Flat Norm Monograph

Curtis Michels   Kevin Vixie

August 13, 2024

Abstract

An intuitive explanation of GMT and the Flat Norm.

1 Introduction

2 Applied GMT

2.1 Measure Theory

Mathematics is (at least in part) an exploration of spaces. We look at interesting structures and try to classify them to the best of our ability (when are two instances the same, when are they different, etc.) a more precise definition of what we mean when we say "space" in a mathematical sense follows:

Definition 2.1 (Space). *A space is a set possibly equipped with additional structure, often written as a tuple. The simplest example of a space is merely a set by itself.*

Example 2.1. *A metric space $X = (\Omega, d)$ is a set Ω (whose elements are often called points) together with a distance function $d : \Omega^2 \to \mathbb{R}$ that tells us the distance between any two points in Ω. To recover our intuition surrounding distance, we require d satisfy the following for all points $x, y, z \in \Omega$:*

1. *The distance from a point to itself is 0*

   $d(x, x) = 0$

2. *The distance between two distinct points is always positive*

   if $x \neq y$ then $d(x, y) > 0$

3. *The distance from x to y is always the same as the distance from y to x:*

   $d(x, y) = d(y, x)$

4. *The distance to go directly from x to y is less than or equal to the distance from x to an intermediate point z plus then the rest of the distance z to y.*

   $d(x, y) \leq d(x, z) + d(z, y)$

Remark 2.1. *Condition 2 is often referred to as "positvity", condition 3 is "symmetry" and condition 4 is called the "triangle inequality"*

Remark 2.2. *Even if a space X has additional structure, when we speak of "the elements of a space X" we are referring to the elements of the set Ω which underlies the space.*

Suppose we have a space X, and we would like some way to assign notions of length, area, or volume to the subsets of X. Denote the set of all subsets of X by 2^X. We would like to then develop a theory that yields useful theorems for a function $\mu : 2^X \to [0,\infty]$ which takes in a subset of X and assigns it to a nonnegative extended real number (possibly infinity).

Example 2.2. *Let $X = (\mathbb{R}^2, d)$ be the two dimensional Euclidean metric space where for two points $x = (x_1, x_2)$, and $y = (y_1, y_2) \in \mathbb{R}^2$ we have the Euclidean metric $d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$. Think about how we might assign areas to possible subsets of X, like the open balls of radius r centered at the origin $B(0,r) := \{x = (x_1, x_2) : x_1^2 + x_2^2 < r^2\}$ and other geometric shapes, without needing to specify them all by hand. i.e. we want a function that will tell us the area of **any** set. This is the problem we are trying to solve.*

The right place to look for such functions is measure theory, in particular the notion of a so-called **outer measure** will be useful to us.

Definition 2.2 (Outer Measure)**.** *Let X be any space. A function $\mu : 2^X \to [0,\infty]$ is called a "outer measure on X" if:*

1. *The measure of the empty set is 0*

   $\mu(\emptyset) = 0$

2. *If A is a subset of B then the measure of A is less than or equal to the measure of B*

   $A \subseteq B \implies \mu(A) \le \mu(B)$

3. *If we have a sequence of subsets of X then the measure of their union is less than or equal to the sum of their measures*

   $A_1, A_2, \ldots \subseteq X \implies \mu\left(\bigcup_{j=1}^{\infty} A_j\right) \le \sum_{j=1}^{\infty} \mu(A_j)$

Remark 2.3. *Condition 2 is sometimes referred to as "monotone" and condition 3 "subadditive"*

Remark 2.4. *Note that for some measures and choice of sets condition 3 may say nothing more than $\infty \le \infty$. No special care with regards to convergence is required to talk about the sum, since we are adding non-negative numbers up. It will either be finite or ∞, there is no oscillation to worry about.*

In working with the above definition for some time, we may recognize that there are certain subsets of a given space that are "nicer" in a sense to work with, that obey a deeper intuition. The following definition is due to Carathéodory who developed a way to identify these so-called measurable sets.

Definition 2.3. *Let μ be a outer measure on X. Then a subset $A \subseteq X$ is "μ-measurable if it reasonably "cuts" every other set. To be precise, for every set $B \subseteq X$ we must have:*

$$\mu(B) = \mu(B \setminus A) + \mu(B \cap A)$$

The above says that a set A is measurable if for every set B, the measure of the piece of A in B plus the measure of the chunk of B without A recovers the entire measure of B. You may think this definition a bit strange, and it may be difficult to think of measures on spaces and sets within them that do not have this property. And indeed, even the proof that sets that are not measurable exist in simple spaces (like the real number line equipped with standard calculus tools) is challenging and difficult. However, such non-measurable sets exist there, and furthermore are fundamental in a sense, as it may be proven that every reasonable measure of length on said number line yields sets that are not measurable.

It has yet to be seen why this is a good definition, but soon enough this will become clear.

Remark 2.5. *For any space X and outer measure μ on X, it is easy to check from the definition above:*

1. *The empty set and the entire space $\emptyset$, X are always μ-measurable.*

2. *All null sets (sets for which $\mu(A) = 0$) are μ-measurable.*

3. *If A is μ-measurable then $X \setminus A$ is also μ-measurable.*

Definition 2.4. *A family (set of sets) $\mathcal{A}$ of subsets of a space X is called a σ-algebra if:*

1. *$\emptyset, X \in \mathcal{A}$*

2. *If $A \in \mathcal{A}$ then so to is its relative complement $X \setminus A \in \mathcal{A}$*

3. *Any sequence $A_1, A_2, \ldots \in \mathcal{A}$ is closed in $\mathcal{A}$ under unions $\bigcup\limits_{j=1}^{\infty} A_j \in \mathcal{A}$*

Remark 2.6. *Although it is not part of the definition, the 3 conditions specified imply that if $A_1, A_2, \ldots \in \mathcal{A}$ then so too is their intersection $\bigcap\limits_{j=1}^{\infty} A_j \in \mathcal{A}$. This follows from De Morgan's laws regarding unions and compliments.*

From the definitions, it is easily checked that given a measure μ on X the collection of all μ-measurable sets form a σ-algebra.

Remark 2.7. *In probability and statistics, very often it is common to instead start with a σ-algebra M on a space X and then given a measure μ, to refer to all sets in M as being measurable.*

Theorem 2.1 (Existence of a smallest σ-algebra). *Let A be a (possibly uncountable) collection of subsets in X. Then there exists a "smallest" sigma algebra $\mathcal{A}$ that contains all of the sets in A.*

Proof. Consider the set of all σ-algebras that contain the subsets of A, indexed by some possibly uncountable set I, denoted $\{\mathcal{A}_\alpha\}_{\alpha \in I}$.

Then it is easily checked by definition that the intersection of all these σ-algebras (which is obviously the smallest), is itself a σ-algebra.

$$\mathcal{A} = \bigcap_{\alpha \in I} \mathcal{A}_\alpha$$

$\square$

Hence given any collection of subsets for a given space, we can form a σ-algebra. There is a very important special case in which we do this all the time:

Definition 2.5. *Let $(X, \mathcal{O})$ be a topological space. Then the smallest σ-algebra containing the sets in $\mathcal{O}$ (often called the open sets in X) is called the Borel σ-algebra.*

We will now attempt to convince ourselves that Carathéodory's definition of a measurable set is a good one in the sense that it leads to powerful measures with lots of measurable sets in a general setting.

Definition 2.6. *Let X be a topological space. A measure μ on X is Borel regular if:*

1. *All Borel sets are μ-measurable. (All of the sets in the Borel σ-algebra)*

2. *For every set $A \subseteq X$ there exists a borel set $A \subseteq B$ with $\mu(B) = \mu(A)$*

Remark 2.8. *It is a common mistake to assume that because $\mu(B) = \mu(A)$ that $\mu(B \setminus A) = 0$. But this is only true under certain circumstances. Suppose A is measurable, that implies $\mu(B) = \mu(B \setminus A) + \mu(B \cap A)$ by the definition of measurability. Since $A \subseteq B$ we have that $B \cap A = A$ and thus $\mu(B) = \mu(B \setminus A) + \mu(A)$, one may be tempted to conclude from this that $\mu(B \setminus A) = 0$, but because we allow measures to be infinite, we cannot simply subtract $\mu(A)$ from both sides lest we get $\infty - \infty$ which is left undefined. Thus we must assume that A is measurable and $\mu(A) < \infty$ to conclude $\mu(B \setminus A) = 0$.*

Theorem 2.2 (Carathéodory's Theorem). *Let X be a metric space and μ be any outer measure on X such that $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $d(A, B) > 0$. Then*

Remark 2.9. *If $d(a, b)$ is the distance between two points then we define the distance between two sets of points $d(A, B) = \inf\{d(a, b) : a \in A, b \in B\}$ then all Borel sets in the Borel σ-algebra generated by the open balls under d are μ-measurable.*

The assumption above is satisfied for many reasonable measures, intuitively because if the sets are separated by some distance then they shouldn't have any overlap - thus their measures should simply add.

Sometimes it is useful to measure the amount of one set that lies in another. For this we define a notion of restriction.

Definition 2.7. *Let μ be a measure on X and let $A \subseteq X$. The we define the restriction of μ to A by:*

$$\mu \llcorner A(B) = \mu(A \cap B)$$

Remark 2.10. *It's easy to see that the above defines a new measure, and that if a set B is μ measurable then it is also $\mu \llcorner A$ measurable.*

2.2 Integration, Generalized

We now turn to defining integration with respect to a measure μ. If you recall, the familiar Riemann integral from undergraduate calculus $\int f(x)dx$ is defined using sums that look like $\sum f(x)\Delta x$. Some may even have nightmares approximating such integrals over some interval (a, b) by hand by drawing boxes under curves at points $\{x_i\}_{i=1}^n \subseteq (a, b)$ of some fixed width Δx. It would be very natural then for us to define our integrals using sums with terms like:

$$f(x_i)\mu((x_i, x_{i+1}))$$

But it turns out that we can do better! In the sense that a slight alteration will allow us to integrate everything the above would lead to, and more. We will instead use sums that have terms in the form of:

$$y_i\mu(f^{-1}((y_i, y_{i+1})))$$

We still have a height times a length, but now we're measuring the preimage of a chunk of the y-axis (which itself is a chunk of the x-axis). This small dance turns out to be rather powerful, and with it emphasis is placed on functions that when we preimage open sets, give us sets we know to be measurable (i.e. we can do this trick!) we choose to call these functions **measurable functions**.

Definition 2.8. *Suppose μ is a measure on X. Let Y be a topological space. Let $f : X \to Y$ be a function from X into Y. Then f is called a $\mu-$measurable function if and only if for every open set $U \subseteq Y$ we have $f^{-1}(U) \subseteq X$ is a μ-measurable set.*

Definition 2.9. *Let μ be a measure on X. Then a function $g : X \to [-\infty, \infty]$ is called a μ-step function (if the context of the measure is clear, simply a "step function") if g is μ-measurable and the image of X under g (aka the range), $g(X) := \{g(x) : x \in X\}$ is countable as a set.*

Remark 2.11. *The reason why we care about step functions is because we know that sums $\sum$ require a countable index set to sum over.*

Definition 2.10. *A μ-step function is called μ-integrable if the sum:*

$$I(g) := \sum_{y \in g(X)} y\mu(g^{-1}(y))$$

Converges in $[-\infty, \infty]$.

Remark 2.12. *In the above definition we will agree that $0 \cdot \infty = 0$, and that $\infty - \infty$ is undefined.*

Definition 2.11. *A μ-integrable function is called μ-summable if $I(g) \in (-\infty, \infty)$.*

Now we can define integration properly.

Definition 2.12. *Let $f : X \to [-\infty, \infty]$ and let S_μ be the set of all μ-integrable step functions. Define the upper integral of f with respect to μ over X to be:*

$$\int^* f d\mu := \inf\{I(g) : g \in S_\mu, f \le g, \mu - a.e.\}$$

Similarly the lower integral:

$$\int_* f d\mu := \sup\{I(g) : g \in S_\mu, g \le f, \mu - a.e.\}$$

Definition 2.13. *Such a function f as above is μ-integrable if and only if:*

$$\int^* f d\mu = \int_* f d\mu$$

And we call this value the integral of f with respect to μ over X and denote it by $\int f d\mu$.

Remark 2.13. *Every non-negative μ-measurable function is μ-integrable.*

Theorem 2.3 (Egoroff)**.** *Let μ be an outer measure on X and let A be $\mu-$measurable with finite measure $\mu(A) < \infty$. If $f_1, f_2, \ldots$ is a sequence of $\mu-$measurable functions $f_k : A \to \mathbb{R}$ with $\lim\limits_{k \to \infty} f_k(x) = 0$ for all $x \in A$, then for every $\varepsilon > 0$ there exists $B \subseteq A$ with $\mu(A \setminus B) < \varepsilon$ and $f_k(x) \to 0$ uniformly on B.*

Remarkably, Egoroff's theorem says that if you have pointwise convergence, you can throw away a set of arbitrarily small measure to get uniform convergence. This is particularly useful for us, as in measure theory theorems often are stated as holding "almost everywhere" (a.e.) meaning up to a possible set of measure 0, where the statement may not hold.

Remark 2.14. *Regarding the theorem above:*

1. *When we say $\lim_{k \to \infty} f_k = 0$, this is merely to save space, it is equivalent to say that there exists a limiting function f for which $\lim_{k \to \infty} f_k = f$, we merely subtracted f from both sides and spoke about the new function $f_k - f$, thus eliminating the need to mention f at all. This is common shorthand and will not be mentioned again.*

2. *Uniform convergence is much stronger than pointwise convergence, in that the sequence of functions must approach the limiting function in such a way that the values convergence for all x in the domain simultaneously (in a tube-like fashion). Pointwise convergence merely means that looking at any particular point, the value of the sequence at that point converges to the limit.*

2.3   Hausdorff Measures

Hausdorff measures are some of the monst important measures we have in Geometric Measure Theory and provide not only a reasonable notion of length, area, and volumes but also dimension in a very general way. Their definition will be in two parts.

Definition 2.14. *Let X be a metric space. Define the diameter of a set A to be $diam(A) = \sup\{d(x,y) : x, y \in A\}$. For any set $S \subseteq X$ and any real numbers $k, \delta \geq 0$, define the k-dimensional measure of S by sets of diameter $\leq \delta$ by:*

$$\mathcal{H}^k_\delta(S) := \inf \left\{ \sum_j \frac{\omega_k \cdot diam(E_j)^k}{2^k} : S \subseteq \bigcup_j E_j, diam(E_j) \leq \delta \right\}$$

Where the infimum is taken over all countable covers $S \supseteq \bigcup_j E_j$ constrained so that each set E_j has a diameter at most δ. ω_k is the volume of the unit k-dimensional ball.

The value of our dimensional constant is $\omega_k := \frac{2^k \pi^{k/2}}{\Gamma(\frac{k}{2}+1)}$, where Γ denotes the gamma function. This constant is chosen so that our measure agrees with intuitive convention.

The above definition may seem quite complex at first, but it is quite simple and natural. Suppose that you are given some flat shape and you want to measure its area, but all you have are post-it notes of varying sizes. One thing

7

you can do is to try to cover the object with the post it notes and then add up the areas of each note. (Of course in the mathematical definition, your post-it notes can be of any shape but of size no more than δ) One thing you will notice if you do this, is that when you use smaller post it notes (make δ smaller) your approximation increases in accuracy (perhaps at the cost of the work of covering the object in more and more notes), and if you use small enough notes you will cover the object nearly perfectly. Hence:

Definition 2.15. *Let S be a subset of a metric space X. We define the k-dimensional Hausdorff measure of S by:*

$$\mathcal{H}^k(S) := \lim_{\delta \to 0} \mathcal{H}^k_\delta(S)$$

Example 2.3. *When $k = 0$, we have that the 0 dimensional Hausdorff measure of a set S, $\mathcal{H}^0(S)$ defined above is nothing more than the number of points in the set. If there is an infinite number of points, $\mathcal{H}^0(S) = \infty$.*

In a sense, δ is the sort of resolution at which we are looking at the set to measure. Imagine we wish to measure an infinite spiral S. We know intuitively that the spiral should have infinite length, so it should be that $\mathcal{H}^1(S) = \infty$; but imagine we try to measure it by covering it with sets of some diameter $\varepsilon > 0$; we would necessarily cover up too much of the tail and get a finite number for $\mathcal{H}^1(S)$. Hence the reason why we need to look at the limiting value as $\delta \to 0$.

Thus far we've talked about a notion of dimension k, and we know how to (at least implicitly) calculate the k-dimensional Hausdorff measure for a given set. But this gives rise to the question; which k should I use to measure a given set I'm interested in? Will the answer change if I choose "incorrectly"? This leads us to the derived notion of Hausdorff dimension.

Definition 2.16. *Let X be a metric space and $S \subseteq X$. Define:*

$$\mathcal{H}\text{-}dim(S) := \sup \left\{ k \geq 0 : \mathcal{H}^k(S) < \infty \right\}$$

Or equivalently

$$\mathcal{H}\text{-}dim(S) := \inf \left\{ k \geq 0 : \mathcal{H}^k(S) = \infty \right\}$$

The above is always defined for any set S. What this tells us is that there is that given a set S, there is a unique value of k that is a sort of Goldilocks number. Suppose we're trying to measure the area of a piece of paper. If we use a $k = 1$ dimensional measure, then we are trying to essentially cover the paper with 1 dimensional sets - i.e. infinitely thin lines. Since we will need an infinite number of these lines to cover the page, we'll get a 1-dimensional measure of ∞. Suppose that we try to measure the piece of paper using a $k = 3$ dimensional measure, covering it with something like tennis balls or cubes. No matter what we choose, the piece of paper will only occupy an infinitely thin slice of the 3d

object - hence its $k = 3$ measure will be 0. In this case, the Goldilocks value of k to use will be 2 - we can again use something like sticky notes to cover the piece of paper and figure out a reasonable idea of what its measure is.

More precisely, there is some value of k where $\mathcal{H}^k(S)$ flips from being ∞ to 0, measuring with sets of too small dimension to too big of dimension. This value of k where it flips is what we call the dimension of S.

Remark 2.15. *Note that at $k = \mathcal{H}\text{-}dim(S)$ the set S may still have measure of 0 or ∞. Imagine measuring the empty set, or perhaps all of $\mathbb{R}^2$ for examples.*

Example 2.4. *Show the cantor set has Hausdorff dimension $\frac{\log 2}{\log 3}$*

2.4 Rectifiable Sets

Consider $\mathbb{R}^n$ as a space equipped with the standard euclidean metric and topology inherited from it.

Definition 2.17. *A Borel subset S of $\mathbb{R}^n$ is called k-rectifiable if S has Hausdorff dimension k and there exists a countable collection of continuously differentiable maps:*

$$f_i : \mathbb{R}^k \to \mathbb{R}^n$$

such that:

$$\mathcal{H}^k \left(S \setminus \bigcup_{i=1}^{\infty} f_i(\mathbb{R}^k) \right) = 0$$

That is to say, a Borel subset in $\mathbb{R}^n$ is k-rectifiable if it is approximately pieces of $\mathbb{R}^k$ (approximately in the sense of up to a set of $\mathcal{H}^k$ measure 0), deformed in a continuously differentiable manner.

Remark 2.16. *Some authors prefer to equivalently define rectifiable sets using Lipschitz maps instead of continuously differentiable ones.*

Remark 2.17. *You can generalize the above notion to (μ, k)-rectifiability where μ is any measure you want, simply replace $\mathcal{H}^k$ with μ in the definition.*

Perhaps more intuitively, rectifiable sets may be thought of as generalizations of k-dimensional C^1-submanifolds. In the sense that they are countable unions of pieces of k-dimensional C^1-submanifolds (possibly with a measure 0 junk set) or of possessing approximate tangent spaces $\mathcal{H}^k$ almost everywhere.

We now seek to define these so-called approximate tangent spaces in terms of a procedure that "blows-up" or "zooms-in" at a point.

Definition 2.18. *Let $S \subseteq \mathbb{R}^n$. Then a subspace T of $\mathbb{R}^n$ is the approximate tangent space to S at the point $p \in \mathbb{R}^n$ if for all compactly supported $f \in C_c(\mathbb{R}^n)$:*

$$\lim_{\lambda \to 0^+} \int_{\eta_{p,\lambda}(S)} f(y) d\mathcal{H}^k(y) = \int_T f(y) d\mathcal{H}^k(y)$$

Where $\eta_{p,\lambda}(y) := \frac{y-p}{\lambda}$.

The above may look daunting, it's actually quite simple and very informative of what is going on. The map $\eta_{p,\lambda}(y)$ for $0 < \lambda < 1$ acts as a blow up by a factor of λ for the set S about the point p. The integrals against test functions is nothing more than the standard notion of weak convergence. Hence we say that T is the tangent plane of S at p if we have weak convergence of the identity 1 over the blow up of S at p to the identity over T as $\lambda \to 0^+$. Even more simply, when we zoom in on the point p, our set S approximately looks like T.

2.5   Forms

A $1-$form is a linear function that assigns a vector to a number. Hence, this "action" on a vector is its defining property. Two 1-forms are considered equal then if and only if they are equal as functions; they act on all vectors in precisely the same way. We define the sum of two 1-forms α and β by:

$$(\alpha + \beta)(v) = \alpha(v) + \beta(v)$$

Just the same as we typically define function addition. Which, because both α, β are linear their sum is still linear and hence the sum of two 1-forms is a 1-form. Likewise, we define scalar multiplication in the obvious way. Thus the set of all 1-forms over some common domain is itself a vector space.

If V is a vector space, then the set of all 1-forms is denoted by V^* and is called the dual vector space of V. The 1-forms are therefore sometimes referred to as covectors.

Example 2.5. *The gradient of a scalar valued function $f : \mathbb{R}^n \to \mathbb{R}$ denoted by $\nabla f = \begin{bmatrix} f_{x_1} & f_{x_2} & \cdots & f_{x_n} \end{bmatrix}$ is a covector whose action (defined by matrix multiplication) with a vector u in $\mathbb{R}^n$ indicates the slope of f in the direction of u.*

Example 2.6. *Any row vector with n entries (above which is a special case) can be viewed as a covector or 1-form with domain $\mathbb{R}^n$.*

Given a basisfor a vector space V there is a natural basis for V^* called the cannonical dual basis. If $v \in V$ is given by $v = (v_1, v_2, ..., v_n)$ then we define the ith basis vector for the dual space ω_i by its action on v given by $\omega_i(v) = v_i$. In other words, the ith component of v which is a real number. Hence ω_i is a real-valued function and it is easy to check its linearity hence it is in fact a

1-form. Clearly then if we use the standard basis for V given by e_i then we have

$$\omega_i(e_j) = \delta_i^j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

Clearly then if we have a vector with components $v_1, v_2, ..., v_n$ with respect to the standard basis we can write either $v = v_1 e_1 + \cdots + v_n e_n$ or $v = \omega_1(v)e_1 + \cdots + \omega_n(v)e_n$.

2.6 Currents

2.7 The Flat Norm

3 L1TV Computes the Flat Norm

4 Answers and Questions

5 Computing the Flat Norm