = \lim C_k = \cap\, C_k$.
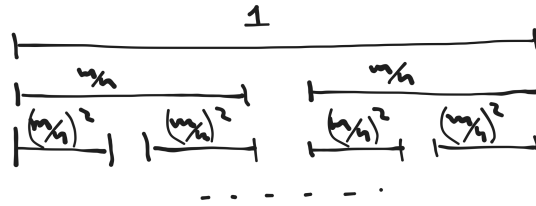
Obviously C_k is a $(m/n)^k$-cover of C, so

$$\mathcal{H}_{(m/n)^k}^s \leq \sum_{I \in C_k} \omega_s \left(\frac{\text{diam}(I)}{2}\right)^s = \omega_s 2^k ((m/n)^k)/2)^s = \omega_s / 2^s (2(m/n)^s)^k$$

And if $s > \log_{n/m}(2)$ we have right side approaching 0 as k tends to infinity. That means that $\dim(C) \leq \log_{n/m}(2)$.

Now we need to prove the inequality in the other direction. Let $s = \log_{n/m}(2)$. And let S be a $(m/n)^k$-cover of C. In fact by the construction C is an intersection of compacts on a real line, so is compact. And by one of the previous propositions we can conceder only open covers. Then by compactness we can leave only a finite number of sets in S and this way we reduce its Hausdorff variance and we can extend the resting elements to closed intervals of the same diameter. This does not change the variance. The new

cover is noted by S'. Now in every interval of S' we can find 2 maximal intervals from some C_i and C_j, so the they are disjoint. If we can't do that, then there are no points of C in this interval and we can throw away that set also. So now we have 2 maximal intervals J and J' in I. They are ordered. Between them we have an interval K and as they are maximal $I \setminus J \setminus K \setminus J'$ does not contain any points from C and we can through those parts away from the covering. By the construction

$$|J|, |J'| \leq \frac{m}{n} \cdot \frac{n}{n-2m}|K| = \frac{m}{n-2m}|K|$$

Now we have $1/2(|J| + |J'|) \leq \frac{m}{n-2m}|K|$

$$|I|^s = (|J|+|J'|+|K|)^s \geq ((1+\frac{n-2m}{2m}))(|J|+|J'|))^s = (\frac{n}{m}1/2(|J|+|J'|))^s = 2(1/2(|J|+|J'|))^s \geq |J|^s+|J'|^s$$

Where the last step is done by concavity of function $x \mapsto x^s$. That means that we can reduce this any cover to a C_k cover which has a smaller s-variation. That means that for dimension $s = \log_{n/m}(2)$ the $\mathcal{H}^s(C)$ is finite as the s-variation of C_k is always $\omega_s/2^s$.

Remark: This is a variation on the proof given in the book "The geometry of fractal sets" by K. J. Flaconer, generelised to the case of arbitrary m and n. In this book the proof is done for the case $m = 1, n = 3$.

Proposition: *There is a subset of* $[0, 1]$ *with a Hausdorff dimension* 1, *but Lebesgue measure 0.*

To show that we shall use Cantor's sets. Let $C_{m/n}$ be a set discussed in a previous paragraph. Then $S = \bigcap C_{m/(2m+1)}$ is a set of dimension 1. As for every $0 \leq s < 1$ there is such m, that $\log_{n/m}(2) = \log_{(2m+1)/m}(2) > s$, as $\log_{(2m+1)/m}(2) \to 1$. And thus $\mathcal{H}^s(S) > \mathcal{H}^s(C_{m/(2m+1)}) = \infty$.

2.4 Integration

Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^k$ be an injective $\mathcal{C}^1$ mapping. Then its jacobian can be written as

$$J(f) = \sqrt{\det(df^T \circ df)}$$

and we have the following integration formula:

$$\int_{\text{im}(f)} g \, d\mathcal{H}^n = \int_U g(f(x))J(f)dx$$

3 Convergence of measures

Measures not only allow us to compute integrals, but they can also be used to model geometric figures and to test different properties of these figures. One of the fundamental ideas at the heart of geometric measure theory is that one can replace figures with the measures induced on these figures.

$$E \rightsquigarrow \mu \llcorner E$$

Now, we need to compare two figures. To do this, we can compare the values of integrals of functions with respect to associated measures; in other words, we will treat measures as linear functionals. Furthermore, if two figures are close to one another, then we want their associated measures to yield sufficiently close values. This implies that we want the function values to remain bounded in small neighborhoods and not vary too much. Therefore, we only consider continuous functions. Lastly, since we want to be able to work with possibly unbounded figures, we would like the integration of measures with respect to functions to be well-defined and finite. Thus, we only use continuous functions with compact support, $\mathcal{C}_c(X)$.

This allows us to establish a notion of convergence of shapes, equivalent to a convergence of measures. We'll see examples of this later, but for now, I'd like to specify the type of convergence we'll be using. We are treating measures as linear functionals on the space $\mathcal{C}_c(X)$. In this context, convergence is defined by the behavior of these functionals on each test function; specifically, we are considering that the integral values converge for every function in $\mathcal{C}_c$. Such convergence is called weak-* convergence. Before discussing this convergence further, I shall first demonstrate that there is a space of measures which serves as the dual space to $\mathcal{C}_c$.

3.1 Topologies on spaces E and E^*

For topological spaces Y_i and a set of functions $f_i : X \to Y_i$, we can define the smallest, coarsest topology on X that makes these functions continuous. By definition such topology is $\tau(\{f_i\}) = \bigcap\{\tau \mid \tau$ is a topology on

X and f_i are continuous}. As an example, the product topology is exactly $\tau(\{\pi_i\})$, where π_i are canonical projections.

Proposition: *Let τ be a topology on X. Then $\tau = \tau(\{f_i\})$ if and only if every function $g : W \to X$ such that $f_i \circ g$ are continuous is continuous.*

Remark: This is a well-known property of caorsest topology, but I checked that it is also an alternative characterisation of such topology.
If $\tau = \tau(\{f_i\})$ and $g : W \to X$ is such function that $f_i \circ g$ are continuous. It's sufficient to check that for all elements of prebase of $\tau(\{f_i\})$ the inverse image is open, but the prebase consists of elements of the form $f_i^{-1}[U]$ and its inverse image is $(f_i \circ g)^{-1}[U]$ which is open by hypotheses.

If τ is a such topology, that for every function $g : W \to X$ it is continuous if and only if $f_i \circ g$ are continuous, then in particular we have id $: (X, \tau) \to (X, \tau)$ continuous and that means that $f_i = f_i \circ$ id are continuous and we have $\tau(\{f_i\}) \subseteq \tau$. On the other hand we have id$' : (X, \tau(\{f_i\})) \to (X, \tau)$ continuous because $f_i = f_i \circ$ id$' : (X, \tau(\{f_i\})) \to Y_i$ are continuous by the definition of coarsest topology. Thus we have id$'$ continuous and that means that $\tau \subseteq \tau(\{f_i\})$. And finally $\tau = \tau(\{f_i\})$.

Tichonoff's Theorem: *Product of compact spaces is compact.*

General structure: Let I be a set of indices and E_i for $i \in I$ be a topological space with a topology τ_i. The prebase of the product topology on $\prod_{i \in I} E_i$ is $\{\pi_i^{-1}[U] \mid i \in I, U \in \tau_i\}$. a set of products of open subspaces of one spaces on others. All the finite intersections form a base of product topology. Its elements are products of open sets where almost all factors are E_i.

Maximal covers: Let's note that a set of covers that does not contain finite sub-covers for a partially ordered set with the relation of inclusion. For every chain we have its union which does not contain a finite sub-cover, which otherwise would have been in some element of chain. Thus each chain has an upper bound. By the Zorn's lemma we find a maximal element M.

Let X be a topological space and $M \subseteq \tau$ a maximal cover that does not contain a finite sub-cover. **Then if** $V \in M^c$, **we have** $U_1, \dots, U_n \in M$ **such that** $V \cup U_1 \cup \dots \cup U_n = X$. Because otherwise we could have added V to M and M would not be maximum. **If** $U, V \in M^c$ **then** $U \cap V \in M^c$. In other words M^c is a multiplicative system, which is similar to the statement that $\mathfrak{p}^c$ is multiplicative for a prime ideal $\mathfrak{p}$. This is true due to the fact that we have $U_1, \dots, U_k \in M$ and $V_1, \dots, V_l \in M$ such that $U \cup U_1 \cup \dots \cup U_n = X = V \cup V_1 \cup \dots \cup V_l$ and thus $(U \cap V) \cup U_1 \cup \dots \cup U_k \cup V_1 \cup \dots \cup V_l = X$, which implies that $U \cap V \in M^c$.

Alexander's lemma about prebase: Let B be a prebase of a topological space X. Then if in every cover of X by elements of B there exists a finite subcover, then the space X is compact. If X is not compact, then we have a M maximal cover that does not contain a finite sub-cover. Then to every $x \in X$ we can associate its neighborhood $V_x \in M$. Then we find some element of a basis $U_x = U_{1,x} \cup \dots \cup U_{n_x,x} \subseteq V_x$ where $U_{i,x} \in B$ are elements of prebase. Thus by maximality $U_x \in M$ as $U_x \subseteq V_x$. But as $U_x = U_{1,x} \cup \dots \cup U_{n_x,x}$ and as M^c is a multiplicative system, for some i we have $U_{i,x} \in M$. It means that in M we have a sub-cover of X by elements of a prebase B. And by hypotheses we can chose a finite sub-cover which gives a contradiction.

Tichonoff theorem's proof: Let $\mathcal{S} = (U_i)_{i \in I}$ be a cover of a product $E = \prod_{j \in J} E_j$ of compact space by elements of canonical prebase. Let's suppose that it does not contain a finite sub-cover. For every $j \in J$ we shall pose $S_j = \{\pi_j^{-1}[V_{i,j}] = U_i \mid V_{i,j} \in \tau_j, i \in I_j\}$. Then $(V_{i,j})_{i \in I}$ cannot be a cover of E_j, because otherwise we can extract a finite sub-cover of E_j and hence of E. So we can chose $x_j \in E_j$ such that $x_j \notin \bigcup_{i \in I_j} V_{i,j}$. Let $x = (x_j)_{j \in J}$ and it does not lie in every set of $\mathcal{S}$, thus it is not a cover and we get a contradiction.

Remark: This is the most non-trivial part of the proof of Banach-Alaoglu theorem and as I had this proof noted I have decided to also put it here.

In this section, E is a normed vector space and E^* is its dual space of continuous 1-forms on E. On the space E, apart from its metric topology, we have the weak topology $\sigma(E, E^*) = \tau(\{f\}_{f \in E^*})$. As $f \in E^*$ is continuous with respect to the regular topology, the topology $\sigma(E, E^*)$ is coarser then the regular topology, which we call strong.

On the space E^*, we also have strong topology with the operator norm. Additionally, we have the weak-$*$ topology $\sigma(E^*, E) = \tau(\{v\}_{v \in E})$.

Proposition: *The weak-$*$ topology is a trace topology from the space $\mathbb{R}^E$ with the product topology.*

Proof: Let $\tau(\{\pi_v\}_{v \in E})$ be the trace topology. Then it is easy to see that $\pi_v = v$ as both function are evaluations at v and thus $\tau(\{\pi_v\}_{v \in E}) = \tau(\{v\}_{v \in E}) = \sigma(E^*, E)$ is a weak-$*$ topology.

Remark: In the book "Functional Analysis" by Haim Brezis, the part above is done by establishing an homeomorpism and a verification of its bicontinuity. As you have seen, there is actually nothing substantial to prove since these are just two notions of the same concept – projection and evaluation in the

dual-space.

Theorem (Banach-Alaoglu): *The closed unit ball $B = \{f \in E^* \mid |f| \leq 1\}$ is compact in the weak-$*$ topology $\sigma(E^*, E)$.*

Proof:

$$B = \left\{ f \in \mathbb{R}^E \mid \begin{cases} |f(x)| < |x|, \ \forall x \in E \\ f(\lambda x) = \lambda f(x), \ \forall \lambda \in \mathbb{R}, x \in E \\ f(x+y) = f(x) + f(y) \ \forall x, y \in E \end{cases} \right\}$$

Hence it is intersection of the following sets $B = K \cap \bigcap_{x,y \in E} A_{x,y} \cap \bigcap_{x \in E, \lambda \in \mathbb{R}} B_{\lambda,x}$, where $K = \{f \in \mathbb{R}^E \mid |f(x)| \leq |x|\} = \prod_{x \in E} [-|x|, |x|]$ is compact by Tichonoff theorem, where for $x, y \in E$, we define $A_{x,y} = \{f \in \mathbb{R}^E \mid f(x+y) - f(x) - f(y) = 0\}$, which is closed since evaluations and addition are continuous, and thus $f \mapsto f(x+y) - f(x) - f(y)$ is continuous and $A_{x,y}$. For similar reasons $B_{\lambda,x} = \{f \in \mathbb{R}^E \mid f(\lambda x) - \lambda f(x) = 0\}$ is closed. This proves that B is compact.

3.2 Vector valued measure

Let X be a topological space and V a Banach space, then $\mu : \mathcal{B}(X) \to V$ is a V-valued Borel measure if

$$\sum_n \mu(E_n) = \mu\left(\bigcup_n E_n\right)$$

for any disjoint countable family $\{E_n\}$ of Borel sets. From that definition we have $\mu(A) + \mu(\emptyset) = \mu(A \cup \emptyset) = \mu(A)$ and thus $\mu(\emptyset) = 0$. This is a quite a strong property as the convergence of the sum does not depend on the order, which in finite dimensions is equivalent to the absolute convergence of that series.

Let μ be a vector valued measure. Then the *total variation* $|\mu|$ of a Borel set A by measure μ is defined by:

$$|\mu|(A) = \sup\left\{ \sum_n |\mu(A_n)| \mid \{A_n\} \text{ countable partition of } A \right\}$$

Proposition: *Total variation is a positive bounded measure.*

It is easy to see that $|\mu|(\emptyset) = 0$ since all partitions of an empty set consist of empty sets which measure is zero. The image of $|\mu|$ by the definition consists of positive numbers. Lastly we verify σ-additivity. Let $\{S_n\}$ be a disjoint countable collection of Borel sets. Then

$$\sum_n |\mu|(S_n) = \sum_n \sup\left\{ \sum_m |\mu(S_{n,m})| \mid (S_{n,m})_m \text{ is a countable Borel partition of } S_n \right\}$$

Then we remark that for each choice of $\{S_{n,m}\}$, it is a countable Borel partition of $S = \bigcup_n S_n$, and thus $|\mu|(S) \geq \sum_n |\mu|(S_n)$. On the other hand if $\{A_k\}$ is a countable Borel partition of S then we have partitions of S_n defined as $\{S_{n,k} = A_k \cap S_n\}_k$ and we have the following inequality:

$$\sum_k |\mu(A_k)| = \sum_k |\sum_n \mu(S_{n,k})| \leq \sum_n \sum_k |\mu(S_{n,k})|$$

which implies $|\mu|(S) \leq \sum_n |\mu|(S_n)$ and we conclude that $|\mu|$ is a positive measure.

Let's verify that total variation is bounded. That is a tricker question and we shall follow the proof from "...". The measure can be partitioned into projection measures $\mu = (\mu_i)_{i=1}^n$. As all the norms are equivalent we can consider $|\cdot| = \|\cdot\|_1$. Then as we have the following inequality:

$$\sup\left\{ \sum_i |\mu(X_i)| \mid X_i \text{ is a borel partition of } X \right\} \leq \sum_j \sup\left\{ \sum_i |\mu_j(X_i)| \mid X_i \text{ is a borel partition of } X \right\}$$

It is sufficient to prove that for real valued measures its total variation is bound. If we suppose it is not, then we have a real valued measure μ, countable Borel partition of X $\{X_m\}_m$ and $n \in \mathbb{N}$ such that

$$\sum_{m=0}^n |\mu(X_m)| > 2(|\mu(X)| + 1)$$

Let $P = \{X_i \mid \mu(X_i) > 0\}$ and $N = \{X_i \mid \mu(X_i) < 0\}$. Then we have $|\mu(\bigcup P)| > |\mu(X)| + 1$ or $|\mu(\bigcup N)| > |\mu(X)| + 1$, thus we have a set E such that $|\mu(E)| > |\mu(X)| + 1$. Then we have $|\mu(E^c)| = |\mu(X) - \mu(E)| \geq |\mu(E)| - |\mu(X)| > 1$. Then by additivity of $|\mu|$ we have $|\mu|(E) = \infty$ or $|\mu|(E^c) = \infty$; supposing the latter we pose $E_1 = E$ (or $= F$) we always have $\mu(E_1) > 1$ and if we continue the same procedure for $X = E^c$

we construct by the choice axiom the following sequence of disjoint sets $(E_i)_i$ and $|\mu|(E_i) > 1$ and thus $\sum \mu(E_i)$ does not converge and we have a contradiction to the definition of vector valued measure. Thus μ is bound.

By setting

$$\mu_+ = \frac{|\mu| + \mu}{2} \qquad\qquad \mu_- = \frac{|\mu| - \mu}{2}$$

we have μ_+ and μ_- positive bounded measures and $\mu = \mu_+ - \mu_-$ which ports a name a *Jordan decomposition*.

The *mass* of μ is set to be $\|\mu\| = |\mu|(X)$.

Proposition: *The set of vector norms with the mass form a normed vector space.*

Proof: Let $\mu : \mathcal{B}(X) \to V$ for V an $\mathbb{R}$-vector space be a vector norm. Then evidently $\|k\mu\| = |k|\|\mu\|$. Let ν be another vector measure then

$$
\begin{aligned}
\|\mu + \nu\| &= \sup\left\{ \sum_{n=0}^{+\infty} |(\mu + \nu)(E_n)| \,\big|\, \{E_n\}_n \text{ countable paritition of } X \right\} \\
&\leq \sup\left\{ \sum_{n=0}^{+\infty} |\mu(E_n)| + |\nu(E_n)| \,\big|\, \{E_n\}_n \text{ countable paritition of } X \right\} \\
&\leq \sup\left\{ \sum_{n=0}^{+\infty} |\mu(E_n)| \,\big|\, \{E_n\}_n \text{ countable paritition of } X \right\} \\
&\quad + \sup\left\{ \sum_{n=0}^{+\infty} |\nu(E_n)| \,\big|\, \{E_n\}_n \text{ countable paritition of } X \right\} \\
&= \|\mu\| + \|\nu\|
\end{aligned}
$$

3.3 Riesz representation theorems for vector valued measure

For an $\mathbb{R}^n$-valued measure μ on X we define an associated functional

$$\Lambda_\mu : \mathcal{C}_0(X, \mathbb{R}^n) \to \mathbb{R}$$
$$f \mapsto \int f \, d\mu$$

Riesz representation theorem: *The map*

$$\Lambda : \mathcal{M}(X, \mathbb{R}^n) \to \mathcal{C}_0(X, \mathbb{R}^n)^*$$
$$\mu \quad \mapsto \quad \Lambda_\mu$$

is an isometry

Proof: The injectivity of Λ is quit obvious. For surjectivity we make an inverse construction, for a given functional L we take its total variation defined by

$$|L|(A) = \sup\{\langle L \mid \phi\rangle \mid \phi \in \mathcal{C}_c(A, \mathbb{R}^n), |\phi| < 1\}$$

for open set A. And for other sets we set

$$|L|(E) = \inf\{|L|(A) \mid E \subseteq A\}$$

Thus $|L|$ is locally finite, because it's continuous and thus bounded. Whats more the second property yields us the regularity of the total variation as we can take a countable intersection of $\{A_n\}$ of open sets such that $|L|(A_n) \to |L|(E)$, thus the total variation is a radon measure. Then the proof of existence of function f such that L is equal to integration with respect to $f|L|$ and $|f| = 1$ $|L|$-a.e. can be found on pages 34-41 [Maggi, 2012]. Let's check that it's an isometry. Let measure μ be represented by a functional L. Then

$$\|L\| = \sup\{L(f) \mid \|f\| \leq 1\}$$
$$\|\mu\| = \sup\{\sum |\mu(B_i)| \mid \{B_i\} \text{ a partition of } X\}$$

For $f \in C_0(X, \mathbb{R}^n)$ we find a series of step functions $f_n = \sum a_i \chi_{B_i}$ such that $f_n \to f$ and $a_i \leq 1$. Thus by dominant convergence $\int f_n \, d\mu \to L(f)$. On the other hand we have

$$| \int f_n \, d\mu = | \sum a_i \cdot \mu(B_i)| \leq \sum |a_i||\mu(B_i)| \leq \sum |\mu(B_i)| \leq \|\mu\|$$

and thus we have $| \int f \, d_\mu| \leq \|\mu\|$ and thus $\|L\| \leq \|\mu\|$.
In the other direction it obviously follows form the fact the C_c^0 is dense in L^1.

Corollary: *Continuous positive functionals are represented by Radon measures*

Because $f = 1$.

3.4 Interpretation of Banach-Alaoglu theorem for vector valued measures

The weak-$*$ convergence can be interpreted as convergence of evaluation of measure on every continuous function on compact sets.

The original statement of Banch-Alaoglu theorem is **the closed unit ball $B = \{f \in E^* \,|\, |f| \leq 1\}$ is compact in the weak-$*$ topology**. If we replace the termes in this proof by measure terms we have the following theorem

Banach-Alaoglu Theorem for $\mathcal{M}(X, \mathbb{R}^n)$: *The set $B = \{\mu \in \mathcal{M}(X, \mathbb{R}^n) \,|\, \|\mu\| \leq C\}$ is compact for every $C \in \mathbb{R}_{>0}$. That's said every bounded sequence of vector measures has a weakly-$*$ converging subsequence.*

Consequence: If (μ_n) is a bounded sequence of vector measures, then it has a converging subsequence.

3.5 Weak-$*$ convergence of measures

Proposition: *Let $(\mu_n)_n$ be a sequence of positives measures converging to μ, then we have*

1. *For any open subset $A \subseteq X$, $\liminf \mu_n(A) \geq \mu(A)$*

2. *For any compact subset $K \subseteq X$, $\limsup \mu_n(K) \leq \mu(K)$*

3. *For any relatively compact $E \subseteq X$ such that $(\partial E) = 0$*

Proof: We will simultaneously demonstrate propositions 1 and 2. Let $K \subset A$, where K is compact and A is open. Consider a function $f \in C_c(X)$ such that $\chi_K \leq f \leq \chi_A$. For a Radon measure ν, we then have:

$$\nu(K) \leq \int f \, d\nu \leq \nu(A)$$

And by considering the limits, we obtain:

$$\limsup \mu_i(K) \leq \limsup \int f \, d\mu_i = \int f \, d\mu \leq \mu(A)$$

$$\mu(K) \leq \int f \, d\mu = \liminf \int f \, d\mu_i \leq \liminf \mu_i(A)$$

Since we are dealing with Radon measures and these inequalities hold for every compact K and every open A, we can pass to the limit. The lines then transform into:

$$\limsup \mu_i(K) \leq \mu(K)$$

$$\mu(A) \leq \liminf \mu_i(A)$$

Point three is a consequence of the two preceding points. Indeed, we have:

$$\limsup \mu_i(\overline{E}) \leq \mu(\overline{E}) = \mu(\text{int}(E)) \leq \liminf \mu_i(\text{int}(E))$$

Remark: The third proposition gives us the property that the measures of a sequence of figures converge to the other figure; thus, locally in the ball, the area also converges.

The following sections draw highly from the lecture notes of [Alberti, 2017]. Most propositions are adopted from this source, however remarks or most proofs are not.

4 Tangents

Similarly to projective spaces $\mathbb{R}P^n$ one can generalise this notion to smaller subspaces than hyperplanes. The set of m dimensional subspaces of a vector space $\mathbb{R}^n$ is called Grassmannian and noted by $G(m, n)$. It is endowed with structures identified from orthogonal projections on m-dimensional subspaces.

In this section, we will finally apply measure theory to the study of smooth geometric properties.

4.1 Approximate Tangent

Definition: *Let α be a fixed angle, let $x \in \mathbb{R}^n$ be a point and let V be a n-dimensional plane in $\mathbb{R}^{n+m}$. The* **cone of angle α around V centered at x** *is defined by setting*

$$\mathcal{C}(x, V, \alpha) = \{x' \in \mathbb{R}^{n+m} \mid |x' - x| \sin(\alpha) \geq d(x - x', V)\}$$

Definition: *Let $V \in G(n + m, n)$ be a d-dimensional plane. If E is a Borel set and $x \in E$ a point, then V is a* **strong tangent plane** *to E at x if and only if for every $\alpha > 0$ there exists a positive radius $r_0 > 0$ such that*

$$E \cap B(x, r_0) \subseteq C(x, V, \alpha)$$

Definition: *Let $V \in G(n + m, n)$ be a d-dimensional plane. If E is a Borel set and $x \in E$ a point, then V is an* **approximate tangent plane** *to E at x if and only if for every $\alpha > 0$ it turns out that*

$$\mathcal{H}^d((E \cap B(x, r)) \setminus \mathcal{C}(x, V, \alpha)) = o(r^d)$$

and

$$\mathcal{H}^d((E \cap B(x, r)) \cap \mathcal{C}(x, V, \alpha)) \sim \omega_d r^d$$

Finally, one can define a tangent space using weak-* convergence. Spaces satisfying this definition are generally also called approximate, but to reduce confusion, here I will call them **limit spaces**.

Definition: *Let $\psi_{x,r} : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m} = x' \mapsto \frac{x' - x}{r}$ be the dilation map. And let $E_{x,r}$ be the image of E under $\psi_{x,r}$. An n-dimensional plane V is a* **limit plane** *to the set E at point x if and only if*

$$\mathcal{H}^n \llcorner E_{x,r} \rightharpoonup \mathcal{H}^n \llcorner V$$

Proposition: *A limit plane is an approximate tangent plane.*

Proof: Let V be a limit plane of E at x. Let $\mu := \mathcal{H}^n \llcorner V$ and $\mu_r := \mathcal{H}^n \llcorner E_{x,r}$. Since μ and μ_r are Radon measures and $B(0, 1)$ is relatively compact and its boundary is μ-negligible, then $\mu_r(B(0, 1)) \to_{r \to 0} \mu(B(0, 1)) = \omega_n$. Furthermore, we have

$$\mu_{x,r}(B(0, 1)) = \mathcal{H}^n(\psi_{x,r}[E \cap B(x, r)]) = \frac{1}{r^n} \mathcal{H}^n(E \cap B(x, r))$$

and thus

$$\mathcal{H}^n(E \cap B(x, r)) \sim \omega_n r^n$$

If, in the preceding constructions, we replace $B(0, 1)$ by $B(0, 1) \cap \mathcal{C}(0, V, \alpha)$, we find

$$\mathcal{H}^n(E \cap B(x, r) \cap \mathcal{C}(x, V, \alpha)) \sim \omega_n r^n$$

And if we take the difference of these two equalities by dividing them by r^n, we find that

$$\frac{\mathcal{H}^n(E \cap B(x, r)) - \mathcal{H}^n(E \cap B(x, r) \cap \mathcal{C}(x, V, \alpha))}{r^n} = \frac{\mathcal{H}^n(E \cap B(x, r) \setminus \mathcal{C}(x, V, \alpha))}{r^n} \sim 0$$

And therefore the plane is approximate.

Actually we can completely eliminate sets and talk about tangent spaces to measures.

Definition: We are saying that μ has a tangent $T \in G(n, m)$ at x if for measures $\mu_{r,x}(A) := r^{-m} \mu(x + rA)$ we find a constant $k \in (0, +\infty)$ such that $\mu_{r,x} \rightharpoonup \mathcal{H}^m|_T$ as $r \to 0^+$.

4.2 Tangent Bundle

Proposition: *Let $\Sigma, \Sigma' \subseteq R^{n+m}$ be n-dimensional surfaces of class C^1. Then tangent planes are equal at $\mathcal{H}^n$-almost every point in the intersection $x \in \Sigma \cap \Sigma'$.*

To prove it, we take a point $x \in \Sigma \cap \Sigma'$ such that $T_x\Sigma \neq T_x\Sigma'$. Then, locally at x surfaces are represented by submersions $F, G : \mathbb{R}^{m+n} \to \mathbb{R}^m$ i.e, $\Sigma \cap A = F^{-1}(0) \cap A$ and $\Sigma' \cap B = G^{-1}(0) \cap B$, where A and B are open neighborhoods of x.

Let's introduce a new function $(F, G) : \mathbb{R}^{m+n} \to \mathbb{R}^{2m} = x \mapsto (F(x), G(x))$. The differential of (F, G) is a matrix of 2 blocks, one above the other. They are placed vertically because, actually, the pair (F, G) is a column and we have $D(F, G) = (DF, DG)^t$. Then $A \cap B \cap \Sigma \cap \Sigma' = (F, G)^{-1}(0)$ and we have a representation of an intersection. Remark that (F, G) is not necessarily a submersion. Let's take in the differential of (F, G) indices $(i_n)_{n \in [\![1,M]\!]}$ of a maximally linear independent set of rows. Its cardinal is at least n because the rows in the differential of F are independent, and it is strictly bigger, because otherwise the tangent spaces at x would coincide.

$$D\left(\begin{array}{c} F \\ G \end{array}\right) = \left(\begin{array}{c} \vdots \\ DF_i \\ \vdots \\ DG_k \\ \vdots \end{array}\right)$$

where $F_i = \pi_i \circ F$ and $G_k = \pi_k \circ G$ are coordinate functions. Then, if we retain only those rows in (F, G) we will have a submersion H

$$H = \left(\left(\begin{array}{c} F \\ G \end{array}\right)_j\right)_{j \in (i_n)}$$

Thus, we have $H : \mathbb{R}^{n+m} \to \mathbb{R}^M$, where $m < M < n + m$ is the rank of H at x. Hence, we obtain $n + m - M < n$ dimensional surface $H^{-1}(0) \cap A \cap B = \Sigma''$ and $\Sigma \cap \Sigma' \cap A \cap B \subset \Sigma''$, because $(F, G)(z) = 0 \Rightarrow H(z) = 0$. Thus $\Sigma \cap \Sigma' \cap A \cap B$ has null $\mathcal{H}^n$ measure.

Finally, we have showed that the target set $S = \{x \in \Sigma \cap \Sigma' \mid T_x\Sigma \neq T_x\Sigma'\}$ around each point has an open ball where its measure is null. Since from every open cover we can extract a countable subcover (because our space is separable), we have proven that the entire set is $\mathcal{H}^n$-null.

Lemma: *Let $f, g \in C^1(\mathbb{R}^n, \mathbb{R})$, then $\nabla f = \nabla g$ $\mathcal{L}^n$-a.e. on $\{f = g\}$.*

For dimensions $n > 1$ it's sufficient to see that points where gradients are not equal form a 1 dimensional surface and its Lebesgue's measure is 0.
For a 1 dimensional case we set $h = f - g \in C^1$. Then we consider a closed set $S = \{h = 0\}$. Let $x \in S$ be such that $\nabla h(x) \neq 0$, then by mean value theorem we find a neighborhood of x that contains only one such x ($\nabla h(x) = 0$). Thus the set of such x is countable and its measure is 0.

Definition: *Let E be a Borel set. A map T from E to the Grassmannian manifold $G(n, d)$ that sends x to $T(x)$ is a **weak tangent bundle** for the set E if and only if for every Σ d-dimensional surface of class C^1 it turns out that $T_x\Sigma = T(x)$ for $\mathcal{H}^d$-almost every $x \in \Sigma \cap E$.*

5 Countably n-rectifiable sets

*Let $M \subseteq X$ be a subset of a metric space. Then M is called n-**rectifiable** if*

$$M \subseteq M_0 \cup \bigcup f_i[\mathbb{R}^n]$$

where $\mathcal{H}^n(M_0) = 0$ and f_i are Lipschitz functions.

Remarque: *Hausdorff dimension of d-rectifiable set is less or equal to d*

This is true due to the fact that Lipschitz maps does not increase the dimension.

Criteria of Rectifiability: *Let $X = \mathbb{R}^{n+m}$, and let $M \subseteq X$ be a Borel set. The following assertions are equivalent:*

1. *The set M is n-rectifiable*

2. *There exist open sets A_i, M_0 $\mathcal{H}^n$-null set and differentiable functions $f_i : A_i \to X$ such that*

$$M \subseteq M_0 \cup \bigcup f_i[A_i]$$

3. There exist open sets A_i, M_0 $\mathcal{H}^n$-null set and diffeomorphisms $f_i : A_i \to X$ such that

$$M \subseteq M_0 \cup \bigcup f_i[A_i]$$

4. There exist n-dimensional surfaces $\Sigma_i \subseteq X$ and M_0 $\mathcal{H}^n$-null set such that

$$M \subseteq M_0 \cup \bigcup \Sigma_i$$

Proposition: *A d-rectifiable Borel set $E \subseteq \mathbb{R}^n$ admits a unique up to $\mathcal{H}^d$-null sets a weak tangent bundle.*

Proof: We have $E \subseteq M_0 \cup \bigcup \Sigma_i$, hence we can define a bundle as following. For $x \in M_0$ we can take what ever we want, for $x \in \Sigma_1$ we take $T_x\Sigma_1$ and for $x \in \Sigma_s \setminus \bigcup_{i=1}^{s-1} \Sigma_i$ we take $T_x\Sigma_s$. This is a necessary condition as planes should be a.e. equal to the planes tangent to those surfaces. The condition for a weak tangent bundle is satisfied due to the previous preposition.

Theorem: *If E is a Borel, n-rectifiable, $\mathcal{H}^n$- locally finite set, then the weak tangent bundle $T(x)$ is the limit plane to E at x for $\mathcal{H}^d$-almost every $x \in E$.*

Proof: We will show that $T_x\Sigma_i$ is a **limit plane** to E at x for $\mathcal{H}^n$-almost every $x \in E \cap \Sigma_i$. Associated with this plane, we consider four measures: $\mu_{x,r} := \mathcal{H}^n \llcorner E_{x,r}$, $\nu_{x,r} := \mathcal{H}^n \llcorner \Sigma_{i,x,r}$, $\eta_{x,r} := \mathcal{H}^n \llcorner (\Sigma_i \setminus E)_{x,r}$ and $\sigma_{x,r} := \mathcal{H}^n \llcorner (E \setminus \Sigma_i)_{x,r}$. We then observe that $\mu_{x,r} = \nu_{x,r} - \eta_{x,r} + \sigma_{x,r}$.

At x, the surface Σ_i is locally represented by an immersion $\phi : T_x\Sigma_i \cap U \to \Sigma_i \cap V$. We can assume that $D\phi(0) = \text{Id}$ and that $B(0,1) \subseteq U, V$. Let $f \in \mathcal{C}_c$, without loss of generality, we can assume that $\text{spt}(f) \subseteq B(0,1)$. Thus, $\psi_{x,r} \circ \phi(h) = (h + o(h))/r$. If we only take $h < r$, we find that $\phi_{x,r} = \psi_{x,r} \circ \phi|_{B(0,r)} \circ \psi_{0,1/r} : B(0,1) \to \Sigma_{i,x,r}$ is given by $h \mapsto (rh + |rh|\epsilon(rh))/r = h + |h|\epsilon(rh)$. Moreover, the differential $D\phi_{x,r}$ converges to the identity:

$$D\phi_{x,r} = rD\phi|_{B(0,r)} 1/r = D\phi|_{B(0,r)} \to \text{Id}$$

Consequently, the integral converges:

$$\int f d\nu_{x,r} = \int_{\Sigma_i \cap B(x,r)} f(s) d\mathcal{H}^n(s) = \int_{B(0,1)} f(\phi_{x,r}(s)) J\phi_{x,r}(s) ds \to_{r \to 0} \int_{T_x\Sigma_i} f(s) ds$$

Thus, we have the weak convergence of measures:

$$\nu_{x,r} \rightharpoonup \mathcal{H}^n \llcorner T_x\Sigma_i$$

Next, we observe that $\lambda_r \rightharpoonup 0 \Leftrightarrow \lambda_r(B_R) \to 0$ for all radii R.

For the measures $\eta_{x,r}$ and $\sigma_{x,r}$, it suffices to consider the case $B(0,1)$, because we are blowing up figures anyway. For $\eta_{x,r}$, we have:

$$\eta_{x,r}(B(0,1)) = \mathcal{H}^n(B(0,1) \cap (\Sigma_{i,x,r} \setminus E_{x,r})) = \frac{1}{r^n} \mathcal{H}^n(B(x,r) \cap (\Sigma_i \setminus E)) \to 0$$

This holds for almost all x, by the first property of the upper Hausdorff measure density, because $x \notin \Sigma_i \setminus E$.

Finally, for $\sigma_{x,r}$, we observe that:

$$\mu_{x,r}(B(0,1)) = \nu_{x,r}(B(0,1)) - \eta_{x,r}(B(0,1)) + \sigma_{x,r}(B(0,1))$$

By passing to the limit, we obtain:

$$\lim_{r \to 0} \mu_{x,r}(B(0,1)) = \omega_n - 0 + \lim_{r \to 0} \sigma_{x,r}(B(0,1))$$

And since, by the second density property, $\limsup_{r \to 0} \mu_{x,r}(B(0,1)) \leq \omega_n$ almost everywhere, we find that $\lim_{r \to 0} \sigma_{x,r}(B(0,1)) = 0$.
Thus, we have the weak convergence:

$$\mu_{x,r} = \nu_{x,r} - \eta_{x,r} + \sigma_{x,r} \rightharpoonup \mathcal{H}^n \llcorner T(x)$$

for almost all x.

Proposition: *Let $M \subseteq X$ be a Borel set with finite n-Hausdorff measure. Then $M = M_r \cup M_u$, where M_r is rectifiable and M_u is unrectifiable.*

Theorem: *Let $E \subseteq \mathbb{R}^{n+m}$ be a Borel set. If E is a d-rectifiable $\mathcal{H}^d$-locally finite set, then the weak tangent bundle $T(x)$ is the approximate tangent plane to E at x for $\mathcal{H}^d$-almost every $x \in E$.*

6 Differential calculus results

In this section, we re-examine C^1 n-dimensional submanifold M of $\mathbb{R}^{n+m}$ to observe several properties of standard constructions. This will then motivate the definition for analogous constructions for measures.

We can define a gradient ∇ through its representation

$$\langle \nabla f, \cdot \rangle = Df$$

Then we define a gradient ∇^M associated with M by

$$\nabla^M f(x) := \pi_{T_x M}(\nabla f(x))$$

and if we have an orthogonal basis $(\mathbf{w}_i)$ of $T_x M$

$$\nabla^M f(x) = (\mathbf{w}_1, \ldots, \mathbf{w}_n)\begin{pmatrix} D_x f(\mathbf{w}_1) \\ \vdots \\ D_x f(\mathbf{w}_1) \end{pmatrix} = (\mathbf{w}_i)(D_x f(\mathbf{w}^i))$$

for an orthogonal basis $(\mathbf{w}_i)$ of TM. If we take that for a definitions, then using matrix notation, it's also easy to show that this notion is independent of the chosen basis, which is already evident from coordinate-free definition. If we have another orthogonal basis $(\mathbf{w}_i) = (\mathbf{u}_i)P$, where P is orthogonal, then $(\mathbf{w}^i) = P^t(\mathbf{u}^i)$ and thus $(\mathbf{w}_i)(Df(\mathbf{w}^i)) = (\mathbf{u}_i)P(DfP^t(\mathbf{u}^i)) = (\mathbf{u}_i)PP^t Df(\mathbf{u}^i) = (u_i)Df(u^i)$

If for two vectors a, b by ab we right their scalar product, then we can introduce divergence of $X : M \to \mathbb{R}^{n+m}$ on M by choosing a basis $(\mathbf{w}_1, \ldots, \mathbf{w}_n)$ of $T_x M$ and setting

$$\operatorname{div}_M X(x) = (\mathbf{w}_i)(\nabla^M X^i(x))$$

where $X = (X^i)_{i \in [\![1,n+m]\!]}$ are coordinates in orthogonal extension of (w_i). We will need a following theorem

Theorem: *Let $M \subset \mathbb{R}^{n+m}$ be a n-dimensional C^1 bounded submanifold-with-boundary. Then we have a following result*

$$\int_M \operatorname{div}_M X d\mathcal{H}^n = \int_{\partial M} X \nu d\mathcal{H}^{n-1} \text{ for every } X \in C^1(M, TM)$$

where ν is an orthogonal unitary vector to the boundary pointing outwards.
This theorem, proposed in [Simon, 1983], is crucial for further development of geometric measure theory notions.

We will now present a proof for the C^2 case, as this regularity allows us to utilize coordinates and integrate by parts. For this section, we assume that $X(x) \in T_x M$. Let us the consider an orthonormal basis $(\mathbf{w}_i)$ of $T_x M$ which is extended by vectors $(\mathbf{w}_i')$ to an orthonormal basis of $\mathbb{R}^{m+n}$. Furthermore, consider curvilinear coordinates $(c^i) : M \cap U \to \mathbb{R}^n \cap V$ (C^2-diffeomorphism), to which we can associate a standard basis

$$\mathbf{c}_i = \mathbf{w}_j \frac{\partial w^j}{\partial c^i} + \mathbf{w}'_j \frac{\partial w'^j}{\partial c^i}$$

which consists of tangents to the coordinate curves. Thus, $(\mathbf{c}_i)$ form a basis, whose vectors vary with position and are associated with coordinates. In coordinate-free language this basis is an image of standard basis $\mathbf{e}_i$ of $\mathbb{R}^n$ by $(D_x(c^i))^{-1}$. By (w^i) here, we denote the standard coordinates associated with basis $(\mathbf{w}_i)$. In the following discussion all expressions are evaluated at a point x; moving to a different point necessitates selecting a new tangent basis. Given that $(\mathbf{c}_i)$ varies smoothly ($\mathbf{C}^1$), we can differentiate them and naturally define the following coefficients, called Christoffel symbol:

$$\mathbf{c}_k \Gamma_{ji}^k = \pi_{T_x M}\left(\frac{\partial \mathbf{c}_i}{\partial c^j}\right)$$

Indeed, we observe symmetry in lower indices since

$$\pi_{T_xM}\left(\frac{\partial \mathbf{c}_i}{\partial c^j}\right) = \frac{\partial}{\partial c^j}\left(\mathbf{w}_l \frac{\partial w^l}{\partial c^i}\right) = \mathbf{w}_l \frac{\partial^2 w^l}{\partial c^j \partial c^i} = \mathbf{w}_l \frac{\partial^2 w^l}{\partial c^i \partial c^j} = \Gamma_{ij}^k$$

Also in a new basis $(\mathbf{c}_k)$ we have scalar product, which is usually denoted by

$$\langle \mathbf{c}_l, \mathbf{c}_k \rangle = g_{lk} = \frac{\partial w^i}{\partial c^l}\frac{\partial w^i}{\partial c^k}$$

which is also symmetric. If we then rewrite the equality for Γ, we get

$$\mathbf{w}_j \frac{\partial^2 w^j}{\partial c^j \partial c^i} = \mathbf{c}_k \Gamma_{ji}^k = \mathbf{w}_j \frac{\partial w^j}{\partial c^k}\Gamma_{ji}^k$$

The metric defines an isomorphism ϕ from our space T_xM to its dual space T_xM^* by

$$\phi(\mathbf{c}_i) := \langle \mathbf{c}_i, \cdot \rangle = g_{lm}\mathbf{c}^{*l} \otimes \mathbf{c}^{*m}\mathbf{c}_i = g_{li}\mathbf{c}^{*l}$$

The components g_{ij} form a matrix representation of ϕ in local coordinates. Consequently, the matrix of ϕ^{-1} is $(g_{ij})^{-1}$. Associated with this ismorphism, we can then construct a dual product by

$$\langle \cdot, \cdot \rangle^* := \phi^{-1}(\langle \cdot, \cdot \rangle) = \phi^{-1}(g_{lm}\mathbf{c}^{*l} \otimes \mathbf{c}^{*m}) = g_{lm}\phi^{-1}(\mathbf{c}^{*l}) \otimes \phi^{-1}(\mathbf{c}^{*m})$$
$$= g_{lm}(g_{lr}^{-1}\mathbf{c}_r) \otimes (g_{mk}^{-1}\mathbf{c}_k) = \delta_m^r \mathbf{c}_r \otimes g_{mk}^{-1}\mathbf{c}_k = g_{rk}^{-1}\mathbf{c}_r \otimes \mathbf{c}_k$$

Usually we write $\langle \cdot, \cdot \rangle^* = g^{mk}\mathbf{w}_m \otimes \mathbf{w}_k$. Let X' denote X rewritten in new coordinates, i.e.

$$\mathbf{X}(w^1, ..., w^n) = \mathbf{X}'(...c^i(w^1, ..., w^n)...)$$

Now we take a partial derivative of both sides. In this step, it is crucial that X has values tangent space, as this allows to write X in the curvilinear basis

$$\frac{\partial}{\partial w^j}\mathbf{w}_i X^i = \pi_{T_xM}\left(\frac{\partial}{\partial w^j}\mathbf{c}_p X'^p\right) = \pi_{T_xM}\left(\frac{\partial}{\partial c^k}\mathbf{c}_p X'^p\right)\frac{\partial c^k}{\partial w^j} = \left(\mathbf{c}_p \frac{\partial X'^p}{\partial c^k} + \pi_{T_xM}\left(\frac{\partial \mathbf{c}_p}{\partial c^k}\right)X'^p\right)\frac{\partial c^k}{\partial w^j}$$
$$= \left(\mathbf{c}_p \frac{\partial X'^p}{\partial c^k} + \mathbf{c}_l\Gamma_{pk}^l X'^p\right)\frac{\partial c^k}{\partial w^j} = \left(\mathbf{c}_p \frac{\partial X'^p}{\partial c^k} + \mathbf{c}_p\Gamma_{lk}^p X'^l\right)\frac{\partial c^k}{\partial w^j} = \mathbf{c}_p\left(\frac{\partial X'^p}{\partial c^k} + \Gamma_{lk}^p X'^l\right)\frac{\partial c^k}{\partial w^j}$$

Then we can calculate divergence formula at x by

$$\text{div}\mathbf{X} = \mathbf{w}_j \cdot \frac{\partial}{\partial w^j}\mathbf{w}_i X^i = \mathbf{w}_j \cdot \mathbf{c}_p\left(\frac{\partial X'^p}{\partial c^k} + \Gamma_{l,k}^p X'^l\right)\frac{\partial c^k}{\partial w^j} = \mathbf{w}_j \cdot \mathbf{w}_i\frac{\partial w^i}{\partial c^p}\left(\frac{\partial X'^p}{\partial c^k} + \Gamma_{l,k}^p X'^l\right)\frac{\partial c^k}{\partial w^j}$$
$$= \frac{\partial c^k}{\partial w^j}\frac{\partial w^j}{\partial c^p}\left(\frac{\partial X'^p}{\partial c^k} + \Gamma_{l,k}^p X'^l\right) = \frac{\partial c^k}{\partial c^p}\left(\frac{\partial X'^p}{\partial c^k} + \Gamma_{l,k}^p X'^l\right) = \frac{\partial X'^p}{\partial c^p} + \Gamma_{l,p}^p X'^l$$

For the next part we are differentiating g_{lm}

$$\frac{\partial}{\partial c^k}g_{lm} = \left(\frac{\partial}{\partial c^k}\frac{\partial w^i}{\partial c^l}\right)\frac{\partial w^i}{\partial c^m} + \frac{\partial w^i}{\partial c^l}\left(\frac{\partial}{\partial c^k}\frac{\partial w^i}{\partial c^m}\right) = \frac{\partial w^i}{\partial c^r}\Gamma_{kl}^r\frac{\partial w^i}{\partial c^m} + \frac{\partial w^i}{\partial c^l}\frac{\partial w^i}{\partial c^r}\Gamma_{km}^r = g_{mr}\Gamma_{kl}^r + g_{lr}\Gamma_{km}^r$$

Now lets consider a following quantity

$$\frac{\partial g_{kl}}{\partial c^m} + \frac{\partial g_{km}}{\partial c^l} - \frac{\partial g_{ml}}{\partial c^k} = g_{kr}\Gamma_{ml}^r + g_{lr}\Gamma_{mk}^r + g_{kr}\Gamma_{lm}^r + g_{mr}\Gamma_{lk}^r - g_{mr}\Gamma_{kl}^r - g_{lr}\Gamma_{km}^r = 2g_{kr}\Gamma_{ml}^r$$

And if we multiply both sides by $1/2g^{ak}$ we get

$$g^{ak}g_{kr}\Gamma_{ml}^r = \delta_r^a \Gamma_{ml}^r = \Gamma_{ml}^a = \frac{1}{2}g^{ak}\left(\frac{\partial g_{kl}}{\partial c^m} + \frac{\partial g_{km}}{\partial c^l} - \frac{\partial g_{ml}}{\partial c^k}\right)$$

Now lets take Cristoffel symbol from divergence formula and develop it

$$\Gamma_{pl}^p = \frac{1}{2}g^{pk}\left(\frac{\partial g_{kl}}{\partial c^p} + \frac{\partial g_{kp}}{\partial c^l} - \frac{\partial g_{pl}}{\partial c^k}\right) = \frac{1}{2}g^{pk}\frac{\partial g_{kp}}{\partial c^l}$$

Now knowing that $\partial_k(\det A) = \det(A)\text{tr}(A^{-1}\partial_k A)$ and writing $g = \det(g_{ij})$ we have

$$g^{pk}\partial_l g_{kp} = \text{tr}((g_{kp})^{-1}\partial_l(g_{kp})) = \frac{\partial_l g}{g}$$

And since $\partial_l \sqrt{g} = \frac{1}{2\sqrt{g}} \partial_l g$ we have

$$\Gamma^p_{pl} = \frac{1}{\sqrt{g}} \partial_l \sqrt{g}$$

and finally we can rewrite divergence formula as

$$\text{div} \mathbf{X} = \partial_p X'^p + \frac{1}{\sqrt{g}} \partial_l \sqrt{g} X'^l = \frac{1}{\sqrt{g}} \sqrt{g} \partial_p X'^p + \frac{1}{\sqrt{g}} \partial_p \sqrt{g} X'^p = \frac{1}{\sqrt{g}} \partial_p (\sqrt{g} X'^p)$$

Proof: Since M is compact we can find a finite open cover U_i of M such that at every U_i we find a local coordinate system. We can associated a C^∞ partition of unit (ψ_i) on M associated to U_i such that $\psi_i \in C^\infty_c(U_i)$ and $\sum \psi_i = 1_M$. We shall write $O_i = U_i \cap M$. And by linearity of equation we want to prove, we can prove it just for $\psi_i X$, or without loss of generality we will write just X. Let coordinates take values in V_i

If patch U_i is disjoint from ∂M, then V_i are open and X is of a compact support inside V_i and we can integrate by parts

$$0 = \int_{V_i} \partial_j 1 X^j \sqrt{g} \, dc = -\int_{V_i} 1 \partial_j (X^j \sqrt{c}) = -\int_{V_i} \frac{1}{\sqrt{g}} \partial_j (X^j \sqrt{g}) \sqrt{g} \, dc = -\int_{U_i} \text{div}_M X \, d\mathcal{H}^n$$

Thus inner patches have no contribution. Lets take a patch U_i that intersects boundary. By definition of submanifold-with-boundary we can introduce coordinates c^i such that $\{c_n = 0\} = \partial M \cap U_i$ and c_n has values only in $(-\infty, 0]$. This time X does not have a compact support inside V_i and thus in integration by parts we have the 2 terms

$$0 = \int_{V_i} \partial_j 1 X^j \sqrt{g} \, dc = -\int_{U_i} \text{div}_M X \, d\mathcal{H}^n + \int_{\mathbb{R}^{n-1}} X^n(c', 0) \sqrt{g(c', 0)} \, dc'$$

And as we can chose such coordinates, that $X^n(c', 0) = \nu \cdot X(c', 0)$ we have

$$\int_{U_i} \text{div}_M X \, d\mathcal{H}^n = \int_{\partial M} \nu \cdot X \, d\mathcal{H}^{n-1}$$

Definition: *Let $\mathbf{v}_i$ be an orthonormal basis of $T_y M$ which is C^1 function of y. Then we can define **the second fundamental form** $B_y : T_y M \times T_y M \to (T_y M)$ by setting $B_y(t, n) := -(n \cdot D_y(\mathbf{v}_i)(t))\mathbf{v}_i$*

Lets verify that this definition is independent from the choice of $\mathbf{v}_i$, thus let $\mathbf{w}_i$ be a different orthonormal basis. Then they are related with an orthogonal matrix $\mathbf{v}_i = O_i^j \mathbf{w}_j$. And if we compute coordinate change we get

$$(n \cdot D_y(\mathbf{v}_i)(t))\mathbf{v}_i = (n \cdot D_y(O_i^j \mathbf{w}_j)(t))\mathbf{O}_i^k \mathbf{w}_k = (n \cdot D_y(O_i^j)(t)\mathbf{w}_j + O_i^j n \cdot D_y(\mathbf{w}_j)(t))\mathbf{O}_i^k \mathbf{w}_k$$

$$= (O_i^j O_i^k (n \cdot D_y(\mathbf{w}_j)(t))\mathbf{w}_k) = (n \cdot D_y(\mathbf{w}_k)(t))\mathbf{w}_k$$

Let $\gamma : I \to M$ be a C^2 curve. It's tangent is $t = \gamma'$ and its curvature is $\kappa = \gamma''$ and we suppose $|t| = 1$. Now lets differentiate a following expression

$$0 = \frac{d}{ds}(t \cdot \mathbf{v}_i) = \frac{d}{ds} t \cdot \mathbf{v}_i + t \cdot \frac{d}{ds} \mathbf{v}_i = \kappa \cdot \mathbf{v}_i + t \cdot D(v_i)(t)$$

Thus we get coordinates for normal curvature $\kappa_N = -(t \cdot D(\mathbf{v}_i)(t))\mathbf{v}_i = B(t, t)$

Similarly, if $\phi : U \subset \mathbb{R}^2 \to M$, then

$$0 = \frac{\partial}{\partial x_2}(\frac{\partial}{\partial x_1} \phi \cdot \mathbb{v}_i) = \frac{\partial^2}{\partial x_2 \partial x_1} \phi \cdot \mathbf{v}_i) + \frac{\partial}{\partial x_1} \phi \cdot \frac{\partial}{\partial x_2} \mathbf{v}_i(\phi(x_1, x_2)) = \frac{\partial^2}{\partial x_2 \partial x_1} \phi \cdot \mathbf{v}_i) + \frac{\partial}{\partial x_1} \phi \cdot D\mathbf{v}_i(\frac{\partial \phi}{\partial x_2})$$

And thus $B(\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2}) = (\frac{\partial^2 \phi}{\partial x_2 \partial x_1})^\perp$ and we remark that B is symmetric.

Definition: *We define the mean curvature vector $\mathbf{H}_y$ of M at y to be trace B_y; thus*

$$\mathbf{H}_y = B_y(\mathbf{t}_i, \mathbf{t}_i)$$

where $(\mathbf{t}_i)$ is an orthonormal basis of $T_y M$.

If we rewrite this definition and recall the definition of divergence, we obtain

$$\mathbf{H}_y = -(t_j \cdot D_{\mathbf{t}_j}(\mathbf{v}_i)(\mathbf{v}_i))\mathbf{v}_i = -\mathrm{div}_M(\mathbf{v}_i)\mathbf{v}_i$$

Let X be a vector field, then $X^\perp = (\mathbf{v}_i \cdot X)\mathbf{v}_i$ and if we take divergence on M, we get

$$\mathrm{div}_M X^\perp = (\mathbf{v}_i \cdot X)\mathrm{div}_M\mathbf{v}_i$$

because other components are in orthogonal space. Thus we conclude

$$\mathrm{div}_M X^\perp = -\mathbf{H} \cdot X$$

and recalling the integration theorem, we get

$$\int_M \mathrm{div}_M X d\mathcal{H}^n = \int_{\partial M} \nu \cdot X d\mathcal{H}^{n-1} - \int_M \mathbf{H} \cdot X d\mathcal{H}^n$$

7 Variations

Let's suppose that we have Σ an n-dimensional $\mathcal{C}^1$ submanifold of $\mathbb{R}^{n+m}$ and an open set U. Let's fix a $\mathcal{C}^2$ map $\phi : (-1, 1) \times U \to U$ and a closed subset $K \subset U$ such that:

1. $\phi_t := \phi(t, \cdot)$ is a $\mathcal{C}^1$-diffeomorphism for every t

2. $\phi_0 = \mathrm{id}$

3. $\phi_t|_{K^c} = \mathrm{id}$ for every t

Thus ϕ is a controlled geometrical variation of K. Let $V := \partial_t\phi(0, \cdot)$ be a velocity vector field and $A := \partial_t^2\phi(0, \cdot)$ an acceleration field, then

$$\phi(t, \cdot) = \mathrm{id} + tV + \frac{t^2}{2}A + o(t^2)$$

Lets note $M := \phi(\cdot, \Sigma \cap K)$ and $\psi_t := \phi_t|_\Sigma$

Definition: By first and second variations of families M we mean following derivatives $\mathcal{H}^n(M)'(0)$ and $\mathcal{H}^n(M)''(0)$.

If we recall the area formula then we can compute variations as a function of ϕ

$$\mathcal{H}^n(M(t)) = \int_{M(0)} (J\psi_t)d\mathcal{H}^n$$

Thus we construct a map

$$D_x\psi_t : T_x\Sigma \to \mathbb{R}^{n+m} = \mathrm{id} + tD_xV + \frac{t^2}{2}D_xA + o(t^2)$$

Lets chose an orthogonal basis τ for T_xM and an orthogonal basis e for $\mathbb{R}^{n+m}$. Then in those basis we can write matrix of $D_x\psi_t$.

$$[D_x\psi_t]_i^l = \tau_i^l + tD_xV^l(\tau_i) + \frac{t^2}{2}D_xA^l(\tau_i) + o(t^2)$$

To calculate $J\psi_t$ we need to know $(D_x\psi_t)^* \circ (D_x\psi_t)$

$$[(D_x\psi_t)^* \circ (D_x\psi_t)]_i^j = \sum_l [D_x\psi_t]_j^l [D_x\psi_t]_i^l = \sum_l (\tau_j^l + tD_xV^l(\tau_j) + \frac{t^2}{2}D_xA^l(\tau_j) + o(t^2))$$

$$(\tau_i^l + tD_xV^l(\tau_i) + \frac{t^2}{2}D_xA^l(\tau_i) + o(t^2))$$

$$= \tau_j \cdot \tau_i + t(\tau_j \cdot D_xV(\tau_i) + \tau_i \cdot D_xV(\tau_j)) + t^2(\frac{1}{2}(\tau_j \cdot D_xA(\tau_i) + \tau_i \cdot D_xA(\tau_j)) + D_xV(\tau_i) \cdot D_xV(\tau_j)) + o(t^2)$$

This can be abstracted as $I + tS + t^2T$ and we want to calculate a development of determinant of it. We have a Jacobi formula $\det(N(t))' = \det(N(t))\mathrm{tr}(N(t)^{-1}N'(t))$ and it's derivative

$$\det(N(t))'' = \det(N(t))'\mathrm{tr}(N(t)^{-1}N'(t)) + \det(N(t))\mathrm{tr}(-N(t)^{-1}N'(t)N(t)^{-1}N'(t) + N(t)^{-1}N''(t))$$

and is we replace $N(t) = I + tS + t^2T$, then $\det(I + tS + t^2T)' = \det(I + tS + t^2T)\mathrm{tr}((I + tS + t^2T)^{-1}(S + 2tT))$ and at zero we have a following equality $\det(N(t))'|_{t=0} = \mathrm{tr}(S)$ and for a second derivative we have

$$\det(N(t))''|_{t=0} = \mathrm{tr}(S)^2 - \mathrm{tr}(S^2) + 2\mathrm{tr}(T)$$

Thus we can write a Taylor Polynomial

$$\det(I + tS + t^2T) = 1 + t\mathrm{tr}(S) + \frac{t^2}{2}(2\mathrm{tr}(T) - \mathrm{tr}(S^2) + \mathrm{tr}(S)^2) + o(t^2)$$

In our case we are calculating $\det((D_x\psi_t)^* \circ (D_x\psi_t)) = (\mathrm{J}(\psi|_t)(x))^2$

$$\mathrm{tr}(S) = \sum_i (\tau_i \cdot D_xV(\tau_i) + \tau_i \cdot D_xV(\tau_i)) = 2\mathrm{div}_\Sigma V(x)$$

$$\mathrm{tr}(T) = \mathrm{div}_\Sigma A + \sum_i |DV(\tau_i)|^2$$

$$\mathrm{tr}(S^2) = (\tau_i \cdot D_xV(\tau_j) + \tau_j \cdot D_xV(\tau_i))(\tau_i \cdot D_xV(\tau_j) + \tau_j \cdot D_xV(\tau_i)) = 2(\tau_i \cdot D_xV(\tau_j))(\tau_j \cdot D_xV(\tau_i))$$

And for Taylor Polynomial of a root we have

$$\sqrt{1 + x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + o(x^2)$$

And composed Taylor Polynomial will be

$$\sqrt{1 + (at + bt^2)} = 1 + \frac{1}{2}at + \frac{t^2}{2}(b - \frac{1}{4}a^2) + o(t^2)$$

In our case the value for a is

$$a = 2\mathrm{div}_\Sigma V(x)$$

And thus we can express the first variation

$$\det(M(t))'|_{t=0} = \mathrm{div}_\Sigma V(x)$$

For b the expression is a little longer

$$b = 2(\mathrm{div}_\Sigma V(x))^2 + (\mathrm{div}_\Sigma A + \sum_i |DV(\tau_i)|^2) - (\tau_i \cdot D_xV(\tau_j))(\tau_j \cdot D_xV(\tau_i))$$

And we have an expression for the second variation

$$\det(M(t))''|_{t=0} = (\mathrm{div}_\Sigma V(x))^2 + (\mathrm{div}_\Sigma A + \sum_i |DV(\tau_i)|^2) - (\tau_i \cdot D_xV(\tau_j))(\tau_j \cdot D_xV(\tau_i))$$

8 Varifold

An m-dimensional varifold V is a Radon measure over $\mathbb{R}^n \times G(n, m)$ endowed with a product topology. We say $\|V\|$ is a measure in $\mathbb{R}^n$ that is reciprocal projection of a varifold V by π_1^{-1}.

Proposition: *For varifolds we consider weak-∗ topology. Then we have a convergence criteria that $V_i \to V$ if and only if*

$$\int f \, dV_i \to \int f \, dV$$

for every continuous function $f : \mathbb{R}^n \times G(m, n) \to R$ with a compact support.

Remark: Varifolds allow us separate the notion of a tangent plane from geometric properties (a similar idea is actually used for normal maps in computer graphics). They allow us to consider possibly many planes at the same point with different masses and also to change the mass of geometric figures.

Thus, if we have a surface M, we can consider varifolds of the kind $\mathcal{H}^n \llcorner M \otimes T$, where T is a Radon measure on the Grassmannian. Or, if we want to be even more precise, we can have a tangent function T that provides a tangent at a given point and a mass function θ. Then, we consider a varifold of type $f \to \int f(x, T(x))\theta(x)d\mathcal{H}^n$. And we are free to choose tangents.

Sometimes, a choice for a tangent space can be made naturally. Then, for example, to an n-rectifiable set E, we can naturally associate a varifold $v(E, \theta)$ by introducing the following functional

$$\langle V_E \mid f \rangle := \int_E f(x, T(E, x)) \theta(x) d\mathcal{H}^m, \quad f \in \mathcal{C}_c(\mathbb{R}^n, G(m, n))$$

if $\theta = 1$ we note such varifold as $v(E)$, if $\theta \in \mathbb{Z}$, we say that varifold is *integral*.

Definition: *We say that a varifold V has a tangent space T with multiplicity $\theta \in (0, \infty)$, using a similar idea as for limit planes. That is to say, if*

$$V_{x,r} \rightharpoonup \theta v(T)$$

Where $V_{x,r}$ is a varifold passed through a dilation mapping at x with renormalisation coefficient, i.e.

$$V_{x,r}(A) := r^{-n} V(\{(ry + x, S) \mid (y, S) \in A\})$$

Definition: By analogy to the variation of submanifolds, one can define first variation of varifolds as a following functional

$$\delta V : X \in \mathcal{C}_c^1(\mathbb{R}^n, \mathbb{R}^n) \to \int_{\mathbb{R}^n \times G_{d,n}} \mathrm{div}_S X dV(x, s)$$

If it happens that first variation is continuous, which is true for rectifiable varifolds. We can decompose first variation by Riesz's representation theorem and Radon-Nykodim theorem into two parts, with respect to mass measure as

$$\delta V = -\mathbf{H} d\|V\| + \delta V^s$$

Where δV^s is a singular part of the decomposition, which is just a restriction of the first variation on the set where Radon-Nykodym derivative explodes. This part corresponds to the integration on the boundary for manifolds. Here, $\mathbf{H}$ is a generalized mean curvature, and in the case when the varifold is actually a rectifiable set, it is precisely its mean curvature because of the uniqueness of the decomposition.

8.1 Examples

Let's consider an arc of radius 1 of center $(0, 0)$ starting at $(1, 0)$ and ending at $(-1, 0)$ with a mass θ and call it A as a varifold, then

$$\mathrm{div}_A X = \partial_{\mathbf{e}_\alpha}(X^\alpha \mathbf{e}_\alpha + X^r \mathbf{e}_r) \cdot e_r = \partial_{\mathbf{e}_\alpha} X^\alpha + X^r$$

Then, the result for a first variation is

$$\delta A(X) = \int_A (\partial_{\mathbf{e}_\alpha} X^\alpha + X^r) \theta d\mathcal{H}^1 = \int_A \partial_{\mathbf{e}_\alpha} X^\alpha \theta d\mathcal{H}^1 + \int_A X^r d\mathcal{H}^1 = [X^\alpha \theta]_0^\pi - \int_A X^\alpha \partial_{\mathbf{e}_\alpha} \theta d\mathcal{H}^1 + \int_A X^r d\mathcal{H}^1$$

$$= [X^\alpha \theta]_0^\pi - \int_A X \cdot (\partial_{\mathbf{e}_\alpha} \theta \mathbf{e}_\alpha - \theta \mathbf{e}_r) d\mathcal{H}^1$$

Thus, a singular term is $[X^\alpha \theta]_0^\pi$ which corresponds to integration on the boundary with a correction for a new mass θ. Then, the curvature vector is $\mathbf{H} = \partial_{\mathbf{e}_\alpha} \theta \mathbf{e}_\alpha - \theta \mathbf{e}_r$, which deviates from a standard one by scaling in the orthogonal direction by mass θ and by a gradient of θ in the tangent direction.

Let's consider a segment $S = [a, b]$ of a horisontal line in $\mathbb{R}^2$, and lets construct a varifold of S which tangent plane is orthogonal to S at every point. Then,

$$\delta V(X) = \int_S \mathrm{div}_{S^\perp} X d\mathcal{H}^1 = \int_S (\partial_2 X) \cdot e_2 d\mathcal{H}^1 = \int_S (\partial_2 X) d\mathcal{H}^1 \cdot e_2 = \partial_2 \int_S X^2 d\mathcal{H}^1$$

And we get a perpendicular direction on the boundary. We can also remark that in this case the functional δV is no longer continuous for a supremum norm on the function space. We can have a sequence of functions that uniformly converges to 0, but perpendicular derivative is positive and uniformly converges to the infinity.

References

[Alberti, 2017] Alberti, P. G. (2017). *Lecture Notes on Geometric Measure Theory*. Pisa University.

[Lawrence C. Evans, 2015] Lawrence C. Evans, R. F. G. (2015). *Measure Theory and Fine Properties of Functions*. Textbooks in Mathematics. CRC Press.

[Maggi, 2012] Maggi, F. (2012). *Sets of Finite Perimeter and Geometric Variational Problems*. Cambridge Studies in Advanced Mathematics. Cambridge University Press.

[Simon, 1983] Simon, L. (1983). *Lecture Notes on Geometric Measure Theory*. Australian National University.