1 Convergence of Expansions

Theorem 1.1. *Suppose either*

1. *X is a metric space of finite diameter, and for any X_d and all $x, y \in X_d$; there is a uniform constant ϵ such that*
$$\frac{d(Tx, Ty)}{d(x, y)} \geq 1 + \epsilon$$
2. *$X \subseteq \mathbb{R}$, a bounded subset, each X_d is an interval, T is continuous on X_d, differentiable on interior of X_d and there exists $\epsilon > 0$ such that $|T'x| \geq 1 + \epsilon$ for all x where the derivative exists.*

Then the expansion converges for all points in X.

Example 1.1. Hurwitz complex CFs

$$X = \{x + iy : x, y \in [-1/2, 1/2)\},$$

$$Tz = \left\{\frac{1}{z}\right\}$$

Then
$$d\left(\frac{1}{z}, \frac{1}{z'}\right) = \frac{d(z, z')}{|z||z'|}$$

Proof. Note case 1 implies case 2 by the mean value theorem.

For case 1, consider a cylinder set
$$C_{[a_1]} = X_{a_1},$$
we know that for all $x, y \in C_{[a_1]}$,
$$\frac{d(Tx, Ty)}{d(x, y)} \geq 1 + \epsilon.$$

Consider $C_{[a_1, a_2]}$. Then if $x \in C_{[a_1, a_2]}$, $Tx \in C_{[a_2]}$.

So for all $x, y \in C_{[a_1, a_2]}$, then
$$\begin{aligned}\frac{d(T^2x, T^2y)}{d(x, y)} &= \frac{d(T^2x, T^2y)}{d(Tx, Ty)} \cdot \frac{d(Tx, Ty)}{d(x, y)} \\ &\geq (1 + \epsilon)^2.\end{aligned}$$

By induction, for all $x, y \in C_{[a_1, \ldots, a_k]}$,
$$\frac{d(T^kx, T^ky)}{d(x, y)} \geq (1 + \epsilon)^k.$$

Since our metric space has finite diameter, $d(\cdot, \cdot) \leq M$ for all inputs (for some M). Thus for all $x, y \in C_{[a_1, \ldots, a_k]}$,
$$d(x, y) \leq \frac{M}{(1 + \epsilon)^k}$$

So as $k \to \infty$,
$$\operatorname{diam}(C_{[a_1, \ldots, a_k]}) \to 0.$$

This completes the proof. ∎

Note that this doesn't work for regular continued fractions, since we have $Tx = \left\{\frac{1}{x}\right\}$, which has derivative -1 at 1.

How do we know the regular continued fractions do converge? There's only a problem at $x = 1$, and near 1 we end up near 0. We can thus do two steps, since things will expand after a possibly bad step.

Question: Where does a cylinder set get its name? Not sure. There's lots of weird terminology.

2 Measures

Bergelson: The more you talk about it, the less I need to!
Definition 2.1. A *measure* is a way to measure the size of sets. (It will not usually measure all sets).

A measure $\mu : \mathcal{B} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$, $\mathcal{B} \subseteq \mathcal{P}(X)$, satisfies

1. $\mu(A) \geq 0$ for all $A \in \mathcal{B}$.
2. $\mu(\varnothing) = 0$.
3. If $A_1, A_2, \ldots, A_n, \ldots$ are all disjoint

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

Why the $\mathcal{B}$? Well, it's not possible to always consistently assign measures to sets when we take all of the subsets, so we restrict ourselves to certain sets that we care about. In particular, $\mathcal{B}$, is a σ*-algebra.* A σ-algebra, $\mathcal{A}$, on a set X must satisfy

1. $\mathcal{A} \subseteq \mathcal{P}(X)$,
2. $X \in \mathcal{A}$,
3. if $A \in \mathcal{A}$, then $X \setminus A \in \mathcal{A}$,
4. and if $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{A}$, then

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}.$$

Measures are always defined on families of sets that form σ-algebras.

Let $\mathscr{A}$ be any collection of subsets of X. Then there is a smallest σ-algebra containing $\mathscr{A}$, $\sigma(\mathscr{A})$. Note. If $\mathcal{F}_1, \mathcal{F}_2$ are σ-algebras, then $\mathcal{F}_1 \cap \mathcal{F}_2$ is also a σ-algebra. This is also true for arbitrary intersections of σ-algebras.

Thus we have that

$$\sigma(\mathscr{A}) = \bigcap_{\mathcal{A} \supseteq \mathscr{A}} \mathcal{A},$$

where the $\mathcal{A}$ are all σ-algebras.
Example 2.1. "Obvious choice of measure on $[0, 1]$" (This will also work for $\mathbb{R}$, $\mathbb{R}^n$.)

For intervals, we should have $\lambda([a, b]) := b - a$.

If we let $\mathscr{A}$ be the collection of all subintervals of $[0, 1]$, $\sigma(\mathscr{A})$ is called the *Borel* σ- algebra on $[0, 1]$, $\mathcal{B}([0, 1])$. We will see λ extends to $\mathcal{B}([0, 1])$. Since λ is defined on $\mathcal{B}([0, 1])$, we say that it is a Borel measure.

3 Semi-algebras

Definition 3.1. A collection $\mathscr{A}$ of subsets of X is a semi-algebra if

1. $X, \varnothing \in \mathscr{A}$,
2. If $A, B \in \mathscr{A}$, then $A \cap B \in \mathscr{A}$.
3. If $A, B \in \mathscr{A}$, then $A \setminus B$ must be a union of a finite number of disjoint sets in $\mathscr{A}$.

Note that (3) implies $\varnothing \in \mathscr{A}$, since $X \setminus X = \varnothing$.

If $\mathscr{A}$ is a semi-algebra, and μ is a "measure" on $\mathscr{A}$, then the measure μ extends to a unique measure μ^* on $\sigma(\mathscr{A})$. (This is Caratheodory's Extension Theorem).

Example 3.1. If $\mathscr{D}$ is finite, then the collection of cylinder sets (and X and $\varnothing$) is a semi-algebra.

Let's verify the rules.

1. By definition, this is satisfied.
2. We'll ignore X and $\varnothing$. Consider $C_s, C_t \in \mathscr{D}$. If $s = [a_1, a_2, \dots, a_k]$, $t = [b_1, b_2, \dots, b_\ell]$, $k \leq \ell$. If $a_i \neq b_i$ for some $1 \leq i \leq k$, then $C_s \cap C_t = \varnothing$. Otherwise $a_i = b_i$ for $1 \leq i \leq k$, so $C_s \cap C_t = C_t$.
3. If $C_s \cap C_t = \varnothing$, then $C_s \setminus C_t = C_s$. Otherwise, if $C_s \cap C_t = C_t$, then $C_t \setminus C_s = \varnothing$, so we may assume we are considering $C_s \setminus C_t$, which is the set of all $x \in X$ whose first k digits are s, but whose first ℓ digits are not t.

   Since $\mathscr{D}$ is finite, there are a finite number of strings of length ℓ whose first k digits are s, but are not equal to t. Then $C_s \setminus C_t$ will be the disjoint union of the finitely many corresponding cylinder sets.

What if our digit set $\mathscr{D}$ is infinite? We can use "extended cylinder sets,"

$$C^*_{[a_1,\dots,a_k]} = \bigcup_{d \geq a_k} C_{[a_1,\dots,a_{k-1},d]}.$$

This is a sort of compactification argument.

This is how we get a nice semi-algebra for continued fractions for example.

3.1 Completing a σ-algebra and measure

Those familiar with measure theory will know that we don't usually work with the Borel σ-algebra or Borel measures. We need to complete the measure.

We have a set X, a σ-algebra $\mathcal{B}$, and a measure μ. We want to add all subsets of zero-measure sets to our σ-algebra and fill everything in accordingly.

When you complete the Borel σ-algebra and measure, you get the Lebesgue σ-algebra and measure, which we will also denote by λ.

3.2 Terminology

Definition 3.2. A function $f : X \to Y$ is *measurable* with respect to σ-algebras $\mathcal{B}$ on Y and $\mathcal{A}$ on X if for all $B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$.

Recall $f^{-1}(A) = \{x \in X : f(x) \in A\}$.

Definition 3.3. A set, A, has *zero* or *null measure* if $\mu(A) = 0$. A has *full measure* if $\mu(A) = \mu(X)$ (assuming $\mu(X) < \infty$).

Definition 3.4. Something happens for *almost all points in A*, or *almost everywhere in A*, if the set of points $x \in A$ where the thing doesn't happen has measure zero.

For example, $\lambda(\mathbb{Q} \cap [0,1]) = 0$, so almost every real number is irrational.

4 Integration of functions with respect to measures

We know about integration of functions on intervals in $\mathbb{R}$:

$$\int_a^b f(x)\,dx \text{ "signed area under curve"}$$

What then is

$$\int_A f\,d\mu?$$

First, if

$$f = 1_A = \begin{cases} 1 & x \in A \\ 0 & x \notin A, \end{cases}$$

then

$$\int_X 1_A\,d\mu = \mu(A).$$

Second, if $f = \sum_{i=1}^k c_i 1_{A_i}$, where the A_i are disjoint, we have

$$\int_X f\,d\mu = \sum_{i=1}^k c_i \int_X 1_{A_i}\,d\mu = \sum_{i=1}^k c_i \mu(A_i).$$

Third, what if f can be expressed as

$$f_k \to f,$$

where f_k are as in the previous step?

Then

$$\int_X f\,d\mu = \lim_{k\to\infty} \int_X f_k\,d\mu.$$

Finally,

$$\int_A f\,d\mu = \int_X f \cdot 1_A\,d\mu.$$

Remark. Most of the time,

$$\int_X f\,d\lambda = \int f\,dx.$$

Some functions can only be integrated in the world of measure, for example $1_{\mathbb{Q}\cap[0,1]}$. This is because the Riemann integral won't converge.

If $\mu(A) = 0$, whenever $\lambda(A) = 0$, then there exists a function h, such that

$$\int_X f\,d\mu = \int_X f \cdot h\,d\lambda.$$

h is the Radon-Nikodym derivative of μ with respect to λ.

For example, there is a nice measure, the Gauss measure, μ, and we get

$$d\mu = \frac{1}{(\log 2)(1+x)} d\lambda$$