DYNAMICAL GEOMETRIC MEASURE THEORY

JULIE ROWLETT, MALTE HARTMANN, KATIE BUCHHORN, KARSTEN BOHLEN, PATRICK RODE, JIL KLÜNDER, PHILIP BLEEKER, YASEMIN ACAR, CEYHUN ÖZDEMIR, AND MARY COSGROVE ROBERTS

[1]

1. What is volume?

It turns out that volume is not actually a singularly defined notion. It depends upon the definition of a measure. Usually when someone talks about volume what they mathematically mean is three dimensional Lebesgue measure. We will see that there are many notions of volume, corresponding to many different measures...

Definition 1.1. Let X be a set. A subset $\mathcal{A} \subset P(X)$ is called an *algebra* if

(1) $X \in \mathcal{A}$
(2) $Y \in \mathcal{A} \implies X \setminus Y =: Y^c \in \mathcal{A}$
(3) $A,\, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$

$\mathcal{A}$ is a σ-algebra if in additon

$$\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}.$$

Remark 1. Note that algebras are always closed under intersections, since for $A, B \in \mathcal{A}$,

$$A \cap B = (A^c \cup B^c)^c \in \mathcal{A},$$

since algebras are closed under complements and unions. Consequently, σ-algebras are closed under countable intersections.

We will often use the symbol σ in describing countably-infinite properties. *Think about a few examples of algebras.*
Example Let X be a topological space. An important example of a σ-algebra is the Borel σ-algebra, which is the smallest which contains all open sets. Prove that this satisfies the above axioms.

Definition 1.2. Let X be a set and $\mathcal{A} \subset P(X)$ a σ-algebra. We will call $(X, \mathcal{A})$ a measure space. We may be a bit laid-back about this and also use measure space to refer to a set, a σ-algebra, as well as a *measure.* A *measure* μ is a countably additive, monotone set function which is defined on our σ-algebra. It must vanish on the empty set. We will only work with non-negative measures, but there is such a thing as a signed measure. Just so you know those beasties are out there.

(1) Monotone means that if $A \subset B$ then $\mu(A) \le \mu(B)$.
(2) Countably additive means that for a countable disjoint collection of sets in the σ-algebra

$$\{A_n\} \subset \mathcal{A} \text{ such that } A_n \cap A_m = \emptyset \forall n \neq m \implies \mu\left(\bigcup A_n\right) = \sum \mu(A_n).$$

Remark 2. Note that we can always disjoin sets. So, if $\{A_n\} \subset \mathcal{A}$ is a countable collection of sets, setting

$$B_1 := A_1, \quad B_n := A_n \setminus \cup_{k=1}^{n-1} A_k, k \ge 2,$$

[1]We are grateful to Mary Cosgrove Roberts for photographing and posting a koala picture to Facebook which we have included to indicate the end of a proof.

we have

$$\cup B_n = \cup A_n, \quad B_n \subset A_n \forall n \implies \mu(B_n) \le \mu(A_n)$$

and by countable additivity and since $\cup B_n = \cup A_n$

$$\mu(\cup A_n) = \mu(\cup B_n) = \sum \mu(B_n) \le \sum \mu(A_n).$$

So, for not-necessarily disjoint sets, we have countable subadditivity.

Example Cook up some examples of σ-algebras and measures on them. An easy one is to take $X = \mathbb{N}$ and the algebra $\mathcal{A} = P(X)$. Note that all elements of $\mathcal{A}$ are either countably infinite, finite, or empty. Define the measure to be 1 on a single element of $\mathbb{N}$ and 0 on the emptyset. Prove that this satisfies the definition of a measure space. It was suggested that we could take the analogously defined point measure on $\mathbb{R}$ and let $\mathcal{A} = P(\mathbb{R})$. Will this work?

Definition 1.3. A measure space is called σ-finite if there exists a collection of sets in the σ-algebra which cover the whole space, each of which has finite *mass.*What is *mass?* That is the notion of "volume" induced by the associated measure. So, the mass of an element of the σ-algebra is simply the value of the measure evaluated on that element. Mathematically for $A \in \mathcal{A}$ the mass of A is $\mu(A)$ where μ is our measure defined on elements of $\mathcal{A}$. We call the elements of $\mathcal{A}$ *measurable* sets. *Why is the whole space always a measurable set? Why is the emptyset always a measurable set?* Now, the whole space need not have finite mass, but if it does, then it's said to have finite mass, in which case one can normalize the measure so that the whole space has mass equal to one. Such a space is called a *probability space,* and the elements of $\mathcal{A}$ are called *events.* The interpretation of the mass of an event is the probability that it's gonna happen.

1.1. **Lebesgue Volume.** The n-dimensional Lebesgue measure is the unique, complete measure which agrees with our intuitive notion of n-dimensional volume. To make this precise, first we define a generalized interval and our notion of intuitive volume.

Definition 1.4. A generalized interval in $\mathbb{R}^n$ is a set for which there exist real numbers $a_k \le b_k$ for $k = 1, \dots n$, such that this set has the form

$$I = \{x \in \mathbb{R}^n, x = \sum x_k e_k, \quad a_k < \text{ or } \le x_k < \text{ or } \le b_k, k = 1, \dots, n\}.$$

Above we are using e_k to denote the standard unit vectors for $\mathbb{R}^n$. The intuitive volume function on $\mathbb{R}^n$ is defined on such a set to be

$$v_n(I) = \prod (b_k - a_k).$$

Next we can extend our intuitive notion of volume to *elementary sets.*

Definition 1.5. An *elementary subset* of $\mathbb{R}^n$ is a set which can be expressed as a finite disjoint union of generalized intervals. The collection of all of these is denoted by $\mathcal{E}_n$.

Exercise: Prove that v_n is well-defined on $\mathcal{E}_n$.
What we shall call Lebesgue's Theorem (note that this is not his only awesome theorem, and his original statement may have been somewhat different) is the following.

Theorem 1.6 (Lebesgue)**.** *There exists a unique complete measure on $\mathbb{R}^n$ which agrees with v_n on $\mathcal{E}_n$ and such that the corresponding σ algebra is the smallest which contains $\mathcal{E}_n$.*

To prove this we will require techniques from another great French mathematician, Carathéodory. One unfortunate fact about measures is that they're not defined on arbitrary sets, only on measurable sets (remember, those are the ones in the associated σ algebra). However, there is a way to define a set function which is almost like a measure and is defined for *every imaginable or unimaginable set.* This thing is called an *outer measure.*

Definition 1.7. Let X be a set. An outer measure μ^* on X is a map from $P(X) \to [0, \infty]$ such that

$$\mu^*(\emptyset) = 0, \quad A \subset B \implies \mu^*(A) \le \mu^*(B),$$

and

$$\mu^*(\cup A_n) \leq \sum \mu^*(A_n).$$

Whenever things are indexed with n or some other letter and are not obviously indicated to be uncountable or finite, we implicitly are referring to a set indexed by the natural numbers (which unlike in France start with 1, not 0).
Now we can finally prove something.

2. Carathéodory's Extension of Outer Measures

Proposition 2.1 (Outer Mass Existence)**.** *Let $E \subset P(X)$ such that $\emptyset \in E$. Let ρ be a map from elements of E to $[0, \infty]$ such that $\rho(\emptyset) = 0$. Then we can define for* every element $A \in P(X)$

$$\rho^*(A) := \inf\{\sum \rho(E_j) : E_j \in E, A \in \cup E_j\},$$

where we assume that $\inf\{\emptyset\} = \infty$, so that if it is impossible to cover a set A by elements of E then $\rho^*(A) := \infty$. So defined, ρ^* is an outer measure.

Proof: Note that ρ^* is defined for every set. Now since $\emptyset \subset \emptyset = \cup E_j$ for all $E_j = \emptyset \in E$ we have that since $\rho \geq 0$

$$0 \leq \rho^*(\emptyset) \leq 0 \implies \rho^*(\emptyset) = 0.$$

This is the first condition an outer measure must satisfy. Next, let's assume $A \subset B$. (By $\subset$ we always mean $\subseteq$). Then, since any covering of B by elements of E is also a covering of A by elements of E, it follows that the infimum over coverings of A is an infimum over a potentially *larger set of objects (namely coverings)* as compared with the infimum over coverings of B. Hence we have

$$\rho^*(A) = \inf\{\sum \rho(E_j) : E_j \in E, A \in \cup E_j\} \leq \inf\{\sum \rho(E_j) : E_j \in E, B \in \cup E_j\} = \rho^*(B).$$

This is the second condition. Finally, we get to do some analysis here. Let $\epsilon > 0$ be arbitrary. Since the definition of ρ^* is by means of an infimum, if we have a countable collection of sets

$$\{A_j\} \subset X,$$

then for each $j \in \mathbb{N}$ there exists a countable collection of sets $\{E_j^k\}_{k=1}^\infty$ where each $E_j^k \in E$, such that

$$\rho^*(A_j) \geq \sum_{k\geq 1} \rho(E_j^k) - \frac{\epsilon}{2^j} \implies \rho^*(A_j) + \frac{\epsilon}{2^j} \geq \sum_{k\geq 1} \rho(E_j^k).$$

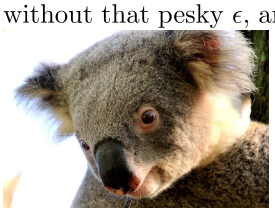
Well then, the collection $\{E_j^k\}$ is a countable collections of elements of E which covers

$$\cup A_j.$$

Therefore by the definition of ρ^* we have

$$\rho^*(\cup A_j) \leq \sum_{j,k\geq 1} \rho(E_j^k) \leq \sum_{j\geq 1} \rho^*(A_j) + \frac{\epsilon}{2^j} = \epsilon + \sum_{j\geq 1} \rho^*(A_j).$$

Since this inequality holds for arbitrary $\epsilon > 0$, we may let $\epsilon \to 0$, and the inequality also holds without that pesky ϵ, and this is precisely the third requirement for ρ^* to be an outer measure.



Now we can say what we mean for a measure to be *complete.*

Definition 2.2. Let $(X, \mathcal{A}, \mu)$ be a measure space. Then, there is a canonically associated outer measure induced by μ defined by

$$\mu^*(A) := \inf\{\sum \mu(E_j), \quad \{E_j\} \subset \mathcal{A},\ A \subset \cup E_j\}.$$

We say that the measure μ is *complete* if, for each $A \subset X$ such that

$$\mu^*(A) = 0 \implies A \in \mathcal{A}.$$

2.1. **Exercise Set 1.**

(1) Prove that v_n is well defined on $\mathcal{E}_n$ for all $n \in \mathbb{N}$. Prove that $\mathcal{E}_n$ is *not* an algebra. Give a construction of the smallest algebra which contains $\mathcal{E}_n$.
(2) Given a measure space $(X, \mathcal{A}, \mu)$ and $E \in \mathcal{A}$, define
$$\mu_E(A) = \mu(A \cap E)$$
for $A \in \mathcal{A}$. Prove that μ_E is a measure.
(3) Prove that the intersection of arbitrarily many σ-algebras is again a σ-algebra. Does the same hold for unions?
(4) Let A be an infinite σ-algebra. Prove that A contains uncountably many elements.

Theorem 2.3 (Carathéodory)**.** *Let μ^* be an outer measure on X. A set $A \subset X$ is called measurable with respect to μ^* $\Leftrightarrow \forall\ E \subset X$ the following equation holds:*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c). \qquad (*)$$

Then $\mathcal{M} := \{A \subset X | A$ is μ^* measurable$\}$ is a σ-algebra and $\mu^*\big|_{\mathcal{M}}$ is a complete measure.

Proof: Note that $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$ because (*) is symmetric in A and A^c. $\emptyset \in \mathcal{M}$ since $\mu^*(\emptyset) = 0$.
Next we will show that $\mathcal{M}$ is complete under finite unions of sets:
For $A, B \in \mathcal{M}$ and $E \subset X$ we get, by multiple use of (*):

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*((E \cap A) \cap B) + \mu^*((E \cap A) \cap B^c) + \mu^*((E \cap A^c) \cap B) + \mu^*((E \cap A^c) \cap B^c).$$

Furthermore, we can write $A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$, which gives us

$$\mu^*(E \cap A \cap B) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A \cap B^c) \geq \mu^*(E \cap (A \cup B))$$

Using this inequality in the above equation gives us:

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$

This inequality is actually an equality, as "$\leq$" follows from the outer measure axioms. Hence $A \cup B \in \mathcal{M}$.
μ^* is finitely-additive: $\forall\, A, B \in \mathcal{M},\ A \cap B = \emptyset \Rightarrow \mu^*(A \cup B) = \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) = \mu^*(A) + \mu^*(B)$

Now we will show that $\mathcal{M}$ is actually a σ-algebra: For $\{A_j\}_{j \in \mathbb{N}} \subset \mathcal{M}$ we can define a sequence of disjoint sets $\{B_j\}_{j \in \mathbb{N}} \subset \mathcal{M}$ fulfilling $\bigcup_{j \in \mathbb{N}} A_j = \bigcup_{j \in \mathbb{N}} B_j$ by:

$$B_1 := A_1, \quad B_2 := A_2 \setminus A_1, \quad B_3 := A_3 \setminus (A_1 \cup A_2) \ ...$$

Let us also define $\tilde{B}_n := \bigcup_{j=1}^n B_j$.
So, we need to show that $\bigcup_{j \in \mathbb{N}} B_j \in \mathcal{M}$. For $E \subset X$:

$$\mu^*(E \cap \tilde{B}_n) \overset{(*)}{=} \mu^*(E \cap \tilde{B}_n \cap B_n) + \mu^*(E \cap \tilde{B}_n \cap {B_n}^c) = \mu^*(E \cap B_n) + \mu^*(E \cap \tilde{B}_{n-1})$$

Using $\mu^*(E \cap \tilde{B}_n) = \mu^*(E \cap B_n) + \mu^*(E \cap \tilde{B}_{n-1})$ inductively we get:

$$\mu^*(E \cap \tilde{B}_n) = \mu^*(E \cap B_n) + \mu^*(E \cap B_{n-1}) + \mu^*(E \cap \tilde{B}_{n-2}) = ... = \sum_{k=1}^n \mu^*(E \cap B_k)$$

Using this result we get:

$$\mu^*(E) = \mu^*(E \cap \tilde{B}_n) + \mu^*(E \cap \tilde{B}_n^c) = \sum_{k=1}^{n} \mu^*(E \cap B_k) + \mu^*(E \cap \tilde{B}_n^c)$$
$$\geq \sum_{k=1}^{n} \mu^*(E \cap B_k) + \mu^*(E \setminus (\bigcup_{k=1}^{\infty} B_k)) \qquad (**)$$

This inequality holds for any $n \in \mathbb{N}$. Taking the limit and using the outer measure axioms gives us

$$\mu^*(E) \geq \sum_{k=1}^{\infty} \mu^*(E \cap B_k) + \mu^*(E \setminus (\bigcup_{k=1}^{\infty} B_k))$$
$$\geq \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)) + \mu^*(E \setminus (\bigcup_{k=1}^{\infty} B_k))$$

Since $\mu^*(E) \leq \mu^*(E \cap Y) + \mu^*(E \setminus Y)$ holds for any $Y \subset X$, the above inequality is an equality:

$$\mu^*(E) = \mu^*(E \cap (\bigcup_{k=1}^{\infty} B_k)) + \mu^*(E \setminus (\bigcup_{k=1}^{\infty} B_k))$$

This shows that $\bigcup_{k=1}^{\infty} B_k \in \mathcal{M}$. Hence $\mathcal{M}$ is a σ-algebra.

Now we want to show that $\mu^*\big|_{\mathcal{M}}$ is countably additive. Let $\{B_k\}_{k\in\mathbb{N}} \subset \mathcal{M}$ be pairwise disjoint sets. Defining $E := \bigcup_{k=1}^{\infty} B_k$ and using (**), we get

$$\mu^*(\bigcup_{k=1}^{\infty} B_k) = \mu^*(E) \overset{(**)}{\geq} \sum_{k=1}^{\infty} \mu^*(E \cap B_k) + \mu^*(\emptyset) = \sum_{k=1}^{\infty} \mu^*(B_k) \geq \mu^*(\bigcup_{k=1}^{\infty} B_k)$$

$$\mu^*(\bigcup_{k=1}^{\infty} B_k) = \sum_{k=1}^{\infty} \mu^*(B_k)$$

So $\mu^*\big|_{\mathcal{M}}$ is a measure. It is even a complete measure: For $Y \subset X$ such that $\mu^*(Y) = 0$ and for arbitrary $E \subset X$ we have

$$\mu^*(E) \leq \mu^*(E \cap Y) + \mu^*(E \cap Y^c) \leq \mu^*(Y) + \mu^*(E) = \mu^*(E)$$

Therefore $Y \in \mathcal{M}$.

Homework: We have ${v_n}^*$ defined von $\mathbb{R}^n$. Carathéodory's theorem shows that we get a complete measure and σ-algebra from it. Is this the Lebesgue measure $\mathcal{L}_n$ and $\mathcal{M}_n$? Prove your answer.

To prove the Lebesgue theorem in an original way..... we will show that

(1) ${v_n}^*$ on ${\mathcal{E}_n}^*$ is a "pre-measure" which is σ-finite. ${\mathcal{E}_n}^*$ is the smallest algebra containing $\mathcal{E}_n$.
(2) Prove another extension theorem which will show that there is a unique extension of ${v_n}^*$ to the smallest σ-algebra containing $\mathcal{E}_n$ which is a measure.
(3) Prove completeness.

3. Completeness of a Measure

The following gives two equivalent definitions for completeness. This is why I couldn't decide which one to use.

Proposition 3.1 (Completeness Proposition)**.** *The following are equivalent for a measure space $(X, \mathcal{M}, \mu)$. If either of these hold, then μ is called* complete.

(1) If there exists $N \in \mathcal{M}$ with $\mu(N) = 0$, and $Y \subset N$ then $Y \in \mathcal{M}$.
(2) If $\mu^*(Y) = 0$ then $Y \in \mathcal{M}$.

Proof:
So you see why I couldn't decide which definition of completeness was correct: they both are! First let us assume (1) holds. Then if $Y \subset X$ with $\mu^*(Y) = 0$, by the definition of μ^* for each $k \in \mathbb{N}$ there exists

$$\{E_n^k\}_{n\geq 1} \subset \mathcal{M}, \quad Y \subset \cup_n E_n^k, \quad \sum_n \mu(E_n^k) < 2^{-k}.$$

Well, then

$$Y \subset N := \cap_k \cup_n E_n^k \in \mathcal{M},$$

and since $N \subset \cup_n E_n^k$ for each $k \in \mathbb{N}$, by monotonicity of the measure

$$\mu(N) \leq \mu(\cup_n E_n^k) < 2^{-k} \forall k \in \mathbb{N} \implies \mu(N) = 0.$$

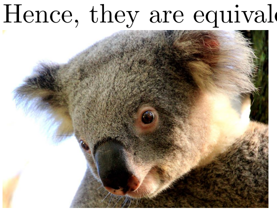
By the assumption of (1) since $Y \subset N \in \mathcal{M}$ and $\mu(N) = 0$, it follows that $Y \in \mathcal{M}$. So, every set with outer measure zero is measurable (that's what (2) says!)
Next, we assume (2) holds. Then if there exists $N \in \mathcal{M}$ with $\mu(N) = 0$ and $Y \subset N$, then

$$Y \subset \cup A_j, \quad A_1 := N, \quad A_j = \emptyset \forall j \geq 2,$$

and $\{A_j\} \subset \mathcal{M}$. So, by definition of outer measure,

$$0 \leq \mu^*(Y) = \inf \ldots \leq \sum \mu(A_j) = \mu(N) = 0.$$

Consequently $\mu^*(Y) = 0$, and by the assumption (2), $Y \in \mathcal{M}$. This shows that (2) $\implies$ (1). Hence, they are equivalent.



Theorem 3.2 (Completion of a measure)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $\mathcal{N} := \{N \in \mathcal{M} \mid \mu(N) = 0\}$ and*

$$\bar{\mathcal{M}} = \{E \cup F \mid E \in \mathcal{M} \text{ and } F \subset N \text{ for some } N \in \mathcal{N}\}.$$

Then $\bar{\mathcal{M}}$ is a σ-algebra and $\exists!$ extension $\bar{\mu}$ of μ to a complete measure on $\bar{\mathcal{M}}$.

Proof:
First, note that every element of $\mathcal{M}$ can be written as itself union with $\emptyset$, and $\emptyset \subset \emptyset \in \mathcal{N}$, so it follows that every element of $\mathcal{M}$ is an element of $\bar{\mathcal{M}}$. Next, if $\{A_n\} \subset \bar{\mathcal{M}}$ and $\{E_n, N_n\} \subset \mathcal{M}$ such that

$$A_n = E_n \cup F_n, \quad F_n \subset N_n \in \mathcal{N}.$$

Then

$$N := \cup N_n \in \mathcal{M}, \quad \text{and} \quad \mu(\cup N_n) \leq \sum \mu(N_n) = 0.$$

Similarly, $E := \cup E_n \in \mathcal{M}$ (why?), and we also have $F := \cup F_n \subset N$. It follows that

$$\cup A_n = E \cup F \in \bar{\mathcal{M}}.$$

Consequently $\bar{\mathcal{M}}$ is closed under countable unions. What about complements? If $A = E \cup F \in \bar{\mathcal{M}}$ with $F \subset N \in \mathcal{N}$ then note that

$$(E \cup F)^c = E^c \cap F^c = ((E^c \cap N) \cup (E^c \cap N^c)) \cap F^c,$$

and since $F \subset N \implies F^c \supset N^c$, the intersection of the last two terms is just $E^c \cap N^c$, so

$$(E \cup F)^c = (E^c \cap N \cap F^c) \cup (E^c \cap N^c).$$

Since $E, N \in \mathcal{M} \implies E^c \cap N^c \in \mathcal{M}$, and $E^c \cap N \cap F^c \subset N \in \mathcal{N}$ we see that $(E \cup F)^c \in \bar{\mathcal{M}}$. So, $\bar{\mathcal{M}}$ is closed under complements. Hence, we have shown that $\bar{\mathcal{M}}$ is a σ-algebra which contains $\mathcal{M}$.

Next, we must demonstrate that $\bar{\mu}$ is a well-defined, complete, and unique extension of μ. It is natural to ignore the subset of the zero-measure set, so we define

$$\bar{\mu}(E \cup F) := \mu(E).$$

If we have another representation of $E \cup F = G \cup H$ with $G \in \mathcal{M}$ and $F, H \subset N, M \in \mathcal{N}$, respectively, then

$$\bar{\mu}(E \cup F) = \mu(E) \leq \mu(G \cup M) \leq \mu(G) = \bar{\mu}(G \cup H) \leq \mu(E \cup N) \leq \mu(E) = \bar{\mu}(E).$$

Hence the whole line is an equality, and $\bar{\mu}$ is well-defined.

Now, let's show that $\bar{\mu}$ is really a measure. By definition, for $E \subset \mathcal{M}$

$$\bar{\mu}(E) = \mu(E), \implies \bar{\mu}(\emptyset) = 0.$$

If $\{A_n\} = \{E_n \cup F_n\} \subset \bar{\mathcal{M}}$ are disjoint, then

$$A_n \cap A_m \supset E_n \cap E_m \implies E_n \cap E_m = \emptyset, \quad \forall n \neq m.$$

Consequently,

$$\bar{\mu}(\cup A_n) = \mu(\cup E_n) = \sum \mu(E_n) = \sum \bar{\mu}(A_n).$$

So, $\bar{\mu}$ is countably additive. Let's show that $\bar{\mu}$ is complete. If $Y \in \bar{\mathcal{M}}$ with $\bar{\mu}(Y) = 0$, and $Z \subset Y$, then there is $N \in \mathcal{N}$ such that $Y \subset N$ with $\mu(N) = 0$. Consequently we also have $Z \subset N$ with $\mu(N) = 0$ hence $Z = \emptyset \cup Z \in \bar{\mathcal{M}}$ (with "E" $= \emptyset$ and $F = Z \subset N \in \mathcal{N}$). Therefore, $\bar{\mu}$ is a complete measure on $\bar{\mathcal{M}}$.

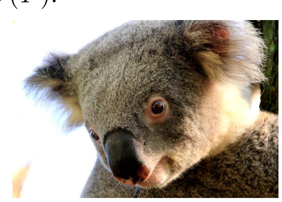
Finally the uniqueness. Let's assume ν also extends μ to a complete measure. Consequently, $\nu(A) = \mu(A)$ for all $A \in \mathcal{M}$. It follows that the elements of $\mathcal{N}$ also have ν-measure zero. By the completeness proposition, all subsets of elements of $\mathcal{N}$ must be elements of the σ-algebra corresponding to ν, and conversely, all subsets of $\mathcal{N}$ must be elements of $\bar{\mathcal{M}}$, and so presuming the σ-algebra corresponding to ν is the smallest possible needed to complete μ, it must coincide with $\bar{\mathcal{M}}$.

For $Y = E \cup F \in \bar{\mathcal{M}}$,

$$\nu(Y) \leq \nu(E) + \nu(N) = \mu(E) + \mu(N) = \mu(E) = \bar{\mu}(Y),$$

and conversely

$$\bar{\mu}(Y) \leq \bar{\mu}(E) + \bar{\mu}(N) = \mu(E) = \nu(E) \leq \nu(Y).$$



So, we've got equality all across, and in particular, $\nu(Y) = \bar{\mu}(Y)$.

Proposition 3.3 (Null Set Proposition)**.** *Let $(X, \mathcal{M}, \mu)$ be a non-trivial measure space, meaning there exist measurable subsets of positive measure. Then*

$$\mathcal{N} := \{Y \in \mathcal{M} : \mu(Y) = 0\}$$

is not *a σ-algebra, but it* is *closed under countable unions.*

Proof: If $\{N_n\} \subset \mathcal{N}$ is a countable collection, then since $\mathcal{M}$ is a σ-algebra,

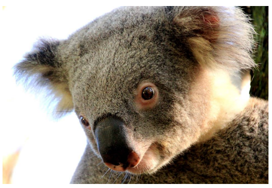
$$\cup N_n \in \mathcal{M}.$$

Moreover, we have

$$\mu(\cup N_n) \leq \sum \mu(N_n) = 0 \implies \mu(\cup N_n) = 0.$$

This shows that $\mathcal{N}$ is closed under countable unions. Why is it however, *not* a σ-algebra? It's not even an algebra! This is because it is not closed under complements. What is always an element of $\mathcal{N}$? The $\emptyset$ is always measurable and has measure zero. Hence $\emptyset \in \mathcal{N}$. What about its complement? This is where the non-triviality hypothesis plays a role. There is some $Y \in \mathcal{M}$ such that $\mu(Y) > 0$. Since $Y \subset X$, by monotonicity

$$\mu(X) \geq \mu(Y) > 0 \implies X = \emptyset^c \notin \mathcal{N}.$$



4. Extension of pre-measures

What seems an intuitive way to prove Lebesgue's theorem is to use our notion of volume v_n defined on disjoint unions of intervals. This happens to be an example of something called *pre-measure.*

Definition 4.1. Let $\mathcal{A} \subset P(X)$ be an algebra. A function $\mu_0 : \mathcal{A} \to [0, \infty]$ is called a *pre-measure* if

(1) $\mu(\emptyset) = 0$
(2) If $\{A_j\}$ is a countable collection of disjoint elements of $\mathcal{A}$ such that
$$\cup A_j \in \mathcal{A},$$
then
$$\mu_0(\cup A_j) = \sum \mu_0(A_j).$$

The name pre-measure is appropriate because it's almost a measure, it's just possibly not countably additive for *every* disjoint countable union, since these need not *always* be contained in a mere algebra (which is not necessarily a σ-algebra). However, Carathéodory can help us to extend pre-measures to measures. First, we require the following.

Proposition 4.2. *If μ_0 is a pre-measure on $\mathcal{A}$ and*

$$\mu * (Y) := \inf\{\sum \mu_0(A_j) : A_j \in \mathcal{A} \forall j, Y \subset \cup A_j\},$$

then (i) $\mu^*(A) = \mu_0(A) \forall A \in \mathcal{A}$ and (ii) every set in $\mathcal{A}$ is $\mu*$ measurable.

Proof: First note that pre-measures are by definition finitely additive since for $A, B \in \mathcal{A}$ with $A \cap B = \emptyset$, then

$$A \cup B = \cup A_j, \quad A_1 = A, A_2 = B, A_j = \emptyset \forall j > 2.$$

We further note that finite additivity imply monotonicity for all elements of the algebra, so if $A \subset B$ are both elements of $\mathcal{A}$, then

$$\mu(B) = \mu(B \setminus A) + \mu(A) \implies \mu(A) = \mu(B) - \mu(B \setminus A) \leq \mu(B).$$

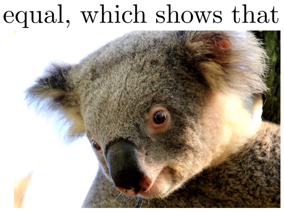
The union is in $\mathcal{A}$ because it's an algebra, and since $\mu_0(\emptyset) = 0$, the definition of pre-measure shows that

$$\mu_0(A \cup B) = \sum \mu_0(A_j) = \mu_0(A) + \mu_0(B).$$

To prove (i) let $E \in \mathcal{A}$. If $E \subset \cup A_j$ with $A_j \in \mathcal{A}\ \forall j$, then let

$$B_n := E \cap (A_n \setminus \cup_1^{n-1} A_j).$$

Then
$$B_n \in \mathcal{A} \forall n, \quad B_n \cap B_m = \emptyset \forall n \neq m.$$
Since the union
$$\cup B_n = E \in \mathcal{A},$$
by definition of pre-measure,
$$\mu_0(E) = \mu_0(\cup B_n) = \sum \mu_0(B_n) \leq \sum \mu_0(A_n),$$
since $B_n \subset A_n \forall n$. Taking the infimum over all such covers of E comprised of elements of $\mathcal{A}$, we have
$$\mu_0(E) \leq \mu * (E).$$
On the other hand, $E \subset \cup A_j$ with $A_1 = E \in \mathcal{A}$, and $A_j = \emptyset \forall j > 1$. Then, this collection is considered in the infimum defining $\mu*$, so
$$\mu * (E) \leq \sum \mu_0(A_j) = \mu(E).$$
We've shown the inequality is true in both directions, hence $\mu^*(E) = \mu_0(E)$.
To show (ii) if $A \in \mathcal{A}$ and $E \subset X$ and $\varepsilon > 0$ there exists $\{B_j\} \subset \mathcal{A}$ with $E \subset \cup B_j$ and
$$\sum \mu_0(B_j) \leq \mu^*(E) + \varepsilon.$$
Since μ_0 is additive on $\mathcal{A}$,
$$\mu*(E)+\varepsilon \geq \sum \mu_0(B_j \cap A)+\mu_0(B_j \cap A^c) = \sum \mu_0(B_j \cap A)+\sum \mu_0(B_j \cap A^c) \geq \mu*(E \cap A)+\mu*(E \cap A^c).$$
This is true for any $\varepsilon > 0$, so we have
$$\mu * (E) \geq \mu * (E \cap A) + \mu * (E \cap A^c) \geq \mu * (E).$$
So, these are all equal, which shows that A satisfies the definition of being μ^* measurable since E was arbitrary.



Now we will prove that we can always extend a pre-measure to a measure.

Theorem 4.3 (Pre-measure extension theorem)**.** *Let $\mathcal{A} \subset P(X)$ be an algebra, μ_0 a pre-measure on $\mathcal{A}$, and $\mathcal{M}$ the smallest σ-algebra generated by $\mathcal{A}$. Then there exists a measure μ on $\mathcal{M}$ which extends μ_0, namely*
$$\mu := \mu^* \text{ restricted to } \mathcal{M}.$$
If ν also extends μ_0 then $\nu(E) \leq \mu(E) \forall E \in \mathcal{M}$ with equality when $\mu(E) < \infty$. If μ_0 is σ-finite, then $\nu \equiv \mu$ on $\mathcal{M}$, so μ is the unique extension.

Proof: The existence of μ follows from Carathéodory's theorem. Note that the σ-algebra in that theorem must contain $\mathcal{A}$, since all elements of $\mathcal{A}$ are μ^* measurable by the pre-measure proposition. Consequently, the Carathéodory σ-algebra, on which μ^* is a measure, contains $\mathcal{M}$, and therefore $\mu*$ restricted to $\mathcal{M}$ is a measure since it must still satisfy the requisite properties on $\mathcal{M}$ which is contained in the possibly larger Carathéodory σ-algebra. We will investigate when in fact these algebras coincide.
So, we only need to consider the statements about a possibly different extension ν which coincides with μ_0 on $\mathcal{A}$ and is a measure on $\mathcal{M}$. If $E \in \mathcal{M}$ and
$$E \subset \cup A_j, \quad A_j \in \mathcal{A} \forall j,$$
then
$$\nu(E) \leq \sum \nu(A_j) = \sum \mu_0(A_j).$$

This holds for any such covering of E by elements of $\mathcal{A}$, so taking the infimum we have

$$\nu(E) \leq \mu*(E) = \mu(E) \text{ since } E \in \mathcal{M}.$$

If $\mu(E) < \infty$, let $\varepsilon > 0$. Then we may choose $\{A_j\} \subset \mathcal{A}$ which are WLOG (without loss of generality) disjoint (why?) such that

$$E \subset \cup A_j, \quad \mu(\cup A_j) = \sum \mu_0(A_j) < \mu^*(E) + \varepsilon = \mu(E) + \varepsilon,$$

since $E \in \mathcal{M}$. Note that for

$$A = \cup A_j, \quad \nu(A) = \lim_{n\to\infty} \nu(\cup_1^n A_j) = \lim_{n\to\infty} \sum_1^n \nu(A_j) = \lim_{n\to\infty} \sum_1^n \mu_0(A_j) = \mu(A).$$

Then we have since $E \in \mathcal{M}$,

$$\mu(A) = \mu(A \cap E) + \mu(A \setminus E) = \mu(E) + \mu(A \setminus E) < \mu(E) + \varepsilon$$

which shows that

$$\mu(A \setminus E) < \varepsilon.$$

Consequently,

$$\mu(E) \leq \mu(A) = \nu(A) = \nu(E \cap A) + \nu(A \setminus E) \leq \nu(E) + \mu(A \setminus E) < \nu(E) + \varepsilon.$$

This holds for all $\varepsilon > 0$, so

$$\mu(E) \leq \nu(E).$$

Consequently in this case $\mu(E) = \nu(E)$.
Finally, if $X = \cup A_j$ with $A_j \in \mathcal{A}$, $\mu_0(A_j) < \infty \forall j$, we may WLOG assume the A_j are disjoint. Then for $E \in \mathcal{M}$,

$$E = \cup(E \cap A_j),$$

which is a disjoint union so by countable additivity

$$\mu(E) = \mu(\cup E \cap A_j) = \sum \mu(E \cap A_j) = \sum \nu(E \cap A_j),$$

since $E \cap A_j \subset A_j$ shows that $\mu(E \cap A_j) \leq \mu(A_j) < \infty$, so $\mu(E \cap A_j) = \nu(E \cap A_j)$.

5. Metric outer measures and Hausdorff measure

In the following we will make a little detour and introduce *metric outer measures*. These are outer measures defined on metric spaces with one crucial additional property. We consider: metric space (X, d) and for $A, B \subset X$ define

$$\operatorname{dist}(A, B) := \inf\{d(x, y) : x \in A, y \in B\}.$$

Define also the diameter of a set $A \subset X$

$$\operatorname{diam}(A) := \sup\{d(x, y) : x, y \in A\}, \operatorname{diam}(\emptyset) := 0.$$

Definition 5.1. Given an outer measure μ^* on (X, d). Then μ^* is called *metric outer measure* iff for each $A, B \subset X$ we have

$$\operatorname{dist}(A, B) > 0 \Rightarrow \mu^*(A \cup B) = \mu^*(A) + \mu^*(B).$$

Recall: $A \subset X$ is μ^*-measurable iff for each $E \subset X$

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^C).$$

Denote by $\mathcal{M}(\mu^*)$ the μ^*-measurable subsets. We now prove a Theorem due to Carathéodory which states that the Borel sets in X are contained in $\mathcal{M}(\mu^*)$. Recall that the Borel sets $\mathcal{B}(X)$ is the smallest σ-algebra generated by the topology of X (induced by the metric).

Theorem 5.2 (Carathéodory)**.** *Let μ^* be a metric outer measure on (X,d). Then we have $\mathcal{B}(X) \subset \mathcal{M}(\mu^*)$.*

Proof:
Note that since $\mathcal{M}(\mu^*)$ is a σ-algebra (by Thm. 2.3) it is enough to prove that every closed set is μ^*-measurable. So let $F \subset X$ be a closed subset. It suffices to show that for any set A

$$\mu(A) \geq \mu^*(A \cap F) + \mu^*(A \setminus F).$$

Define the sets

$$A_k := \{x \in A : \operatorname{dist}(x, F) \geq \frac{1}{k}\}.$$

Then $\operatorname{dist}(A_k, A \cap F) \geq \frac{1}{k}$, so since μ^* is metric we have

$$\mu^*(A \cap F) + \mu^*(A_k) = \mu^*(\underbrace{(A \cap F) \cup A_k}_{\subset A}) \leq \mu^*(A). \tag{+}$$

Then

$$A \setminus F = \bigcup A_k$$

since F is closed (which gives $\forall_{x \in A \setminus F} \operatorname{dist}(x, F) > 0$) and (A_k) is increasing.
The main and last step in the proof is to calculate the limit in (+). If the limit is infinity there is nothing to do. Hence assume the limit exists.
For this define a pairwise disjoint cover like this: $B_1 := A_1, B_2 := A_2 \setminus A_1, B_3 := A_3 \setminus A_2$ etc..
Then we show that for $|j - k| \geq 2$ we have $\operatorname{dist}(B_i, B_j) > 0$. This follows from the inclusions for $i \geq j + 2$

$$B_i \subset A \setminus (F \cup A_{i-1}) \subset A \setminus (F \cup A_{j+1}).$$

But $x \in A \setminus (F \cup A_{j+1})$ implies that there is a $z \in F$ with

$$d(x, z) \geq \frac{1}{j+1}$$

hence

$$d(x, y) \geq d(x, z) - d(y, z) \geq \frac{1}{j} - \frac{1}{j+1} > 0.$$

And thus $\operatorname{dist}(B_i, B_j) > 0$.
This means we can apply the metric property (for even and odd indices) and by induction we conclude that

$$\mu^*\left(\bigcup_{k=1}^{n} B_{2k-1}\right) = \sum_{k=1}^{n} \mu^*(B_{2k-1}),$$
$$\mu^*\left(\bigcup_{k=1}^{n} B_{2k}\right) = \sum_{k=1}^{n} \mu^*(B_{2k}).$$

Because these unions are contained in A_{2n} the sums are $\leq \mu^*(A_{2n})$. The values $\mu^*(A_{2n})$ are increasing and by assumption bounded. Hence both sums are convergent for $n \to \infty$.

Therefore we conclude for any j

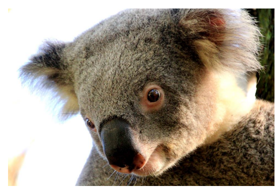
$$\begin{aligned}
\mu^*(A \setminus F) &= \mu^*\left(\bigcup_i A_i\right) \\
&= \mu^*\left(A_j \cup \bigcup_{k \geq j+1} B_k\right) \\
&\leq \mu^*(A_j) + \sum_{k=j+1}^{\infty} \mu^*(B_k) \\
&\leq \lim_{n\to\infty} \mu^*(A_n) + \underbrace{\sum_{k=j+1}^{\infty} \mu^*(B_j)}_{\to 0, j\to\infty}.
\end{aligned}$$

Since the latter sum goes to 0 by convergence we obtain

$$\mu^*(A \setminus F) \leq \lim_{n\to\infty} \mu^*(A_n).$$

Together with (+) this yields

$$\mu^*(A) \geq \lim_{k\to\infty} \mu^*(A_k) + \mu^*(A \cap F) \geq \mu^*(A \setminus F) + \mu^*(A \cap F)$$

which is the desired inequality. 

We let $\mathcal{C}$ denote a collection of sets which cover X. Then for each $A \subset X$ we denote by $\mathcal{CC}(A)$ the collection of sets in $\mathcal{C}$ such that there is an at most countable sequence of sets $\{E_n\}_{n\in\mathbb{N}} \in \mathcal{CC}(A)$ such that

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

These are the *countable covers* of A by sets belonging to $\mathcal{C}$.

Definition 5.3. *i)* Fix on the metric space a set function $\nu\colon \mathcal{C} \to [0,\infty]$ with $\nu(\emptyset) = 0$. We define the following set function depending on $\mathcal{C}, \nu$

$$\mu^*_{\nu,\mathcal{C}}(A) := \inf_{\mathcal{D}\in\mathcal{CC}(A)} \sum_{D\in\mathcal{D}} \nu(D). \tag{5.1}$$

Theorem 5.4. *The measure given by* (5.1) *is the unique outer measure μ^* on X such that*

$$\mu^*(A) \leq \nu(A), \;\; A \in \mathcal{C}$$

and for any other outer measure $\tilde{\mu}^*$ with the above condition we have

$$\tilde{\mu}^*(A) \leq \mu^*(A), \;\; A \subset X.$$

The proof follows basically the same lines as the construction of Lebesgue outer measure and is therefore omitted.

Given the same data as in the above definition we define for $\epsilon > 0$

$$\mathcal{C}_\epsilon := \{A \in \mathcal{C} : \operatorname{diam}(A) < \epsilon\}$$

and assume this is a cover for X (i.e. each $x \in X$ is covered by a $C \in \mathcal{C}$ with $\operatorname{diam}(C) < \epsilon$). Now define the measure depending on this cover as a special case of (5.1), in particular we set

$$\mu^*_\epsilon(A) := \mu_{\nu,\mathcal{C}_\epsilon}(A).$$

As $\mathcal{C}_\epsilon \subset \mathcal{C}_{\epsilon'}$ for $\epsilon' < \epsilon$ we have

$$\mu_{\epsilon'}(A) \geq \mu_\epsilon(A).$$

Theorem 5.5. *The limit* $\mu_0^*(A) := \lim_{\epsilon\to 0} \mu_\epsilon^*(A), A \subset X$ *defines a metric outer measure.*

Proof: The outer measure property is preserved under this limit (Exercise).
Let $A, B \subset X$ be such that $\operatorname{dist}(A,B) > 0$ and hence $\operatorname{dist}(A,B) > \frac{1}{n}$ for $n > n_0$. Let $\delta > 0$ be given and cover the union $A \cup B$ with sets E_k^n such that

$$\mu^*_{\frac{1}{n}}(A \cup B) + \delta \geq \sum_{k=1}^{\infty} \nu(E_k^n)$$

and such that for each k we have $\operatorname{diam}(E_k^n) \leq \frac{1}{n}$. Hence we have that the E_k^n intersect either A or B and not both in the sense that

$$E_k^n \cap A \neq \emptyset \Rightarrow E_k^n \cap B = \emptyset,\ E_k^n \cap B \neq \emptyset \Rightarrow E_k^n \cap A = \emptyset.$$

Denote by $\{\tilde{E}_k^n\} \subset \{E_k^n\}$ the subsequence of sets such that $\tilde{E}_k^n \cap (A \cup B) \neq \emptyset$ for each k.
Define also the subsequences

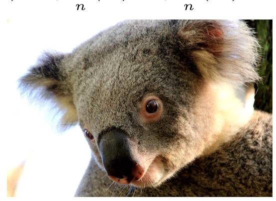
$$E^n(A) := \{\tilde{E}_k^n : \tilde{E}_k^n \cap A \neq \emptyset\}, E^n(B) := \{\tilde{E}_k^n : \tilde{E}_k^n \cap B \neq \emptyset\}$$

and as already remarked $E^n(A)$ and $E^n(B)$ have no sets in common and together they yield the sequence $(\tilde{E}_k^n)_{k=1}^{\infty}$.
We can then write

$$\mu^*_{\frac{1}{n}}(A) \geq \sum_{E \in E^n(A)} \nu(E) - \frac{\delta}{2}, \mu^*_{\frac{1}{n}}(B) \geq \sum_{E \in E^n(B)} \nu(E) - \frac{\delta}{2}.$$

Hence by exclusion of extraneous sets in the cover we can show that (for n sufficiently large) it follows

$$\mu^*_{\frac{1}{n}}(A \cup B) \geq \mu^*_{\frac{1}{n}}(A) + \mu^*_{\frac{1}{n}}(B).$$



It follows in particular that μ_0^* is metric.
A particular case of the canonical metric outer measure is the so-called Hausdorff measure.

Definition 5.6 (Hausdorff pre-measures)**.** Let (X,d) be a metric space, $S \subset X, \delta > 0$ and $t \in [0,\infty)$, then define the set function

$$\mathcal{H}_\delta^t(S) := \inf \left\{ \sum_{i=1}^{\infty} (\operatorname{diam} U_i)^t | \bigcup_{i=1}^{\infty} U_i \supseteq S, \operatorname{diam}(U_i) < \delta \right\}$$

where the infimum is taken over all countable covers of S by sets $U_i \subset X$ with $\operatorname{diam}(U_i) < \delta$.

Remark 3.
- Setting $\nu(U) := \operatorname{diam}(U)^t$ then $\mathcal{H}_\delta^t(S) = \mu^*_{\nu,\mathcal{C}_\delta}(S)$ is just a special case of our canonical outer measure.

  We therefore immediately know some things. First:

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

  makes sense as a definition of outer measure and is called *Hausdorff-measure*.

  We know that $\mathcal{H}^t$ is a metric outer measure 5.5, all the Borel sets are $\mathcal{H}^t$-measurable 2.3 and the $\mathcal{H}^t$-measurable sets form a σ-algebra 5.5.
- If we consider the special case $(\mathbb{R}^n, |\cdot|) = (X,d)$ with the standard euclidean metric then the Hausdorff measure $\mathcal{H}^n$ agrees for $n \in \mathbb{N}$ (up to a scaling factor for $n > 1$) with Lebesgue outer measure λ^n (Exercise).

6. Existence and uniqueness of Lebesgue measure

Theorem 6.1. *There exists a unique complete measure on $\mathbb{R}^n$ known as the* Lebesgue Measure*, which extends ν_n to the smallest σ-algebra containing ε_n such that the extension on this σ-algebra is complete.*

Proof:
To make an algebra containing ε_n, in particular the smallest algebra containing ε_n, it is necessary to include compliments. Define

$$\mathcal{A} := \{Y \subseteq \mathbb{R}^n \mid Y \in \varepsilon_n \text{ or } \exists Z \in \varepsilon_n \text{ s.t. } Y = Z^c\}$$

Claim 1. *$\mathcal{A}$ is an algebra.*

Proof:

(1) $\emptyset = \prod \mathrm{I}x, x\mathrm{I}$ for $x \in \mathbb{R}^n$. *Notation:* we use $\mathrm{I}a, b\mathrm{I}$ to denote either $]a, b[$, $[a, b]$, $]a, b]$ or $[a, b[$. Notation which is unnecessary shall be simplified when possible.
(2) By definition, $\mathcal{A}$ is closed under compliments
(3) Let $A, B \in \mathcal{A}$. If $A, B \in \varepsilon_n$ then first consider the case where A, B are each single intervals i.e. $A = \prod \mathrm{I}a_i, \alpha_i\mathrm{I}, B = \prod \mathrm{I}b_i, \beta_i\mathrm{I}$ for $a_i \leq \alpha_i,\ b_i \leq \beta_i$. For each i, if $\mathrm{I}b_i, \beta_i\mathrm{I} \subset \mathrm{I}a_i, \alpha_i\mathrm{I}$ then note that

$$\mathrm{I}a_i, \alpha_i\mathrm{I} \backslash \mathrm{I}b_i, \beta_i\mathrm{I} = \mathrm{I}a_i, b_i\mathrm{I} \cup \mathrm{I}\beta_i, \alpha_i\mathrm{I}$$

If $\mathrm{I}b_i, \beta_i\mathrm{I} \not\subset \mathrm{I}a_i, \alpha_i\mathrm{I}$, then either $\mathrm{I}b_i, \beta_i\mathrm{I} \cap \mathrm{I}a_i, \alpha_i\mathrm{I} = \emptyset$ in which case $\mathrm{I}a_i, \alpha_i\mathrm{I} \backslash \mathrm{I}b_i, \beta_i\mathrm{I} = \mathrm{I}a_i, \alpha_i\mathrm{I}$, or $\mathrm{I}b_i, \beta_i\mathrm{I} \cap \mathrm{I}a_i, \alpha_i\mathrm{I} \neq \emptyset$ so that

$$\mathrm{I}a_i, \alpha_i\mathrm{I} \backslash \mathrm{I}b_i, \beta_i\mathrm{I} = \begin{cases} \mathrm{I}a_i, b_i\mathrm{I} \text{ if } b_i \leq \alpha_i (\Rightarrow \beta_i > \alpha_i) \\ \mathrm{I}\beta_i, \alpha_i\mathrm{I} \text{ if } a_i \leq beta_i (\Rightarrow b_i < a_i) \end{cases}$$

In both cases $\mathrm{I}a_i, \alpha_i\mathrm{I} \backslash \mathrm{I}b_i, \beta_i\mathrm{I}$ is the disjoint union of intervals. Repeating for each $i = 1, ..., n$ gives $A \backslash B \in \varepsilon_n$, and similarly $B \backslash A \in \varepsilon_n$. Note that $A \cap B = \prod \mathrm{I}x_i, y_i\mathrm{I}$ with $x_i = \max\{a_i, b_i\}$, $y_i = \min\{\alpha_i, \beta_i\}$ (and should $x_i \geq y_i$ then it is understood that $\mathrm{I}x_i, y_i\mathrm{I} = \emptyset$. Therefore,

$$A \cup B = (A \backslash B) \cup (B \backslash A) \cup (A \cap B) \in \varepsilon_n.$$

In fact, for $A = \prod \mathrm{I}a_i, \alpha_i\mathrm{I} \in \varepsilon_n$ note that

$$\begin{aligned} A^c =& \mathbb{R}^n \backslash A \\ =& \prod \mathrm{I} -\infty, a_i\mathrm{I} \cup \prod \mathrm{I}\alpha_i, \infty\mathrm{I} \end{aligned}$$

Allowing the endpoints x_i and/or y_i of $\mathrm{I}x_i, y_i\mathrm{I}$ to be $\pm\infty$, the same arguments for A, B as above show that $A^c \cup B$ and $A^c \cup B^c$ are elements of $\mathcal{A}$.

More generally, for $A = \bigcup_{j=1}^{k} I_j \in \varepsilon_n$ with $I_j \bigcap_{k \neq j} I_k = \emptyset$ and $B = \bigcup_{l=1}^{m} J_l \in \varepsilon_n$ with $J_l \bigcap_{m \neq l} J_m = \emptyset$ with end points possibly $\pm\infty$, repeated application of the above arguments shows that $I_1 \cup J_1 \in \varepsilon_n$, $(I_1 \cup J_1) \cup I_2 \in \varepsilon_n$, and so forth. Therefore, $A \cup B \in \varepsilon_n$. So $\mathcal{A}$ is closed under finite unions and hence $\mathcal{A}$ is an algebra.

Homework: Show that ν_n is well-defined on $\mathcal{A}$ where

$$\nu_n(\prod^{n} \mathrm{I}a_i, \alpha_i\mathrm{I}) := \begin{cases} 0, \text{ if } a_i = \alpha_i \text{ for some } i \\ \prod(\alpha_i - a_i), \text{ else} \end{cases}$$

Claim 2. *ν_n is a pre-measure on $\mathcal{A}$.*

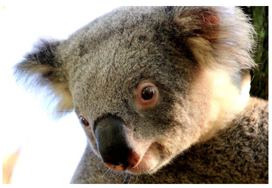
Proof:

(1) $\nu_n(\emptyset) = 0$, by definition.
(2) Let $\{A_m\}_{m\geq 1} \subset \mathcal{A}$ such that $\underset{m\geq 1}{\cup} A_m \in \mathcal{A}$, $A_m \underset{k\neq m}{\cap} A_k = \emptyset$ then $\exists \{I_j\}_{j=1}^{k}$ disjoint in $\mathcal{A}$ such that $\bigcup_{j=1}^{k} I_j = \bigcup_{m=1}^{\infty} A_m$.

By definition, $\nu_n(\bigcup_{m=1}^{M} A_m) = \sum_{m=1}^{M} v_n(A_m) \leq \nu_n(\bigcup_{j=1}^{k} I_j) = \sum_{j=1}^{k} v_n(I_j)$

$\forall M \in \mathbb{N}$, $\sum_{m=1}^{M} v_n(A_m) \leq \sum_{j=1}^{k} v_n(I_j) = \nu_n(\bigcup_{m=1}^{\infty} A_m) \leq \sum_{m=1}^{M} v_n(A_m)$

$\Rightarrow \nu_n(\bigcup_{m=1}^{\infty} A_m) = \sum_{m=1}^{M} v_n(A_m)$



So ν_n is a pre-measure on the algebra $\mathcal{A}$. Note by the definition of $\mathcal{A}$, it is the smallest algebra which contains ε_n. By the pre-measure extension theorem, since ν_n is σ-finite on $\mathcal{A}$, there exists a unique extension of ν_n to a measure $\bar{\mathcal{M}}$ on the smallest σ-algebra containing ε_n. Unique, because $\mathbb{R}^n = \underset{m\geq 1}{\cup}[-M, M]^n = \underset{m\geq 1}{\cup} I_M$ and $\nu_m(I_M) = (2M)^n < \infty$ for each M.

Canonically completing this measure to $\mathcal{M}$ by applying the completion theorem yields the Lebesgue measure and the Lebesgue σ-algebra, the smallest σ-algebra generated by ε_n such that the extension of ν_n to a measure with respect to this σ-algebra is complete.

Remark 4. In the completion theorem, $\bar{\mathcal{M}}$ is the smallest σ-algebra containing $\mathcal{M}$ such that $\tilde{\mu}$ is complete, recalling

$$\bar{\mathcal{M}} = \{A \subset X \mid \exists E \in \mathcal{M}, F \subset N \in \mathcal{M}, A\ E \cup F\}$$
$$\mathcal{N} = \{Y \in \mathcal{M} \mid \mu(Y) = 0\}.$$

This follows from $\bar{\mu}$ is complete $\Leftrightarrow \forall Z \subset X$ such that $\exists Y \in \bar{\mathcal{M}}$ with $\bar{\mu}(Y) = 0$, $Z \subset Y \Rightarrow X \in \bar{\mathcal{M}} \Rightarrow \forall Z \subset X$ with $Z \subset Y \in \mathcal{M} \in \bar{\mathcal{M}}$ such that $\mu(Y) = 0 \Rightarrow Z \in \bar{\mathcal{M}}$. So if $A \subset X$ such that $\exists E \in \mathcal{M}$ and $F \subset N\nabla\mathcal{M}$ such that $A = E \cup F$ then F must be $\bar{\mu}$ measurable. By definition of the extension E is $\bar{\mu}$ measurable, which implies $E \cup F$ must be $\bar{\mu}$ measurable. The σ-algebra for any complete extension of μ must contain all such sets, therefore contains $\bar{\mathcal{M}}$.

6.1. **Properties of the Lebesgue σ-algebra.**

(1) Borel sets are Lebesgue measurable. To prove this, it suffices to show that open sets are Lebesgue measurable. So, let $\mathcal{O} \subset \mathbb{R}^n$ be open. Then we will show that $\mathcal{O} \in \mathcal{M}$.

First consider $\mathcal{O} = \prod]a_i, \alpha_i[\in \varepsilon_n \subset \mathcal{M}$. For an arbitrary open set $\mathcal{O}$, for each $x \in \mathcal{O}$ there exists $\varepsilon \in \mathbb{Q}, \varepsilon > 0$ such that $x \in \prod\ ]q_m - \varepsilon, q_m + \varepsilon[\subset \mathcal{O}$, $q_m \in \mathbb{Q}$, $m = 1, ..., n$.

Taking the union of all such intervals, namely those contained in $\mathcal{O}$ such that endpoints are rational is a countable union. Countability of course follows since $\mathbb{Q}^n \subset \mathbb{R}^n$ is countable and $\mathbb{Q}$ is countable so a union of intervals with endpoints in $\mathbb{Q}^n$ is countable. Therefore, $\mathcal{O} \in \mathcal{M}$.

(2) *Exercise*: Prove $\mathcal{B} \subsetneq \mathcal{M}$
(3) It is difficult to construct sets $\not\subset \mathcal{M}$, but actually there are many natural examples... *Exercise*: Construct a subset of $\mathbb{R}^n$ which is not measurable. Recall that $f : \mathbb{R}^n \to \mathbb{R}^m$ is "measurable" usually is understood to mean that $\forall B \in \mathcal{B}^m$, $f^{-1}(B) \in \mathcal{M}^n$. More precisely, f is $(\mathbb{R}^n, \mathcal{B}^n), (\mathbb{R}^m, \mathcal{B}^m)$ measurable. In general, $f : X \to Y$ is $(X, \mathcal{A}), (Y, \mathcal{B})$ measurable if $\forall B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$, where $\mathcal{A}$ and $\mathcal{B}$ are σ-algebras.
(4) $n-1$ dimensional sets have $\mathcal{L}^n$ measure 0. Note that by completeness, this implies that lower than $n-1$ dimensional sets have $\mathcal{L}^n = 0$. WLOG $Y = \{x_1, .., x_n\} \in \mathbb{R}^n \mid x_q = 0\}$.
$$Y = [0,0]x] - \infty, \infty[^{n-1} \Rightarrow \nu_n(Y) = \mathcal{L}^n(Y) = 0$$
All subsets are then measurable and have measure 0.

7. Hausdorff measures revisited

In geometric analysis it is useful to have a method for describing the size of lower dimensional sets in $\mathbb{R}^n$, such as curves and surfaces in $\mathbb{R}^3$. - Gerald Folland

We have seen that all such sets have $\mathcal{L}^3$ measure equal to 0. So, $\mathcal{L}^n$ is too coarse for lower dimensional subsets.

Definition 7.1 (Metric outer measures)**.** On a metric space (X, ρ) we call μ^* a *metric* outer measure if $A, B \subset X$ and $\rho(A, B) > 0$ implies $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$.

Proposition 7.2. *Borel sets are always μ^* measurable, for any metric outer measure.*

Proof:
It suffices to show that closed sets are μ^* measurable because they generate $\mathcal{B}$. Let $F \subset X$ be closed. For $A \subset X$, we want to show that $\mu^*(A) = \mu^*(A \cap F) + \mu^*(A \setminus F)$.
For sets A with $\mu^*(A) = \infty$, we have
$$\mu^*(A) \geq \mu^*(A \cap F) + \mu^*(A \setminus F) \geq \mu^*(A)$$
which implies the equality. From now we assume $\mu^*(A) < \infty$. Let
$$B_n = \{x \in A \setminus F : \rho(x, F) \geq \frac{1}{n}\},$$
so we obtain $B_n \subset B_{n+1} \subset \cdots$. If $x \in A \setminus F$, then noting that $A \setminus F = A \cap F^c$, and F^c is open, there exists $n \in \mathbb{N}$ such that $B_{\frac{1}{n}}(x) \in F^c$ and therefore $\rho(x, F) \geq \frac{1}{n}$. So due to the definition of B_n, we have $x \in B_n$. Hence
$$\bigcup_{n \geq 1} B_n \subset A \setminus F \subset \bigcup_{n \geq 1} B_n \quad \Rightarrow \quad A \setminus F = \bigcup_{n \geq 1} B_n.$$
For each n, $\rho(B_n, F) \geq \frac{1}{n}$, so by definition since $(A \cap F) \cup B_n \subset A$, we have
$$\mu^*(A) \geq \mu^*((A \cap F) \cup B_n) = \mu^*(A \cap F) + \mu^*(B_n).$$
This follows by the definition of metric outer measure because $\rho(B_n, A \cap F) \geq \frac{1}{n}$. Therefore it suffices to show that
$$\mu^*(B_n) \to \mu^*(A \setminus F).$$
Let $C_n := B_{n+1} \setminus B_n$. If $x \in C_{n+1}$ and $\rho(x, y) < \frac{1}{n(n+1)}$ then
$$\rho(y, F) \leq \rho(x, y) + \rho(x, F) < \frac{1}{n(n+1)} + \frac{1}{n+1} = \frac{1}{n}.$$
Hence $\rho(y, F) < \frac{1}{n} \Rightarrow y \notin B_n$. So any y with $\rho(x, y) < \frac{1}{n(n+1)}$ is not in B_n, where $x \in C_{n+1}$. Therefore $\rho(x, y) \geq \frac{1}{n(n+1)}$ holds for all $y \in B_n$ and for each $x \in C_{n+1}$. $\Rightarrow \rho(C_{n+1}, B_n) \geq$

$\frac{1}{n(n+1)}$. Then

$$\begin{aligned}\mu^*(B_{2k+1}) &\geq \mu^*(C_{2k} \cup B_{2k-1}) = \mu^*(C_{2k}) + \mu^*(B_{2k-1}) \\ &\geq \mu^*(C_{2k}) + \mu^*(C_{2k-2} \cup B_{2k-3}) \\ &\geq \cdots \geq \sum_{j=1}^{k} \mu^*(C_{2j}) + \mu^*(B_1).\end{aligned}$$

Also,

$$\begin{aligned}\mu^*(B_{2k}) &\geq \mu^*(C_{2k-1} \cup B_{2k-2} = \mu^*(C_{2k-1}) + \mu^*(B_{2k-2}) \\ &\geq \mu^*(C_{2k-1}) + \mu^*(C_{2k-3} \cup B_{2k-4}) \\ &\geq \cdots \geq \sum_{j=1}^{k} \mu^*(C_{2j-1}) + \mu^*(B_1).\end{aligned}$$

Now since $B_n \subset A$ is true for all n, we get $\mu^*(B_n) \leq \mu^*(A)$ for all n. Then we obtain

$$\begin{aligned}\mu^*(A) &\geq \sum_{j=1}^{k} \mu^*(C_{2j}) + \mu^*(B_1) \text{ and} \\ \mu^*(A) &\geq \sum_{j=1}^{k} \mu^*(C_{2j-1}) + \mu^*(B_1) \text{ for all } k.\end{aligned}$$

As a consequence, $\sum_{j=1}^{\infty} \mu^*(C_j)$ converges and therefore $\sum_{j=n}^{\infty} \mu^*(C_j) \to 0$ with $n \to \infty$. We have

$$A \setminus F \subset \bigcup_{n\geq 1} B_n = B_1 \cup \bigcup_{n \geq 1} C_n = B_n \cup \bigcup_{k \geq n} C_k.$$

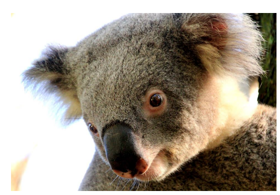
Now back to $\mu^*(A \setminus F)$:

$$\mu^*(A \setminus F) \leq \mu^*(B_n) + \sum_{j=n}^{\infty} \mu^*(C_j)$$

holds and it follows that

$$\mu^*(A \setminus F) \leq \lim_{m\to\infty} \inf_{n \geq m} \mu^*(B_n) \leq \lim_{m \to \infty} \sup_{n \geq m} \mu^*(B_n) \leq \mu^*(A \setminus F),$$

since $\mu^*(B_n) \leq \mu^*(A \setminus F)$ is true for all n. Therefore F is μ^*-measurable.



Definition 7.3 (Hausdorff measure)**.** Let (X, ρ) be a metric space, $p \geq 0$ and $\delta > 0$. For $A \subset X$, let

$$H_{p,\delta}(A) := \inf \left\{ \sum_{j=1}^{\infty} (\operatorname{diam}(B_j))^p : A \subset \bigcup_{j=1}^{\infty} B_j \text{ and } \operatorname{diam}(B_j) \leq \delta \right\}.$$

Recall and define

$$\operatorname{diam}(B) = \sup_{x,y \in B} \rho(x,y) \quad \text{ and } \quad \inf\{\emptyset\} = +\infty.$$

Now define the Hausdorff measure

$$H_p(A) := \lim_{\delta \to 0} H_{p,\delta}(A).$$

Remark 5. (1) If one requires the B_j's to be closed, the result is the same because

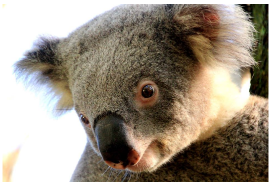
$$\operatorname{diam}(B_j) = \operatorname{diam}(\overline{B_j}).$$

(2) If one requires the B_j's to be open, the result is the same because we can replace B_j by
$$U_j = \{x \in X : \rho(x, B_j) < \epsilon \cdot 2^{-j-1}\} \quad \text{for } \epsilon > 0.$$
Thendiam$(U_j) \leq \mathrm{diam}(B_j) + \epsilon \cdot 2^{-j}$ and therefore
$$\sum_{j=1}^{\infty}(\mathrm{diam}(U_j))^p \leq \sum_{j=1}^{\infty}(\mathrm{diam}(B_j) + \epsilon \cdot 2^{-j})^p \leq \sum_{j=1}^{\infty}(\mathrm{diam}(B_j))^p + c(\epsilon),$$
where c depends only on p and ϵ. We have $c \to 0$ as $\epsilon \to 0$. Hence we get the same result in $H_{p,\delta}$.
(3) The intuition is the following: if $p \in \mathbb{N}$ and A is p-dimensional, then the amount of A contained in a region of diam $= r$" should be proportional to r^p. This is because a ball in p-dimensional space has volume proportional to r^p. We'll see more about this!
(4) We need to let the diameters $\to 0$ to capture irregularly shaped sets!

Example 7.4. Let $A_m := \{(x, \sin(mx)) : |x| \leq \pi\} \subset \mathbb{R}^2$.
It is $\mathrm{diam}(A_m) \leq (4 + 4\pi^2)^{\frac{1}{2}}$ for all m. If we didn't take $\delta \to 0$, we would cover A_m by A_m and measure ($p = 1$) would be bounded. We need $\delta << \frac{1}{m}$ before $H_{1,\delta}(A_m)$ actually measured the length, which diverges to infinity as $m \to \infty$.

Proposition 7.5. *H_p is a metric outer measure.*

Proof:
We will use the outer measure Proposition 1 (Outer Mass Existence). Observe that $H_{p,\delta}(\emptyset) = 0$, because $\mathrm{diam}(\emptyset) = 0$. Therefore as defined, $H_{p,\delta}$ is an outer measure. We now want to carry this to H_p.
If $\rho(A, B) > 0$ and $A \cup B \subset \bigcup C_j$ such that $\mathrm{diam}(C_j) \leq \delta < \rho(A, B)$ for all j, then
$$C_j \cap A \neq \emptyset \Rightarrow C_j \cap B = \emptyset, \qquad C_j \cap B \neq \emptyset \Rightarrow C_j \cap A = \emptyset$$
$$\Rightarrow \text{Let } D_j = C_j. \qquad \Rightarrow \text{Let } E_j = C_j.$$
We have $A \subset \bigcup D_j$, $B \subset \bigcup E_j$, and
$$\sum \mathrm{diam}(C_j)^p = \sum \mathrm{diam}(D_j)^p + \sum \mathrm{diam}(E_j)^p \geq H_{p,\delta}(A) + H_{p,\delta}(B)$$
for all $\delta < \rho(A, B)$. Now taking the infimum, we have
$$H_p(A \cup B) \geq H_p(A) + H_p(B).$$
Next we will show that H_p is also an outer measure. This will imply the inequality
$$H_p(A) + H_p(B) \geq H_p(A \cup B),$$
and so combined with the reverse inequality shows that H_p is then a metric outer measure.
(1) Note that $H_p(\emptyset) = 0$.
(2) If $A \subset B$ then $H_{p,\delta}(A) \leq H_{p,\delta}(B)$ for all $\delta > 0$, which implies $H_p(A) \leq H_p(B)$.
(3) We have
$$H_{p,\delta}(\bigcup A_j) \leq \sum H_{p,\delta}(A_j) \quad \forall \delta > 0.$$
On the right side we can simply replace $H_{p,\delta}(A_j)$ by $H_p(A_j)$ because $H_{p,\delta} \uparrow H_p$ as $\delta \downarrow 0$. So, for any $\delta > 0$
$$H_{p,\delta}(\bigcup A_j) \leq \sum H_p(A_j).$$
Letting $\delta \downarrow 0$ on the left, we now have the countable subadditivity (clever no?)
$$H_p(\bigcup A_j) \leq \sum H_p(A_j).$$
This completes the proof that H_p is an outer measure, and the above arguments also show that it is a metric outer measure.



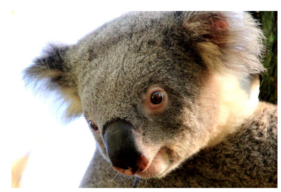
Proposition 7.6. *H_p is invariant under isometries. If $f, g : Y \to X$ satisfying $\rho(f(y), f(z)) \leq C_\rho(g(y), g(z))$ $\forall y, z \in Y$, then $H_p(f(A)) \leq C^p H_p(g(A))$.*

Proof:
Let $I : (X, \rho) \to X$ be an isometry. If $A \subset \underset{j\geq 1}{\cup} E_j$, $\text{diam}(E_j) \leq \delta$ $\forall j$ then $I(A) \leq \underset{j\geq 1}{\cup} I(E_j)$, $\text{diam}(I(E_j)) \leq \delta$ $\forall j$. This implies that $H_{p,\delta}(A) = H_{p,\delta}(I(A))$, taking $\delta \to 0$ then gives $H_p(A) = H_p(I(A))$.
Let $\varepsilon, \delta > 0$, $A \subset Y$. Assume $g(A) \subset \overset{\infty}{\underset{j=1}{\cup}} B_j$ such that $\text{diam}(B_j) \leq \frac{\delta}{C}\nu_j$ and $\sum\limits_{j\geq 1} \text{diam}(B_j)^p \leq H_p(g(A)) + \varepsilon$. Then for any $y \in A$, there exists j such that $g(y) \in B_j \Rightarrow y \in g^{-1}(B_j)$. Then $f(y) \in f(g^{-1}(B_j)) =: B_j$ therefore $f(A) \subset \overset{\infty}{\underset{j=1}{\cup}} B_j$.
By hypothesis, $\forall f(y), f(z) \in B_j$, $\rho(f(y), f(z)) \leq C_\rho(g(y), g(z))$. By the definition of B_j, $\exists u, v \in B_j$ such that $f(y) = f(g^{-1}(u)) \in f(g^{-1}(B_j))$ and $f(z) = f(g^{-1}(v)) \in f(g^{-1}(B_j))$. We have

$$C_\rho(u, v) = C_\rho(g(y), g(z)) \leq C\text{diam}(B_j) = \frac{C\delta}{C} = \delta$$

So $H_{p,\delta}(f(A)) \leq \sum\limits_{j\geq 1} (\text{diam}(B_j)^p$. We also have

$$(\text{diam}(B_j))^p \leq C^p(\text{diam}(B_j))^p \Rightarrow \sum_{j\geq 1}(\text{diam}(B_j))^p \leq C^p \sum_{j\geq 1}(\text{diam}(B_j))^p \leq C^p(H_p(g(A))) + \varepsilon$$

Since this is true for all ε, we have that $H_p(f(A)) \leq C^p H_p(g(A))$ 

Proposition 7.7 (H-Dimension Proposition)**.** *If $H_p(A) < \infty$, then $H_q(A) = 0$ $\forall q > p$. If $H_p(A) > 0$, then $H_q(A) = \infty$ $\forall q < p$.*

Proof:
For the first statement, assume $H_p(A) < \infty$. Then $\forall \delta > 0$, $\exists \{B_j\}_{j\geq 1}$ with $A \subset \overset{\infty}{\underset{j=1}{\cup}} B_j$, $\text{diam}(B_j) \leq \delta$ and $H_{p,\delta}(A) \leq \sum\limits_{j\geq 1} (\text{diam}(B_j))^p \leq H_p(A) + 1$. If $q > p$, then

$$H_{q,\delta}(A) \leq \sum_{j\geq 1}(\text{diam}(B_j))^q = \sum_{j\geq 1}(\text{diam}(B_j))^{p+q-p} \leq \sum_{j\geq 1}(\text{diam}(B_j))^p \delta^{q-p} \leq \delta^{q-p}(H_p(A) + 1) \to 0$$

as $\delta \to 0$. Hence $H_q(A) = 0$. Then second statement is the contrapositive argument, hence $H_q(A) > 0$ for some $q > p$ gives $H_p(A) = \infty$. 

Let $A \subset X$. By the above *H-Dimension Proposition*,

$$\inf\{p \geq 0 \mid H_p(A) = 0\} = \sup\{p \geq 0 \mid H_p(A) = \infty\}$$

This defines the *Hausdorff dimension* of A.

Theorem 7.8. *There exists a constant c_n such that $\mathcal{H}_n = c_n\mathcal{L}_n$*

To prove this will require some more work and theory.....

Definition 7.9. Let ν and μ be measures on $(X, \mathcal{M})$. Then ν is *absolutely continuous* with respect to μ and we write $\nu << \mu$ if $\nu(Y) = 0\ \forall Y \in \mathcal{M}$ with $\mu(Y) = 0$. We say that μ and ν are *mutually singular* and write $\mu\nu$ if there exists $E, F \in \mathcal{M}$ with $E \cap F = \emptyset$, $E \cup F = X$, $\mu(E) = 0$, $\nu(F) = 0$.

Proposition 7.10. *$H_n << \mathcal{L}_n$ and $\mathcal{L}_n << H_n$.*

Proof:
First, we consider $I = \prod \mathrm{I}a_i, b_i\mathrm{I}$, $l_i := b_i - a_i$. If any $l_i = 0$ let's WLOG assume that $l_i, ..., l_k$ are all non-zero and $l_{k+1} = ... = l_n = 0$. Then $\forall \varepsilon > 0$, we can cover an interval of length L by $\frac{L}{\varepsilon}$ balls (one-dimensional) of radius ε. Similarly, we can cover I by $\prod\limits_{i=1}^{k} \frac{l_i}{\varepsilon}$ balls of radius ε. It follows that

$$\forall \delta \le \varepsilon, H_{n,\delta}(I) \le \prod_{i=1}^{k} \frac{l_i}{\delta}(2\delta)^n = \delta^{n-k}2^n \prod_{i=1}^{k} l_i,$$
$$\delta \to 0 \Rightarrow H_n(I) = 0.$$

If $l_i = 0$ for all i, then I is either a point or the empty set which both have $H_n = 0$. Finally, if for all i, $l_i \neq 0$, then we can cover I by $\prod\limits_{i=1}^{n} \frac{l_i}{\varepsilon}$ balls of radius ε. Then

$$\forall \delta \ge \varepsilon, H_{n,\delta}(I) \le \prod_{i=1}^{n} \frac{l_i}{\delta}(2\delta)^n = 2^n\mathcal{L}_n(I)$$

If $\mathcal{L}_n(I) = 0$, then $H_{n,\delta} = 0$ which implies that $H_n(I) = 0$.
If $\mathcal{L}_n(A) = 0$, then $\exists \{I_j\}_{j\ge1}$ such that $A \subset \bigcup\limits_{j\ge1} I_j$ and, for a fixed $\varepsilon > 0$, $\sum\limits_{j\ge1} \mathcal{L}_n(I_j) < \frac{\varepsilon}{2^n}$. Then

$$H_n(A) \le \sum_{j\ge1} H_n(I_j) \le 2^n \sum_{j\ge1} \mathcal{L}_n(I_j) < \varepsilon$$

Hence $H_n(A) = 0$. Therefore $H_n << \mathcal{L}_n$.

Now, we want to proof the second statement of the proposition, i. e. $\mathcal{L}_n << \mathcal{H}_n$.
Therefore, let $A \subset \mathbb{R}^n$ such that $\mathcal{H}_n(A) = 0$, where $A \in \mathcal{B}$. Then, since $\mathcal{H}_{n,\delta} \le \mathcal{H}_n$,

$$\mathcal{H}_{n,\delta}(A) = 0\, \forall \delta > 0$$

$\Rightarrow \exists$ a sequence $\{B_j\}_{j\ge1}$, which is closed in $\mathbb{R}^n$, such that $A \subset \bigcup\limits_{j=1}^{\infty} B_j$ and $\sum\limits_{j\ge1} (\mathrm{diam}(B_j))^n < \varepsilon$, where $\varepsilon > 0$. Note that for $x \in B_j$, $\rho(x, y) \le \delta_j = \mathrm{diam}(B_j) \forall x \in B_j$. So we can fix $x_j \in B_j$, and we get $B_j \subseteq \bar{B}_{\delta_j}(x_j)$.
So we have

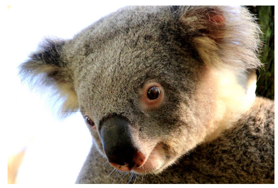
$$\mathcal{L}_n(B_j) \le \mathcal{L}_n(B_{\delta_j}(x_j)) = w_n\delta_j^n$$

where w_n = Vol $(B_1(0))$ denotes the volume of the unit ball with radius 1 (around zero). Alltogether, we get

$$\varepsilon > \sum_{j\ge1} \mathrm{diam}(B_j)^n = \sum_{j\ge1} \frac{\mathcal{L}_n(B_{\delta_j}(x_j))}{w_n} \ge \frac{1}{w_n} \sum_{j\ge1} \mathcal{L}_n(B_j)$$

and since $A \subset \bigcup\limits_{j=1}^{\infty}$ we get

$$\varepsilon > \frac{1}{w_n} \sum_{j \geq 1} \mathcal{L}_n(B_j) \geq \frac{1}{w_n} \mathcal{L}_n$$



Letting $\varepsilon \downarrow 0 \Rightarrow \mathcal{L}_n(A) = 0$.

Proposition 7.11. *The volume of the unit ball in $\mathbb{R}^n$ is*

$$w_n = Vol\ (B_1(0)) = \frac{2\pi^{\frac{n}{2}}}{n \cdot \Gamma(\frac{n}{2})}$$

Our goal is to compute

$$\int\limits_{S_1(0)} \int\limits_{0}^{1} r^{n-1}\ dr\, d\sigma$$

Proof:
We split the proof into four steps:

(i) First, we claim that

$$\int\limits_{\mathbb{R}^n} e^{-\pi |x|^2} dx = 1.$$

Note that $I_n = (I_1)^n$ by Fubini-Tonelli. Therefore, $I_n = (I_2)^{\frac{n}{2}}$ and

$$\begin{aligned} I_2 &= \int\limits_{0}^{2\pi} e^{-\pi r^2} r\ \mathrm{d}r\mathrm{d}\theta \\ &= 2\pi \cdot \int\limits_{0}^{\infty} e^{-\pi r^2} r\ \mathrm{d}r \\ &= 2\pi \cdot \left[\left(\frac{e^{-\pi r^2}}{-2\pi} \right) \right]_0^{\infty} \\ &= 1 \end{aligned}$$

$\Rightarrow I_1 = \sqrt{I_2} = 1$ and $I_k = 1 \forall k \in \mathbb{N}$.

(ii) Let

$$\Gamma(s) = \int\limits_{0}^{\infty} t^{s-1} e^{-t} \mathrm{d}t$$

for $s > 0$ (extends to $\mathbb{C}$). Then

$$\Gamma(s+1) = \int\limits_{0}^{\infty} t^s e^{-t} \mathrm{d}t = \left[-t^s e^{-t} \right] - \int\limits_{0}^{\infty} -e^{-t} s t^{s-1} dt = s \cdot \Gamma(s) \tag{7.1}$$

Applying equation 7.1 to the natural numbers, one gets

$$\Gamma(k) = (k-1)! \quad \forall k \in \mathbb{N}$$

and

$$\Gamma(k+\frac{1}{2}) = (k-\frac{1}{2})(k-\frac{3}{2})\cdots\Gamma(\frac{1}{2})$$

Question 1. *But what is $\Gamma(\frac{1}{2})$?*

By definition, first of all it is

$$\Gamma(\frac{1}{2}) = \int_0^\infty t^{\frac{1}{2}} e^{-t} \mathrm{d}t \tag{7.2}$$

Let $s = t^{\frac{1}{2}}$, $\mathrm{d}s = \frac{1}{2}t^{-\frac{1}{2}}\mathrm{d}t$, $t = s^2$, $2\mathrm{d}s = t^{-\frac{1}{2}}\mathrm{d}t$. Then we get $\Gamma(\frac{1}{2}) = \int_0^\infty e^{-s^2} 2\mathrm{d}s$. Let furthermore $u = \frac{s}{\sqrt{\pi}}$, $\sqrt{\pi}u = s$. Then, 7.2 can be simplified to

$$\begin{aligned}\Gamma(\frac{1}{2}) &= 2\int_0^\infty e^{-\pi u^2 \sqrt{\pi}}\mathrm{d}u \\ &= \int_{-\infty}^\infty e^{-\pi^2 u}\sqrt{\pi}\mathrm{d}u \\ &= \sqrt{\pi}I_1 = \sqrt{\pi}\end{aligned}$$

Therefore, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(iii) Now, we want to compute σ_n, the area of $S_1(0)$. First, we know

$$1 = \int_{\mathbb{R}^n} e^{-\pi|x|^2}\ dx = \int_{S_1(0)} \int_0^\infty e^{-\pi r^2} r^{n-1}\ \mathrm{d}r\mathrm{d}\sigma = \sigma_n \int_0^\infty e^{-\pi r^2} r^{n-1}\ \mathrm{d}r$$

Letting $s = r^2\pi$, one gets $\mathrm{d}s = 2r\pi dr$, and therefore

$$1 = \frac{\sigma_n}{2\pi}\int_0^\infty e^{-s}\left(\frac{s}{\pi}\right)^{\frac{n-1}{2}} \frac{ds}{(\frac{s}{2\pi})^{\frac{1}{2}}}$$

Since $\frac{s}{\pi}^{\frac{1}{2}} = r$, $\frac{\mathrm{d}s}{2\pi r} =\ \mathrm{d}r$.
Now, we have

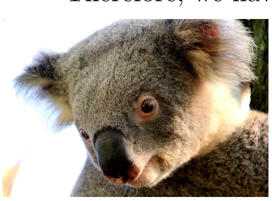
$$1 = \frac{\sigma_n}{2\pi\cdot\pi^{\frac{n}{2}-\frac{1}{2}}}\int_0^\infty e^{-s}s^{n/2-1}\mathrm{d}s = \frac{\sigma_n}{2\pi^{n/2}}\Gamma(\frac{n}{2})$$

$\Rightarrow \sigma_n = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}$

(iv) In this step, we want to compute w_n.

$$\int_{B_1(0)} dx = \mathrm{Vol}\ (B_1(0)) = \int_{S_1(0)}\int_0^1 r^{n-1}\ \mathrm{d}r\mathrm{d}\sigma = \sigma_n\int_0^1 r^{n-1}\ \mathrm{d}r = \left[\sigma_n\frac{r^n}{n}\right]_0^1 = \frac{\sigma_n}{n} = w_n$$

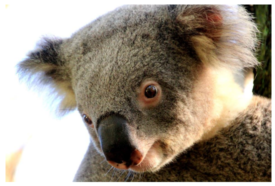
Therefore, we have $w_n = \frac{2\pi^{n/2}}{n\cdot\Gamma(\frac{n}{2})}$, which finishes our proof.



Corollary 7.12. *$\forall x \in \mathbb{R}^n$ and $r > 0$, the area of $S_r(x)$ is $r^{n-1}\sigma_n$ and Vol $(B_r(x)) = w_n r^n$.*

Proof:

$$\int_{S_r(x)} \mathrm{d}\sigma = \int_{S_r(0)} \mathrm{d}\sigma = \int_{S_1(0)} r^{n-1}\mathrm{d}\sigma = r^{n-1}\sigma_n$$



Analogously for $B_r(x)$.

Theorem 7.13. *The relationship between Hausdorff and Lebesgue measures in non-negative integer dimensions restricted to Borel sets is as follows.*

(i) $H_0 =$ *counting measure*
(ii) $H_1 = \mathcal{L}_1$
(iii) $H_n = \frac{2^n \mathcal{L}_n}{w_n}$ *for* $n \geq 2$

Proof:

(i) Exercise.

(ii) For $n = 1$, note that for any interval I, $L_1(I) = \mathrm{diam}(I)$. For any $\delta > 0$, $I = (a, b)$, where $(\cdot, \cdot)$ denotes an interval, which is either opened or closed, can be covered by intervals $(a + k\delta, a + (k+1)\delta)$. For $k = 0$, one gets $\frac{b-a}{\delta} = 1$, and therefore

$$H_{1,\delta}(I) \leq \frac{\delta \cdot (b-a)}{\delta} = b - a = \mathcal{L}_1(I)$$

$\Rightarrow H_1(I) \leq \mathcal{L}_1(I) = H_{1,b-a}(I) \leq H_1(I)$, which we get by covering I by I. $\Rightarrow H_1(I) = \mathcal{L}_1(I)$.
Since intervals generate $\mathcal{B}$, we get $\Rightarrow H_1|_{\mathcal{B}} = \mathcal{L}_1|_{\mathcal{B}}$, because H_1 and $\mathcal{L}_1$ are both measures on $\mathcal{B}$. Alternatively, apply pre-measure extension Theorem to $H_1|_{\mathcal{A}} = \mathcal{L}_1|_{\mathcal{A}}$. Since $\mathbb{R}$ is σ-finite, we also get $H_1 = \mathcal{L}_1$ on $\mathcal{B}$.

(iii) For $n \geq 2$, note that $\mathcal{H}_n(B_r) \geq \mathcal{H}_{n,r}(B_r) = (2r)^n = 2^n r^n = 2^n \mathcal{L}_n(B_r)\frac{1}{w_n}$
$\Rightarrow \frac{w_n \mathcal{H}_n(B_r)}{2^n} \geq \mathcal{L}_n(B_r)$ for any ball $B_r \subset \mathbb{R}^n$.
Let $\varepsilon > 0$. By definition of $\mathcal{L}_n^*$, $\exists$ symmetric intervals $\{R_j\}$ such that

$$\mathcal{L}_n(B_r) \geq \sum_j \mathcal{L}_n(R_j) - \varepsilon$$

where $B_r \subset \bigcup_{j\geq 1} R_j$, $\mathrm{diam}(R_j) \leq 2r$. Furthermore, $\exists \{A_j\}_{j\geq 1} \subset \mathcal{E}_n$ such that all are finite and $B_r \subset \bigcup_{j=1}^{n} A_j$, and since $A_j = \bigcup_{k=1}^{m_j} I_k^j$ each I_k^j can be choped into finitely many symmetric intervals. In particular, $B_r \subset \bigcup_{j=1}^{\infty} A_j$ and

$$\mathcal{L}_n(B_r) \geq \sum \mathcal{L}_n(A_j) - \varepsilon = \sum \mathcal{L}_n(R_j) - \varepsilon$$

Therefore, we get $\bigcup A_j = \bigcup R_j$.
If R_j has side length l_j, then $\mathrm{diam}(R_j) = (\sum_1^n l_j^2)^{\frac{1}{2}} = n^{\frac{1}{2}} l_j \Rightarrow \mathcal{L}_n(R_j) = l_j^n = \frac{\mathrm{diam}(R_j)^n}{n^{\frac{n}{2}}}$.

$$\mathcal{L}_n(B_r) \geq \sum (\frac{\mathrm{diam}(R_j)}{\sqrt{n}})^n - \varepsilon \geq \frac{H_{n,\delta}(B_r)}{n^{n/2}} - \varepsilon$$

Because we can make diam R_j as small as we like, we can let $\delta \downarrow 0$. If then also $\varepsilon \downarrow 0$, we get

$$\mathcal{L}_n(B_r) \geq \frac{H_n(B_r)}{n^{n/2}}$$

The same argument shows $\forall A \in \mathcal{B}$ with $\mathcal{L}_n(A) < \infty$ the following inequality

$$\mathcal{L}_n(A) \geq \frac{\mathcal{H}_n(A)}{n^{n/2}} \tag{7.3}$$

Let $\delta > 0$. Claim: $\exists$ balls $\{B_j^k\}$ with $\operatorname{diam} B_j^* \leq \delta$ such that $\mathcal{L}_n(R_j \setminus \bigcup_1^\infty B_j^k) = 0$ and $B_j^k \cap B_j^{k'} = \emptyset \ \forall k \neq k'$. Then

$$\mathcal{L}_n(B_r) + \varepsilon \geq \sum_{j,k} \mathcal{L}_n(B_j^k) = \frac{w_n}{2^n} \sum \operatorname{diam}(B_j^k)^n \geq \frac{w_n}{2^n} H_{n,\delta}(\bigcup B_j^k)$$

Claim and 7.3 let us conclude

$$\Rightarrow \mathcal{L}_n(\bigcup R_j \setminus \bigcup_{j,k} B_j^k) = 0 = \mathcal{H}_n(\bigcup R_j \setminus \bigcup B_j^k)$$

$\Rightarrow H_{n,\delta}(\bigcup R_j \setminus \bigcup B_j k^{)} = 0$ and therefore also $H_{n,\delta}(\bigcup B_j^k) + H_{n,\delta}(\bigcup R_j \setminus \bigcup B_j^k) = H_{n,\delta}(\bigcup R_j)$ $H_{n,\delta}(\bigcup B_j^k) = H_{n,\delta}(\bigcup R_j) \geq H_{n,\delta}(B_r)$. Therefore, we get

$$\mathcal{L}_n(B_r) + \varepsilon \geq \frac{w_n}{2^n} H_{n,\delta(\bigcup B_j^k)} \geq \frac{w_n}{2^n} H_{n,\delta}(B) \quad \forall \delta > 0 \quad \forall r > 0$$

$\delta, \varepsilon \downarrow 0 \Rightarrow L_n(B_r) \geq w_n/2^n H_n(B_r) \geq L_n(B_r) \Rightarrow L_n = H_n$ on balls $\Rightarrow$ generate on $\mathcal{B}$.

To complete the proof relating Hausdorff and Lebesgue measures on the Borel σ-algebra for dimensions in $\mathbb{N}$, we require the following, which may already be clear, but we include for completeness.

Claim 3. *For any interval $I \subset \mathbb{R}^n$, there exists a series $\{B_j\}_{j\geq 1}$ such that*

(1) Each B_j is a ball in I.
(2) It is $B_j \cap B_k = \emptyset$ for all $j \neq k$.
(3) We have $\mathcal{L}_n(I \setminus \bigcup B_j) = 0$ (and therefore $\mathcal{L}_n(I) = \mathcal{L}_n(\bigcup B_j)$).

Proof:
First note that $\mathcal{L}_n(I \setminus \mathring{I}) = 0$. So without loss of generality we can assume that I is open. For $x \in I$, there is $\delta \in \mathbb{Q}, \delta > 0$ such that $B_\delta(x) \subset I$. Also there exists $q \in \mathbb{Q}^n$ such that $|x - q| < \delta \cdot 10^{-6}$. This implies for every y with $|y - q| < (1 - 10^{-6})\delta$,

$$|y - x| \leq |y - q| + |x - q| < \delta \implies y \in B_\delta(x) \subset I.$$

So we have

$$B_1 := B_{(1-10^{-6})\delta}(q) \subset I.$$

For $N \geq 1$ and $x \in I$, it is either $x \in \bigcup_{k=1}^N \overline{B_k}$ or not. We are assuming $\{B_k\}^N \subset I$ are disjoint balls with rational radii and rational centers (centers are elements of $\mathbb{Q}^n$). If $x \in \bigcup_{k=1}^N \overline{B_k}$ we consider $x \in I \setminus \bigcup_{k=1}^N \overline{B_k}$. Note that this set is *open*. So, if there exists $x \in I \setminus \bigcup_{k=1}^N \overline{B_k}$, then the same argument shows that there is a new ball,

$$x \in \ B_{N+1} \subset I \setminus \bigcup_{k=1}^N \overline{B_k}$$

with the center and radius of B_{N+1} rational (same argument as above). Then we note further that the set of balls

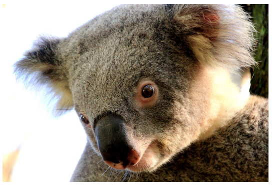
$$\{B_\delta(q) : \quad \delta \in \mathbb{Q}, \quad \text{ and } q \in \mathbb{Q}^n\}$$

is countable. Consequently, we require at most countably many of these balls to ensure that

$$I \subset \bigcup_{k=1}^\infty \overline{B_k} \ \text{ and } \ \mathcal{L}_n(\overline{B_k} \setminus B_k) = 0 \text{ for all } k \ \Rightarrow \ \mathcal{L}_n(\bigcup(\overline{B_k} \setminus B_k)) = 0.$$

So we get

$$\mathcal{L}_n(I) = \mathcal{L}_n(I \cap \bigcup B_k) + \mathcal{L}_n(I \setminus \bigcup B_k) = \mathcal{L}_n(\bigcup B_k) + \mathcal{L}_n(\bigcup \overline{B_k} \setminus B_k) = \mathcal{L}_n(\bigcup B_k).$$



8. Self-similarity and Hausdorff dimension

Definition 8.1. For $r > 0$, a similitude with scaling factor r is a map $S : \mathbb{R}^n \to \mathbb{R}^n$ of the form

$$S(x) = r\mathcal{O}(x) + b,$$

where $\mathcal{O}$ is an orthogonal transformation (rotation, reflection, or composition of these), and $b \in \mathbb{R}^n$. If $S = (S_1, \cdots S_m)$ is a family of similitudes with common scaling factor $r < 1$, for $E \subset \mathbb{R}^n$ we define

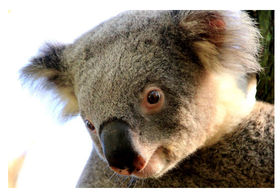
$$S^0(E) = E, \qquad S(E) = \bigcup_{j=1}^{m} S_j(E), \qquad S^k(E) = S(S^{k-1}(E)) \text{ for } k > 1.$$

We say that E is invariant under S if $S(E) = E$.

Lemma 8.2. *If $S(E) = E$, then $S^k(E) = E$ for all $k \geq 0$.*

Proof:
It is $S(E) = \bigcup_{j=1}^m S_j(E) = E$ and also

$$S^2(E) = \bigcup_{j=1}^{m} S_j \left(\bigcup_{j=1}^{m} S_j(E) \right) = \bigcup_{j=1}^{m} S_j(E) = E.$$

By induction we have $S^k(E) = E$ for $k \geq 2$. 
But what does that mean?

Well, the scaling factor r is less than one, so applying each S_j spins/flips/shrinks and slides E. Hence E looks like, for each k, m^k copies of itself which are scaled down by a factor of r^k. If these copies are disjoint or have little (negligible) overlap, E is "self-similar".

Example 8.3. Let $\beta \in (0,1)$ and $\mathrm{I}_0 = [0,1]$. Now define

$$\beta(a,b) = \left(\frac{a+b}{2} - \beta \left(\frac{b-a}{2} \right), \frac{a+b}{2} + \beta \left(\frac{b-a}{2} \right) \right).$$

Let $I_1 := I_0 \setminus \beta \mathring{I}_0$. This is closed and the union of two intervals, written $I_1 = \bigcup_{j=1}^2 I_j^1$. Then we define

$$I_2 := \bigcup_{j=1}^{2} I_j^1 \setminus \beta \mathring{I}_j^1,$$

which is a union of two disjoint unions of two closed intervals. Again we write $I_2 = \bigcup_{j=1}^4 I_j^2$. In general we write and define

$$I_k = \bigcup_{j=1}^{2^k} I_j^k \quad \text{and} \quad I_{k+1} := \bigcup_{j=1}^{2^k} I_j^k \setminus \beta \mathring{I}_j^k$$

As defined note that

$$I_0 \supset I_1 \supset \ldots \supset I_k \supset I_{k+1}$$

are a sequence of nested compact sets in $\mathbb{R}$ which is complete. Consequently,

$$\bigcap I_k = \lim_{k\to\infty} I_k =: C_\beta$$

is a compact subset of $\mathbb{R}$. Note that

$$\mathcal{L}_1(I_0) = 1, \quad \mathcal{L}_1(I_1) = (1-\beta)\mathcal{L}_1(I_0), \quad \mathcal{L}_1(I_{k+1}) = (1-\beta)\mathcal{L}_1(I_k),$$

and so

$$\mathcal{L}_1(C_\beta) = \lim_{k\to\infty}(1-\beta)^k = 0,$$

since $\beta \in (0,1)$. Note that more generally, one can let β vary at each step, so that

$$I_1 = I_0 \setminus \beta_0 I_0 = \bigcup_{j=1}^{2} I_j^k,$$

and in general

$$I_{k+1} = \bigcup_{j=1}^{2^k} I_j^k \setminus \beta_k \mathring{I}_j^k.$$

Similarly we have nested compact sets and so

$$C := \lim_{k\to\infty} I_k \text{ is a compact subset of } \mathbb{R}.$$

This is known as a *generalized Cantor set.* The Lebesgue measure

$$\mathcal{L}_1(C) = \prod_{k\geq 0}(1-\beta_k).$$

Now, let's see that when the scale factor β is constant, the Cantor set C_β is invariant under a similitude. Let

$$S := (S_1, S_2), \quad S_1(x) := \beta x, \quad S_2(x) = \beta x + (1-\beta).$$

Then we compute

$$S(I_0) = S_1(I_0) \cup S_2(I_0) = [0,\beta] \cup [1-\beta, 1] = I_1.$$

Analogously we have

$$S(I_1) = I_2 = S^2(I_0), \quad I_{k+1} = S^{k+1}(I_0).$$

So, since each S_i is continuous we have

$$S(\lim_{k\to\infty} S^k(I_0)) = S(C_\beta) = \lim_{k\to\infty} S^{k+1}(I_0) = C_\beta.$$

Consequently we see that C_β is invariant under the family of similitudes $S = (S_1, S_2)$.

Lemma 8.4. *Let $A \subset \mathbb{R}^n$. Then we have $\dim_H(A) \leq n$. More generally, if $A \subset B$, then* $\dim(A) \leq \dim(B)$.

Proof:
If $A \subset B$, and $H_p(B) = 0$, then $H_p(A) = 0$. Therefore

$$\dim(B) = \inf\{p \geq 0 | H_p(B) = 0\} \geq \inf\{p \geq 0 | H_p(A) = 0\} = \dim(A)$$

$\Rightarrow \dim(A) \leq \dim(\mathbb{R}^n)$. We can write the euclidian space $\mathbb{R}^n$ as $\mathbb{R}^n = \bigcup_{m\geq 1} B_m$, where B_m are balls of radius m centered at the origin. For $p < n$,

$$H_{p,\delta}(B_m) = \inf\{\sum_j \operatorname{diam}(E_j)^p | B_m \subset \bigcup_j E_j \text{ with } \operatorname{diam}(E_j) \leq \delta\}$$

We have proven that

$$H_n(B_m) = c_n \mathcal{L}_n = c_n m^n w_n.$$

We can conclude, that $\forall\varepsilon > 0, \exists\,\delta' > 0$ and $\{E_j\}_{j\geq 1} \subset \mathbb{R}^n$ such that $\mathrm{diam}(E_j) \leq \delta \leq \delta'$, $B_m \subset \bigcup_j E_j$ and $\sum_j \mathrm{diam}(E_j)^n \geq H_n(B_m) - \varepsilon$.

W.L.O.G., we may assume $\delta' < 1$. Let $\delta_j := \mathrm{diam}(E_j)$. Then we have $\delta_j \leq \delta \leq \delta' < 1$. Let $q := n - p$. Then $\frac{1}{\delta_j} > \frac{1}{\delta} > 1$ and therefore

$$\frac{1}{\delta_j^q} \geq \frac{1}{\delta^q}$$

Thus, we get

$$\sum_j \delta_j^p = \sum_j \delta_j^{n-q} = \sum_j \delta_j^n \frac{1}{\delta_j^q} \geq \frac{1}{\delta^q} \sum_j \delta_j^n = \frac{1}{\delta^q} \sum_j \mathrm{diam}(E_j)^n \geq \frac{1}{\delta^q}(H_n(B_m) - \varepsilon)$$

The RHS tends to ∞ as $\delta \downarrow 0 \Rightarrow H_p(B_m) = \infty$.

$\Rightarrow \forall p < n$, $H_p(B_m) = \infty$ and $H_n(B_m) = c_n m^n w_n < \infty$. Hence, $\sup\{p \geq 0 | H_p(B_m) = \infty\} \leq n$. For $p > n$, we can show that $H_p(B_n) = 0$ using the same argument. We have

$$\sum_j \delta_j^p = \sum_j \delta_j^{n+p-n} \leq \sum_j \delta_j^n \delta_j^{p-n} \leq \sum_j \delta_j^n \delta^{p-n},$$

since each $\delta_j \leq \delta$. Pulling the constant factor δ^{p-n} out front, we have

$$\sum_j \delta_j^p \leq \delta^{p-n} \sum_j \delta_j^n \leq \delta^{p-n} H_n(B_m) \to 0 \text{ as } \delta \downarrow 0.$$

Homework: $\dim(\mathbb{R}^n) = \sup \dim(B_m) = \sup\{n\} = n$. (Proof given in remarks below).

Remark 6.

(i) (PK) One can also use $H_p(A) \geq H_q(A)$ for $q \geq p$ to simplify the proof.

(ii) Since we have already shown

$$\dim(\bigcup_j E_j) \geq \dim(E_j) \quad \forall j \in \mathbb{N}$$

we get

$$\dim(\bigcup_j E_j) \geq \sup_j \dim E_j$$

Now, if $q > \sup_j \dim(E_j)$, then we get the following chain of implications

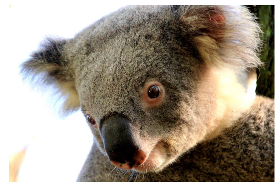
$$\begin{aligned} H_q(E_j) = 0\,\forall j \Rightarrow H_q(\bigcup_j E_j) &\leq \sum_j H_q(E_j) = 0 \\ \Rightarrow \dim(\bigcup_j E_j) &< q \quad \forall q > \sup_j \dim(E_j) \end{aligned}$$

So the Hausdorff-measure is logical because it sees the maximal "fatness" of the union of sets.

8.1. **Subsets of $\mathbb{R}^n$.**

Lemma 8.5. *Let $E \subset \mathbb{R}^n$ such that $\dim(E) < n$. Then $\mathring{E} = \emptyset$.*

Proof: If $\overset{\circ}{E} \neq \emptyset$, then there $\exists r > 0$ and $x \in E$ such that $B_r(x) \subset E$. $\Rightarrow \dim(E) \geq \dim(B_r(x)) = n$

So we get $n \geq \dim E \geq n \Rightarrow \dim E = n$. 

Lemma 8.6. *Let $E \subset \mathbb{R}^n$. If $\dim(E) > 0$, then E is uncountable.*

Proof: If E is countable, then $E = \bigcup_j e_j$, where $e_j \in \mathbb{R}^n$ is a point. Then we get

$$0 \leq \dim(E) = \sup \dim(\{e_j\}) = \sup(0) = 0$$

Since $\{e_j\} \subset B_\delta(e_j)\, \forall \delta > 0$, $\forall p > 0\, H_{p,\delta}(\{e_j\}) \leq (2\delta)^p$, which tends to 0 as $\delta \downarrow 0$.

Therefore, $H_p(\{e_j\}) = 0 \forall p > 0$. 

Remark 7. The Hausdorff Dimension of a subset $E \subset \mathbb{R}^n$ is the same if we consider E as a subset of $\mathbb{R}^m$ for any $m \geq n$ via the canonical embedding $(\mathbb{R}^n \mapsto \mathbb{R}^n \times \{0\}$. In this sense, if we have a set E which naturally lives in k-dimensions, if we view the set E as living in 10 zillion dimensions, the Hausdorff dimension of E remains the same. This is simply because the Hausdorff dimension, which is determined by the Hausdorff (outer) measure is defined in terms of *diameter,* and the diameter of sets does not change if we embed the sets into higher dimensional Euclidean space. That is another reason the Hausdorff dimension is "a good notion of dimension," because it is invariant of the ambient space.

Similitudes are finite families of maps of the form $r \cdot O(x) + b$, where $O(x)$ is an orthogonal transformation, and b is a vector in $\mathbb{R}^n$. These are therefore affine linear maps. We would like to understand how similitudes and invariant sets under similitudes relate to Hausdorff measure which motivates the following.

Proposition 8.7. *If $k \leq n$, $A \subset \mathbb{R}^k$ and $T\colon \mathbb{R}^k \to \mathbb{R}^n$ is an affine linear map, then $H_k(T(A)) = \mathcal{T}(T) H_k(A)$, where $\mathcal{T}(T) = \sqrt{\det(M^* M)}$, $Tx = Mx + b$ and $M^* = M^T$.*

Proof: First note that H_k is translation invariant because $H_k(A+b) = H_k(A)$ since $A \subset \bigcup_j E_j \Leftrightarrow A + b \subset \bigcup (E_j + b)$ and $\operatorname{diam}(E_j) = \operatorname{diam}(E_j + b)$. If $n = k$, then

$$H_n(T(A)) = c_n \mathcal{L}_n(T(A)) = c_n \int_{T(A)} d\mathcal{L}_n = c_n \int_A \mathcal{T}(T) d\mathcal{L}_n$$

(Caused by the translation invariance, we may assume $b = 0$.)

$$= c_n \mathcal{T}(T) \mathcal{L}_n(A) = \mathcal{T}(T) H_n(A)$$

If $k < n$, then let R be an isometry of $\mathbb{R}^n$, such that

$$R\colon T(\mathbb{R}^k) \to \mathbb{R}^k \times \{0\} = \{y \in \mathbb{R}^n | y = \sum < y_j e_j, y_j = 0 \forall j > k\}$$

Note that $Tx = Mx + b$. Since M is $k \times n$, Mx is a linear combination of the columns, and these columns are all contained in $\mathbb{R}^k$. Consequently the span of the columns has dimension at most k, and therefore the image $M\mathbb{R}^k + b$ has dimension at most k. For this reason there exists an isometry R of $\mathbb{R}^n$ (a change of coordinates composed with a translation) which maps $T(\mathbb{R}^k)$

to the canonical embedding of $\mathbb{R}^k$ in $\mathbb{R}^n$. Now, to reduce to the case in which we map between the same dimensional Euclidean space, let $\Phi\colon \mathbb{R}^n \to \mathbb{R}^k$ be the orthogonal projection,

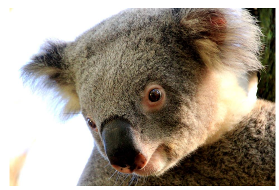
$$\Phi(\sum_i^n y_i e_i) = \sum_i^k y_i e_i$$

and let $S := \Phi \circ R \circ T\colon \mathbb{R}^k \to \mathbb{R}^k$.
Note that we assumed $b = 0$. Therefore, $Tx = Mx \Rightarrow T\mathbb{R}^k = M\mathbb{R}^k$.
We can write $R = Ux$, where U is unitary. Thus, $\Phi \cong [\delta_{ij}]_{i=1,\dots,k;j=1,\dots,n}$.
By the first case,

$$H_k(S(A)) = \mathcal{T}(S) H_k(A)$$

and

$$\begin{aligned}\mathcal{T}(S) &= \sqrt{\det(S^*S)} = \sqrt{\det(\Phi \circ R \circ T)^*(\Phi \circ R \circ T)} \\ &= \sqrt{\det(M^*U^*[\delta_{ij}]^T[\delta_{ij}]UM)} = \sqrt{\det M^*U^*UM} = \sqrt{\det M^*M} \\ &= \mathcal{T}(T)\end{aligned}$$



We have used the fact that U is unitary and therefore $U * U = I$.

8.2. **Ingredients: Useful analytic tools.**

Lemma 8.8 (Urysohn-light)**.** *Let (X, ρ) be a complete metric space and $A, B \subset X$ non-empty, closed sets with $A \cap B = \emptyset$. Assume that either A and B are both compact or that A and B are at a positive distance apart. Then $\exists f \in C(X)$ s.t.*

$$f|_A = 0 \quad f|_B = 1.$$

Proof: First we know that the distance between A and B is finite because $\exists a \in A$, $b \in B \quad \rho(A, B) \leqslant \rho(a, b) < \infty$.

In the case that A and B are compact, if they were at a distance of zero, then we would have sequences $\rho(a_n, b_n) \to 0$ for $\{a_n\} \subset A$, $\{b_n\} \subset B \Rightarrow \rho(a_n, B) \to 0$. By compactness which implies sequential compactness, we may assume without loss of generality that $a_n \to a \in A$ and $b_n \to b \in B$. Then by the triangle inequality,

$$\rho(a_n, b) \leq \rho(a_n, b_n) + \rho(b_n, b) \to 0 \implies a_n \to b \implies a = b \in A \cap B,$$

which contradicts the assumption that A and B are disjoint. So in all cases there exists $\delta > 0$ such that $\rho(A, B) = \delta$.
Let

$$U_r := \{x \in X |\ \rho(x, B) \geqslant (1 - r)\delta\}, \quad r \in (0, 1), \quad U_1 := X,$$

and

$$f(x) := \inf\{r \in (0, 1] |\ x \in U_r\}.$$

Note that $f(x)$ is well defined because it's an infimum and defined $\forall x \in X$ since every $x \in U_1$.
If $x \in B$, then $\rho(x, B) = 0$
$\Rightarrow \forall r \in (0, 1)$ is $\rho(x, B) < (1 - r)\delta$
$\Rightarrow f(x) = 1$ because $x \in U_1 = X$ but not in U_r $\forall r \in (0, 1)$.

If $x \in A$, then $\rho(x, B) \geqslant \delta$
$\Rightarrow \forall r > 0$, is $\rho(x, B) \geqslant (1 - r)\delta \ \Rightarrow x \in U_r\ \forall r \in (0, 1]$.
$\Rightarrow f(x) = 0$.

We need to show now that $f \in C(X)$. Let $x \in X$. If $x_n \to x$, then note that this means $\rho(x_n, x) \to 0$. We estimate

$$\rho(x, B) \le \rho(x, x_n) + \rho(x_n, B), \quad \rho(x_n, B) \le \rho(x_n, x) + \rho(x, B),$$

and therefore

$$|\rho(x_n, B) - \rho(x, B)| \le \rho(x_n, x) \to 0 \text{ as } n \to \infty.$$

It follows that $\rho(x_n, B) \to \rho(x, B)$ as $n \to \infty$. Since B is closed, $\rho(x, B) = 0 \iff x \in B$. In this case, since $x_n \to x$,

$$\rho(x_n, B) \le \rho(x_n, x) \to 0 \implies \text{ for } r \in (0,1)$$

there exists $N \in \mathbb{N}$ such that for any $n \ge N$,

$$\rho(x_n, x) < (1-r)\delta \implies \forall r' < r,$$

$$\rho(x_n, x) < (1-r)\delta < (1-r')\delta \implies x_n \notin U_r'$$

and therefore

$$f(x_n) \ge r \forall n \ge N.$$

Letting $r \to 1$ shows that $f(x_n) \to 1 = f(x)$. If $x \notin B$, then first we assume $\rho(x, B) \ge \delta \implies f(x) = 0$. Since

$$\rho(x_n, B) \to \rho(x, B),$$

for every $r \in (0,1)$ there is $N \in \mathbb{N}$ such that for all $n \ge N$,

$$\rho(x_n, B) > (1-r)\delta \implies f(x_n) \le r.$$

We can let $r \to 0$ which shows that $f(x_n) \to 0 = f(x)$. If on the other hand $\rho(x, B) < \delta$, then there is some $r \in (0,1)$ such that

$$\rho(x, B) = (1-r)\delta.$$

It follows from the definition of f that

$$f(x) = r.$$

Without loss of generality we may assume that, since $\rho(x_n, B) \to \rho(x, B)$ we have $\rho(x_n, B) = r_n < \delta \ \forall n$. Consequently, we also have

$$f(x_n) = r_n \to r = f(x).$$

Therefore, for any sequence $x_n \to x \in X$, we have shown $f(x_n) \to f(x)$ and f is consequently continuous. 

We require Urysohn's Lemma (at least on metric spaces; it holds in the more general setting of a normal topological space under the assumption that the sets are closed and disjoint) to prove one of Riesz's Representation Theorems.

Theorem 8.9 (Riesz Representation for $C_c(X)'$)**.** *If* $0 \le I \in C_c(X)' \Rightarrow \exists$ *measure* μ *on* X *s.t.*

$$I(f) = \int_X f \, d\mu$$

and Borel sets are μ *measurable.*

Proof:
Write $f \prec U$ if U open, $f \in C_c(X)$, $0 \leqslant f \leqslant 1$, and $supp(f) \subset U$.
(Rmk: $f \in C_c(X) \Leftrightarrow \ supp(f) \Subset X$)
$\mu(U) := \sup\{I(f)|\ f \prec U\}$, $\mu(\emptyset) := 0$ defined for U open.
This is a non-negative set-function with $\mu(\emptyset) = 0$
$\overset{make\ an}{\Rightarrow}$ outer measure.
$\mu^*(E) := \inf\{\mu(U)|\ E \subset U,\ U \text{ open}\}$
Note: Urysohn gives existence of such f
$\rho(supp(f), U^c) > 0 \ \Rightarrow \exists f \equiv 1$ on $K \Subset U$, $f \equiv 0$ on U^c.

If $\rho(A,B) > 0$ then $\mu^*(A \cup B) = \inf\{\mu(U)|\ A \cup B \subset U, U \text{ open}\}$

$$\begin{aligned}\mu(U) &= \sup\{I(f)|\ f \prec U \supset A \cup B\}\\ &\geqslant \sup\{I(f+g) = I(f) + I(g)|\ f \prec V \supset A,\ g \prec W \supset B\ V \cap W = \emptyset\}\\ &= \mu(V) + \mu(W)\end{aligned}$$

Taking the infimum over V and W which contain A and B respectively,
$\Rightarrow \mu(U) \geqslant \inf\{\mu(V)|V \text{ open},\ A \subset V\} + \inf\{\mu(W)|W \text{ open},\ B \subset W\} = \mu^*(A) + \mu^*(B)$.
Next taking the infimum over U, $\overset{\inf\ over\ U}{\Rightarrow}$ $\mu^*(A \cup B) \geqslant \mu^*(A) + \mu^*(B) \geqslant \mu^*(A \cup B)$
$\Rightarrow \mu^*$ is a metric outer measure.
We have proven that all $\mathcal{B}$ are measurable $\forall$ metric outer measure.
$\Rightarrow \mathcal{B}$ is μ^* measurable

To show $I(f) = \int f\ d\mu\ \forall f \in C_c(X)$, we first show

$$\mu(K) = \inf\{I(f)|\ f \in C_c(X),\ f \geqslant \chi_K\} \quad \forall K \Subset X.$$

(Note: $\int \chi_K\ d\mu = \mu(K)$ by def.)

Let $U_\varepsilon := \{x|\ f(x) > 1 - \varepsilon\}$ for such an $f \in C_c(X)$, $f \geqslant \chi_K$. U_ε is open.
If $g \prec U_\varepsilon \ \Rightarrow (1-\varepsilon)^{-1} f - g \geqslant 0 \ \Rightarrow I\left((1-\varepsilon)^{-1} f - g\right) \geqslant 0$
$\Rightarrow (1-\varepsilon)^{-1} I(f) \geqslant I(g)$
$\Rightarrow \mu(K) \underset{K \subset U_\varepsilon}{\leqslant} \mu(U_\varepsilon) \underset{\inf\ over\ g}{\leqslant} (1-\varepsilon)^{-1} I(f)$
$\overset{\varepsilon \downarrow 0}{\Rightarrow} \mu(K) \leqslant I(f)$
On the other hand for U open with $U \supset K$, by Urysohn

$$\exists f \in C_c(X) \text{ s.t. } f \geqslant \chi_K \text{ and } f \prec U$$

$\Rightarrow I(f) \leqslant \mu(U)$ (by def. of μ).
$\mu(K) = \inf\{\mu(U)|\ U \supset K, U \text{ open}\}$
$\Rightarrow \mu(K) \leqslant I(f) \leqslant \mu(U) \quad \forall U \text{ open } U \supset K$
$\overset{\inf\ on\ RHS}{\Rightarrow} \mu(K) \leqslant I(f) \leqslant \mu(K)$
$\Rightarrow \mu(K) = \inf\{I(f)|\ f \in C_c(X),\ f \geqslant \chi_K\} \quad \forall K \subset X$.

It is therefore enough to show

$$I(f) = \int f\ d\mu \text{ for } f \in C_c(X, [0,1))$$

since C_c is the linear span of such f, and both I and the integral $\int\ d\mu$ are linear functionals on C_c.

For $N \in \mathbb{N}$, $1 \leqslant j \leqslant N$ let $K_j := \{x|\ f(x) \geqslant \frac{j}{N}\}$ and $K_0 := supp(f)$.
Then note that

$$K_0 \supset K_1 \supset K_2 \supset \ldots.$$

Define

$$f_j(x) := \begin{cases} 0 & \text{if } x \notin K_{j-1} \\ f(x) - \frac{(j-1)}{N} & \text{if } x \in K_{j-1} \setminus K_j \\ \frac{1}{N} & \text{if } x \in K_j \end{cases}$$

So defined, f_j vanishes on K_{j-1}^c, and on K_j, $f_j = \frac{1}{N}$, whereas on $K_{j-1} \setminus K_j$, since

$$\frac{j-1}{N} \le f < \frac{j}{N} \implies 0 < f_j < 1/N.$$

Consequently,

$$N^{-1}\chi_{K_j} \leqslant f_j \leqslant N^{-1}\chi_{K_{j-1}} \tag{8.1}$$

$$\Rightarrow \frac{1}{N}\mu(K_j) \leqslant \int f_j \, d\mu \leqslant \frac{1}{N}\mu(K_{j-1}). \tag{8.2}$$

If U is open and $U \supset K_{j-1}$, then

$$N f_j \prec U,$$

because the support of f_j is K_{j-1} which is compactly contained in U. Therefore, by the definition of $\mu(U)$ as the supremum over all such f_j, we have

$$I(f_j) \leqslant N^{-1}\mu(U).$$

Now since for a compact set (which we note K_j is) we showed that $\mu(K_j)$ is the infimum over $I(f)$ for all $f \in C_c$ with $f \ge \chi_{K_j}$, by (8.1)

$$\frac{1}{N}\mu(K_j) \le I(f_j) \le N^{-1}\mu(U).$$

Taking the infimum over all open U which contain K_{j-1} as in the definition of μ we then have

$$\frac{1}{N}\mu(K_j) \le I(f_j) \le \frac{1}{N}\mu(K_{j-1}). \tag{8.3}$$

Note that so defined

$$f = \sum_{j=1}^{N} f_j,$$

so summing over (8.2) by linearity of the integral,

$$\Rightarrow \frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leqslant \sum_{j=1}^{N} I(f_j) \leqslant \frac{1}{N}\sum_{j=0}^{N-1}\mu(K_j).$$

Next we sum over (8.3) using the linearity of the functional I,

$$\frac{1}{N}\sum_{j=1}^{N}\mu(K_j) \leqslant \int f \, d\mu \leqslant \frac{1}{N}\sum_{j=0}^{N-1}\mu(K_j).$$

Finally, we subtract these inequalities which leaves only the first and last terms, and so

$$\Rightarrow |I(f) - \int f \, d\mu| \leqslant \frac{\mu(K_0) - \mu(K_N)}{N} \le \frac{\mu(supp(f))}{N} \to 0, \text{ as } N \to \infty.$$

Note that the measure of the support of f is finite because the support is compact, and for compact sets, $\mu(K)$ is defined as the infimum of $I(f)$, and I is a linear functional (which implies I is continuous and hence has bounded norm). Therefore we have $I(f) = \int f \, d\mu$.

Proposition 8.10. *Let S be a family of similarities with common scaling factor $r \in (0,1)$. If there exists U open, non-empty and bounded such that $S(U) \subset U$, then there exists a unique $X \subset\subset \mathbb{R}^n$ such that $S(X) = X \neq \emptyset$. More generally, if there exists $X \subset\subset \mathbb{R}^n$ such that $S(X) = X$, $X \neq \emptyset$, then it is unique.*

Proof:
$\bar{U}$ is compact (BWHB). Each S_i is an affine linear function from $\mathbb{R}^n \to \mathbb{R}^n$.
$S_i(x) = rO_i(x) + b_i$, so $S_i(\bar{U})$ is compact. By continuity, $S_i(\bar{U}) \subset \bar{S_i(U)}$ which implies that $\bigcup_{i=1}^{\infty} S_i(\bar{U}) = S(\bar{U}) \subset \bar{S(U)} \subset \bar{U}$. So $S(\bar{U})$ is compact.
For $X := \bigcap_{k\geq 0} S^k(\bar{U})$, $S^2(\bar{U}) \subset S(\bar{U})$ and $S^k(\bar{U}) \subset S^{k-1}(\bar{U})$. $S^k(\bar{U})$ is compact and non-empty which implies that $X \neq \emptyset$ and compact.

$$\begin{aligned} X &= \lim_{k\to\infty} S^k(\bar{U}) \\ &= \lim_{k\to\infty} S^{k+1}(\bar{U}) \\ &= S(\lim_{k\to\infty} S^k(\bar{U})) \\ &= S(X) \end{aligned}$$

Therefore, X is invariant.

If $Y \neq \emptyset$ is compact, and $S(Y) = Y$, we wish to show that $Y = X$. We have,

$$\begin{aligned} d(Y,X) :=& \sup_{y\in Y} \rho(y,X) \\ \Rightarrow d(S_i(Y), S_i(X)) =& d(rO_i(Y) + b_i, rO_i(X) + b_i) = rd(Y,X) \\ Y =& \bigcup_{i=1}^{m} S_i(Y) = S(Y) \\ \Rightarrow d(Y,X) =& \max_{1\leq i\leq m} d(S_iY, X) = d(S_jY,X) \text{ for some } j \in \{1,...,m\}. \end{aligned}$$

For fixed

$$y \in Y,\ \rho(S_jy, X) = \inf_{x\in X, 1\leq k\leq m} \rho(S_jy, S_kx) \leq \inf_{x\in X} \rho(S_jy, S_jx) = \rho(S_jy, S_jX).$$

Taking the supremum over $y \in Y$, we have

$$d(S_jY, X) \leq d(S_jY, S_jX) = rd(Y,X).$$

Since $r < 1$ this is only possible if

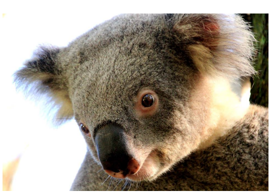
$$d(Y,X) = 0 \Rightarrow \sup_{y\in Y} \rho(y,X) = 0 \Rightarrow \rho(y,X) = 0, \quad \forall y \in Y.$$

The same argument shows that $d(X,Y) = 0$. By compactness of X and Y, which implies sequential compactness,

$$\rho(y,X) = 0 \Rightarrow y \in X, \forall y \in Y \Rightarrow Y \subset X$$

and similarly,

$$\rho(x,Y) = 0 \Rightarrow x \in Y \forall x \in X \Rightarrow X \subset Y \Rightarrow Y = X.$$



Definition 8.11. For $x \in \mathbb{R}^n$, $E \subset \mathbb{R}^n$, a measure μ, $\{i_1, ..., i_k\} \subset \{1, ..., m\}$ we define
(1) $x_{i_1..i_k} := S_{i_1} \circ ... \circ S_{i_k}(x)$,
(2) $E_{i_1..i_k} := S_{i_1} \circ ... \circ S_{i_k}(E)$, and

(3) $\mu_{i_1..i_k} := \mu((S_{i_1} \circ ... \circ S_{i_k})^{-1}(E)$.

Theorem 8.12. *$S = (S_1, ..., S_m)$ is a family of similitudes with common scaling factor $r \in (0,1)$ $X \subset\subset \mathbb{R}^n$, $X \neq \emptyset$, $S(X) = X$. Then there exists a Borel measure μ on $\mathbb{R}^n$ such that $\mu(\mathbb{R}^n) = 1$, $supp(\mu) = X$, and*

$$\forall k \in \mathbb{N}, \quad \mu \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} \mu_{i_1..i_k}.$$

Proof:
We will construct μ on X extend it to $\mathbb{R}^n \setminus X$ to be 0. Let $x \in X$,

$$\delta_x(E) := \begin{cases} 1, & x \in E \\ 0, & x \notin E \end{cases}.$$

For $\{E_j\}_{j\geq 1}$ disjoint then either there exists i, j such that

$$x \in E_j \Rightarrow \delta_x(\underset{j\geq 1}{\cup} E_j) = 1 = \sum_{j\geq 1} \delta_x(E_j),$$

or not; in which case

$$\delta_x(\underset{j\geq 1}{\cup} E_j) = 0 = \sum_{j\geq 1} \delta_x(E_j).$$

Consequently, we have for any $A, B \in \mathbb{R}^n$, $\delta_x(A) = \delta_x(A \cap B) + \delta_x(A \setminus B)$. This shows that every set in $\mathbb{R}^n$ is measurable for δ_x.
We define

$$\mu^k := \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} [\delta_x]_{i_1..i_k}.$$

Then note that

$$[\delta_x]_{i_1..i_k}(E) = \delta_x(S_{i_1} \circ ... \circ S_{i_k}(E)) = \begin{cases} 1, & x \in (S_{i_1} \circ ... \circ S_{i_k})^{-1}(E) \Leftrightarrow S_{i_1} \circ ... \circ S_{i_k}(x) \in E \\ 0, & otherwise \end{cases}$$

For $f \in \mathcal{C}$ (or more generally $f \in \mathcal{C}_c(\mathbb{R}^n)$)

$$\int f d\mu^k = \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} f(x_{i_1..i_k}),$$

and

$$\mu^k(\mathbb{R}^n) = \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} 1 = \frac{m^k}{m^k} = 1.$$

Let $\varepsilon > 0$. That X compact implies $\exists k > 0$ such that

$$|x - y| \leq r^k \mathrm{diam}(X), \quad x, y \in X \implies |f(x) - f(y)| < \varepsilon.$$

Above we have used the fact that $r < 1$ hence $r^k \to 0$ as $k \to \infty$.
If $l > k \geq K$, then since

$$x_{i_1..i_l} \in X_{i_1..i_l} = S_{i_1} \circ ... \circ S_{i_l}(X) = S_{i_1} \circ ... \circ S_{i_k} ... \circ S_{i_l}(X)$$

$$S_{i_{k+1}...i_l}(X) \subset X, \implies S_{i_1} \circ ... \circ S_{i_k} ... \circ S_{i_l}(X) \subset S_{i_1} \circ ... \circ S_{i_k}(X),$$

and

$$\mathrm{diam} X_{i_1..i_k} = r^k \mathrm{diam} X,$$

we have

$$|f(x_{i_1..i_k}) - f(x_{i_1..i_l})| < \varepsilon$$

which follows because $f(x_{i_1..i_k})$ and $f(x_{i_1..i_l})$ are both in $X_{i_1..i_k}$, so

$$|x_{i_1..i_k} - x_{i_1..i_l}| \leq \mathrm{diam}(X_{i_1..i_k}) = r^k \mathrm{diam}(X).$$

Summing over $i_{k+1}..i_l$, and using the trick

$$f(x_{i_1..i_k}) = \frac{1}{m^{l-k}} \sum_{i_{k+1}..i_l=1}^{m} f(x_{i_1..i_k}),$$

because the sum on the right is simply $f(x_{i_1..i_k})$ repeated m^{l-l} times, we have

$$\begin{aligned}\left|f(x_{i_1..i_k}) - \frac{1}{m^{l-k}} \sum_{i_{k+1}..i_l=1}^{m} f(x_{i_1..i_l})\right| =& \left|\Big(\sum_{i_{k+1}..i_l=1}^{m} f(x_{i_1..i_k}) - f(x_{i_1..i_l})\Big)\frac{1}{m^{l-k}}\right| \\ \leq& \frac{1}{m^{l-k}} \sum_{i_{k+1}..i_l=1}^{m} |f(x_{i_1..i_k}) - f(x_{i_1..i_l}))| \\ <& \frac{m^{l-k}\varepsilon}{m^{l-k}}.\end{aligned}$$

Next we sum over $i_1..i_k$ and use the estimate above

$$\begin{aligned}&\left|\frac{1}{m^k}\sum_{i_1..i_k=1}^{m} f(x_{i_1..i_k}) - \frac{1}{m^k}\sum_{i_1..i_k=1}^{m}\Big(\sum_{i_{k+1}..i_l=1}^{m} f(x_{i_1..i_l})\Big)\frac{1}{m^{l-k}}\right| \\ =&\left|m^{-k}\Big(\sum_{i_1..i_k=1}^{m}\Big(f(x_{i_1..i_k}) - \frac{1}{m^{l-k}}\sum_{i_{k+1}..i_l=1}^{m} f(x_{i_1..i_l})\Big)\right| \\ \leq& m^{-k}\sum_{i_1..i_k=1}^{m}\left|f(x_{i_1..i_k}) - \frac{1}{l-k}\sum_{i_{k+1}..i_l=1}^{m} f(x_{i_1..i_l})\right| \\ <& \frac{m^k\varepsilon}{\varepsilon} = \varepsilon.\end{aligned}$$

Since

$$\int f d\mu^l = \frac{1}{m^l}\sum_{i_1..i_l=1}^{m} f(x_{i_1..i_l}), \quad \int f d\mu^k = \frac{1}{m^k}\sum_{i_1..i_k=1}^{m} f(x_{i_1..i_k}),$$

we have

$$\left|\int f d\mu^k - \int f d\mu^l\right| < \varepsilon.$$

We have therefore shown that for any $\varepsilon > 0$ there exists $K \in \mathbb{N}$ such that for $l > k \geq K$,

$$|\int f d\mu^k - \int f d\mu^l| < \varepsilon \Rightarrow \{\int f d\mu^k\}_{k\geq 1} \subset \mathbb{R}$$

is Cauchy and therefore converges.
Consequently we define a bounded linear functional on $\mathbb{C}(\mathbb{R}^n)$ by

$$I(f) := \lim_{k\to\infty}\int f d\mu^k.$$

If

$$f \geq 0 \Rightarrow \int f d\mu^k \geq 0 \forall k \Rightarrow I(f) \geq 0.$$

So, I is non-negative. For $g \in \mathcal{C}_c(\mathbb{R}^n)$,

$$I(f+g) = \lim_{k\to\infty}\int (f+g) d\mu^k = \lim_{k\to\infty}\int f d\mu^k + \lim_{k\to\infty}\int g d\mu^k = I(f) + I(g).$$

Similarly, for $\lambda \in \mathbb{R}$,

$$I(\lambda f) = \lim_{k\to\infty}\int \lambda f d\mu^k = \lambda I(f).$$

Therefore I is linear and non-negative. The functional is bounded because

$$\left|\int f d\mu^k\right| = \left|\int_X f d\mu^k\right| \leq ||f||_\infty \mu^k(\mathbb{R}^n) = ||f||_\infty,$$

which implies

$$|I(f)| \leq ||f||_\infty \quad \forall f \in \mathbb{C}(\mathbb{R}^n),$$

where we note that

$$||f||_\infty := \sup_{x \in X} |f(x)|.$$

By Reisz Representation Theorem there exists a Borel measure μ such that

$$I(f) = \int f d\mu, \quad \forall f \in \mathcal{C}(\mathbb{R}^n).$$

Note since

$$\mu^k(\mathbb{R}^n \setminus X) = 0 \quad \forall k,$$

if a function f has support in $\mathbb{R}^n \setminus X$, then

$$\int f d\mu^k = 0 \, \forall \, k \implies \int f d\mu = 0.$$

Since we can approximate the characteristic function of any compact subset of $\mathbb{R}^n$ by continuous, non-negative functions, it follows that

$$\mu^k(E) \to \mu(E) \quad \text{ for any } E \subset\subset \mathbb{R}^n \implies \mu(E) = 0 \, \forall \, E \subset \mathbb{R}^n \setminus X.$$

Therefore we have

$$\operatorname{supp}(\mu)^c = \cup G, \quad G \subset \mathbb{R}^n \text{ open, such that } \mu(G) = 0,$$
$$\operatorname{supp}(\mu)^c \supset \mathbb{R}^n \setminus X \implies \operatorname{supp}(\mu) \subset X.$$

By the Lebesgue dominated convergence theorem,

$$\int 1 d\mu = \mu(\mathbb{R}^n) = \lim_{k \to \infty} \int 1 d\mu^k = \mu^k(\mathbb{R}^n) = 1.$$

By definition,

$$x_{i_1..i_k} \in X_{i_1..i_k}, \quad \text{ for each } k \in \mathbb{N}.$$

We also have

$$\operatorname{diam}(X_{i_1..i_k}) = r^k \operatorname{diam}(X) \to 0 \text{ as } k \to \infty.$$

By the invariance of X under the family S, we have

$$X = \cup_{i_1 \ldots i_k = 1}^m X_{i_1..i_k}.$$

Then note that for any $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that

$$\operatorname{diam}(X_{i_1..i_k}) = r^k \operatorname{diam}(X) < \varepsilon.$$

This means that for any point $y \in X$, since

$$y \in X = \cup_{i_1 \ldots i_k = 1}^m X_{i_1..i_k},$$

the point y lies in at least one of the elements in the union,

$$y \in X_{i_1..i_k} \implies |y - x_{i_1 \ldots i_k}| \leq \operatorname{diam}(X_{i_1..i_k}) = r^k \operatorname{diam}(X) < \varepsilon.$$

This shows that the collection of points

$$\{\{x_{i_1 \ldots i_k}\}_{i_1 \ldots i_k = 1}^m\}_{k \geq 1}$$

is dense in X, and hence the closure of this collection of points is X. By the definition of μ^k,

$$\operatorname{supp}(\mu^k) = \{x_{i_1 \ldots i_k}\}_{i_1 \ldots i_k = 1}^m.$$

Let p be one of these points, and let f be a compactly supported continuous function with $f(p) = 1$, and $0 \leq f \leq 1$. Then there exists $\epsilon > 0$ and $N \in \mathbb{N}$ such that

$$|y - p| < \epsilon \implies f(y) > 1/2, \quad k \geq N \implies r^k \operatorname{diam}(X) < \epsilon, \quad p \in X_{i_1 \ldots i_N}.$$

Note that we have already seen

$$X_{i_1\dots i_k\dots i_l} \subset X_{i_1\dots i_k} \implies \cup_{i_1\dots i_k\dots i_l} \subset \cup X_{i_1\dots i_k}.$$

Consequently for any $l \geq N$ we know that $p \in X_{i_1\dots i_N}$ and consequently

$$x_{i_1\dots i_N} \in X_{i_1\dots i_N} \implies f(x_{i_1\dots i_N}) \geq 1/2.$$

Similarly, we also have

$$f(x_{i_1..i_N\dots i_l}) \geq 1/2 \quad \forall i_{N+1}\dots i_l.$$

Then we also have for any $k \leq l$,

$$\int f d\mu^l = \frac{1}{m^{l-k}} \sum_{i_{k+1}\dots i_l=1}^{m} \frac{1}{m^k} \sum_{i_1\dots i_k=1}^{m} f(x_{i_1\dots .i_l})$$

and in the second sum taking the specific choice $i_1\dots i_N$ we have

$$\geq \frac{1}{m^{l-N}} \sum_{i_{N+1}\dots i_l=1}^{m} \frac{1}{m^N} f(x_{i_1\dots i_N\dots i_l}) \geq \frac{m^{l-N}}{2m^{l-N}m^N} = \frac{1}{2m^N}.$$

Keeping N fixed and letting $l \to \infty$, this shows that

$$\int f d\mu = \lim_{l\to\infty} \int f d\mu^l \geq \frac{1}{2m^N}.$$

If we had $p \in \operatorname{supp}(\mu)^c$, then since by definition this is an open set, there would be an open neighborhood of this point contained in $\operatorname{supp}(\mu)^c$, and so for such an f with support contained in this neighborhood we'd have

$$\int f d\mu \leq \mu(\operatorname{supp}(f)) = 0.$$

That is a contradiction. Hence the entire set of points

$$\{\{x_{i_1\dots i_k}\}_{i_1\dots i_k=1}^m\}_{k\geq 1} \subset \operatorname{supp}(\mu),$$

and by definition $\operatorname{supp}(\mu)$ is closed so $\operatorname{supp}(\mu)$ contains the closure of these points which is X. We have already seen that $\operatorname{supp}(\mu) \subset X$, so this shows that we have equality.
Finally, by definition,

$$\mu^{k+l} = \frac{1}{m^{k+l}} \sum_{i_1..i_{k+l}=1}^{m} [\delta_x]_{i_1..i_{k+l}},$$

and

$$\mu^l = \frac{1}{m^l} \sum_{i_1..i_l=1}^{m} [\delta_x]_{i_1..i_l}$$

$$\Rightarrow [\mu^l]_{i_1..i_k} = \frac{1}{m^l} \sum_{j_1..j_l=1}^{m} [[\delta_x]_{j_1..j_l}]_{i_1..i_k}$$

$$\begin{aligned}
[\delta_x]_{j_1..j_l}((S_{i_1} \circ .. \circ S_{i_k})^{-1}(E)) &= \delta_x((S_{j_1} \circ .. \circ S_{j_k})^{-1}(S_{i_1} \circ .. \circ S_{i_k})^{-1}(E)) \\
&= \delta_x((S_{j_1}^{-1} \circ .. \circ S_{j_k}^{-1} \circ S_{i_1}^{-1} \circ .. \circ S_{i_k}^{-1}(E)) \\
j_1 := i_{k+1}, j_2 := i_{k+2}\dots j_l := i_{k+l} \Rightarrow &= [\delta_x]_{i_1..i_{k+l}}(E)
\end{aligned}$$

Therefore

$$\mu^{k+l} = \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} [\mu^l]_{i_1..i_k}.$$

Note that

$$\mathcal{X}_{\varphi^{-1}(E)}(x) = \begin{cases} 1, x \in \varphi^{-1}(E) \\ 0, \text{else} \end{cases}$$

and

$$\mathcal{X}_E \circ \varphi(x) = \begin{cases} 1, \varphi(x) \in E \Leftrightarrow x \in \varphi^{-1}(E) \\ 0, \text{else} \end{cases}$$

Therefore,

$$\mathcal{X}_{\varphi^{-1}(E)} = \mathcal{X}_E \circ \varphi.$$

Analogously, (integration is the limit over simple functions i.e sums) and using the definition of μ,

$$\begin{aligned} \int f[d\mu^l]_{i_1..i_k} &= \int f \circ S_{i_1} \circ .. \circ S_{i_k} d\mu^l \\ &\underset{l\to\infty}{\longrightarrow} \int f \circ S_{i_1} \circ .. \circ S_{i_k} d\mu \\ &= \int f[d\mu]_{i_1..i_k} \end{aligned}$$

Let us now assume k is fixed. By the above calculation relating μ^{k+l} and μ^l and the linearity of the integral,

$$\begin{aligned} \int f d\mu^{k+l} &= \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} \int f[d\mu^l]_{i_1..i_k} \\ &\underset{l\to\infty}{\longrightarrow} \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} \int f[d\mu]_{i_1..i_k} \end{aligned}$$

Since

$$\lim_{l\to\infty} \int f d\mu^{k+l} = \int f d\mu$$

by definition, this shows that

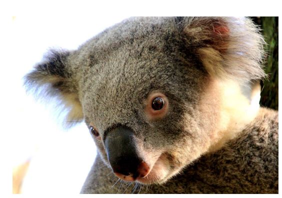
$$\int f d\mu = \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} \int f[d\mu]_{i_1..i_k}.$$

This means that on the right side, we also have a linear functional, namely

$$f \mapsto \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} \int f[d\mu]_{i_1..i_k},$$

which coinicides with our linear functional I. By the proof of the Riesz representation theorem assuming the measure associated with our functional above is constructed in the same way, these measures are therefore the same, and so

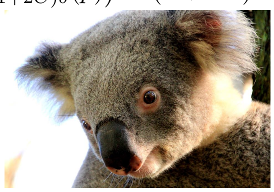
$$\mu = \frac{1}{m^k} \sum_{i_1..i_k=1}^{m} [\mu]_{i_1..i_k}.$$



The $k \in \mathbb{N}$ was arbitrary and fixed, hence this holds for all $k \in \mathbb{N}$.

Lemma 8.13 (Ball counting Lemma). *Let $c, C, \delta > 0$. $\{U_\alpha\}$ open, disjoint, s.t. a ball of radius $c\delta \subset U_\alpha \subset$ ball of radius $C\delta$. Then no ball of radius δ intersects more than $(1+2C)^n c^{-n}$ of the sets $\overline{U}_\alpha$ (<u>note:</u> we are in $\mathbb{R}^n$).*

Proof:
If B is a ball of radius δ, and $B \cap \overline{U}_\alpha \neq \emptyset$, then let p be the center of B, so $B = B_\delta(p)$. $\exists q \in U_\alpha$ s.t. $U_\alpha \subset B_{C\delta}(q)$.
If $x \in U_\alpha$, then for $z \in B \cap \overline{U}_\alpha$, $|z-p| < \delta$ and so $|x-p| \leqslant |x-z| + |z-p| < \operatorname{diam}(B_{C\delta}(q)) + \delta = (1+2C)\delta$.
$\overset{for\ any\ x \in U_\alpha}{\Rightarrow} U_\alpha \subset B_{(1+2C)\delta}(p)$
If N of the $\overline{U}_\alpha$'s intersect B (i.e. have $\neq \emptyset$ intersection), then since they are disjoint and each contains a ball of radius $c\delta$, and they are all contained in $B_{(1+2C)\delta}(p)$.
$\Rightarrow$ adding up the Lebesgue measures of all these N disjoint balls of radius $c\delta$
$\Rightarrow N(c\delta)^n \omega_n \leqslant \mathcal{L}_n\left(B_{(1+2C)\delta}(p)\right) = (1+2C)^n \delta^n \omega_n$



$\Rightarrow N \leqslant (1+2C)^n c^{-n}$.

Theorem 8.14 (Dimension of self-similar sets)**.** *$S = (S_1, ..., S_m)$ is a family of similitudes with common scale factor $r \in (0,1)$. Let U be a separating set, that is an open set, bounded, with $S(U) \subset U$, and $S_i(U) \cap S_j(U) = \emptyset$ if $i \neq j$. Let X be the unique, non-empty, compact set s.t. $S(X) = X$. Let $p := \log_{\frac{1}{r}}(m)$*

i) $H_p(X) \in (0,\infty) \;\Rightarrow p = \dim(X)$
ii) $H_p\left(S_i(X) \cap S_j(X)\right) = 0 \;\; \neq j$

Proof:
For any $k \in \mathbb{N}$, $X = S^k(X) = \bigcup\limits_{i_1,...,i_k=1}^{m} S_{i_1} \circ ... \circ S_{i_k}(X) = \bigcup\limits_{i_1,...,i_k=1}^{m} X_{i_1,...,i_k}$
Each of these $X_{i_1,...,i_k}$ has diameter $= r^k \operatorname{diam}(X)$, so if $\delta_k = r^k \operatorname{diam}(X)$, then

$$H_{p,\delta_k}(X) \leqslant \sum_{i_1,...,i_k=1}^{m} \left(\operatorname{diam}(X_{i_1,...,i_k})\right)^p = m^k r^{pk} \operatorname{diam}(X)^p$$

By definition $p = \log_{\frac{1}{r}}(m) \;\Rightarrow \left(\frac{1}{r}\right)^p = m \Rightarrow m^k = r^{-pk} \Rightarrow H_{p,\delta_k}(X) \leqslant \operatorname{diam}(X)^p$.
$\overset{\delta_k \downarrow 0}{\Rightarrow} H_p(X) \leqslant \operatorname{diam}(X)^p < \infty$, because X is compact hence bounded.
Let $0 < c < C$ s.t. U contains a ball of radius $\frac{c}{r}$ and is contained in a ball of radius $C \left(= \frac{Cr}{r}\right)$.
Let $N = (1+2C)^n c^{-n}$. We will prove that $H_p(X) \geqslant \frac{1}{2^p N}$ by showing that if $\{E_j\}_{j \geqslant 1}$ cover X with $\operatorname{diam}(E_j) \leqslant 1 \forall j$, then $\sum \operatorname{diam}(E_j)^p \geqslant \frac{1}{N 2^p}$.
Since any set E of diameter δ is contained in a ball (closed) of radius $\delta \Rightarrow \operatorname{diam}(E) = \frac{\operatorname{diam}(B_\delta)}{2} \Rightarrow$ $\sum \operatorname{diam}(E_j)^p = \sum \left(\frac{\operatorname{diam}(B_\delta)}{2}\right)^p = \frac{1}{2^p} \sum \operatorname{diam}(B_\delta)^p$
$\Rightarrow$ enough to show that if $X \subset \cup B_j = \cup B_{\delta_j}$ with $\delta_j \leqslant 1 \forall j$, then $\sum\limits_{j=1}^{\infty} \delta_j^p \geqslant \frac{1}{N}$. To prove this, we will prove:
$\star$ If radius of B is $\delta \leqslant 1$ then $\mu(B) \leqslant N\delta^p$.
This shows that $1 = \mu(X) \leqslant \sum \mu(B_j) \leqslant N \sum \delta_j^p$.
To prove $\star$ let $k \in \mathbb{N}$ s.t. $r^k < \delta \leqslant r^{k-1}$. Then $\mu(B) = \frac{1}{m^k} \sum\limits_{i_1,...,i_k=1}^{m} \mu_{i_1,...,i_k}(B)$. Since $X \subset \overline{U}$, $\operatorname{supp}(\mu_{i_1,...,i_k}) = X_{i_1,...,i_k} \subset \overline{U}_{i_1,...,i_k}$. $\Rightarrow \mu_{i_1,...,i_k}(B) \neq 0 \Rightarrow B \cap \overline{U}_{i_1,...,i_k} \neq \emptyset$.
$S_i(U) \cap S_j(U) = \emptyset \; i \neq j \;\Rightarrow$ Since $S(U) \subset U$, we have

$$\begin{matrix} S_k(S_i(U)) \subset S_k(U) \\ S_l(S_i(U)) \subset S_l(U) \end{matrix} \;\Rightarrow S_k(S_i(U)) \cap S_l(S_i(U)) = \emptyset \text{ if } k \neq l$$

$S_k(S_i(U)), S_k(S_j(U))$, $i \neq j$ are also disjoint because $S_i(U) \cap S_J(U) = \emptyset$ if $i \neq j$ and S_k is injective. This shows that if $i_1, ..., i_k \neq j_1, ..., j_k$, then $U_{i_1,...,i_k} \cap U_{j_1,...,j_k} = \emptyset$.

U contains a ball of radius $\frac{c}{r} \Rightarrow U_{i_1,\dots,i_k}$ contains a ball of radius $\frac{c}{r}r^k = cr^{k-1}$
Note: that $cr^{k-1} \geqslant c\delta$ and $Cr^k < C\delta$.
$U_{i_1,\dots,i_k}$ contains a ball of radius $cr^{k-1} \geqslant c\delta$, and is contained in a ball of radius $Cr^k < C\delta$.
$\stackrel{Ball\ counting\ Lemma}{\Rightarrow} B$ can intersect at most $N = (1+2C)^n c^{-n}$ of the $\{\overline{U_{i_1,\dots,i_k}}\}_{i_1,\dots,i_k=1}^m$.

$$\Rightarrow \mu(B) = \frac{1}{m^k} \sum_{i_1,\dots,i_k=1}^{m} \mu_{i_1,\dots,i_k}(B) \leqslant Nm^{-k}.$$

Note that for the last inequality we have used the fact that $\mu_{i_1,..,i_k}$ is supported in $X_{i_1,\dots,i_k} \subset \overline{U_{i_1,\dots,i_k}}$, and the mass of each of these is at most 1 because the total mass is one. Since B intersects at most N of them, the right side of the inequality $m^{-k}N$ follows. Now, recalling that
$p = \log_{\frac{1}{r}}(m) \Rightarrow m^{-k} = r^{kp} \Rightarrow \mu(B) \leqslant Nr^{kp} \leqslant N\delta^p$, since $r^k < \delta \leq 1$. This is $\star$.
Finally since S_j scales by r, we have proven that $H_p(S_j(X)) = r^p H_p(X) = m^{-1}H_p(X)$, using the definition of p. Consequently,

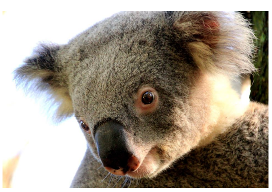
$$\Rightarrow H_p(X) = \sum_{j=1}^{m} H_p(S_j(X)).$$

Since

$$X = \bigcup_{j=1}^{m} S_j(X),$$

this holds iff $H_p(S_i(X) \cap S_j(X)) = 0$ whenever $i \neq j$. More generally, for any measure ν, measurable sets A and B

$$\nu(A \cup B) = \nu(A) + \nu(B) \Leftrightarrow \nu(A \cap B) = 0$$



9. Complex Analysis All-Stars

Definition 9.1. A function f is holomorphic in a neighbourhood $D_r(z_0)$ of z_0, iff $\forall z \in D_r(z_0)$ exists

$$\lim_{w \to z} \frac{f(w) - f(z)}{w - z}$$

(Note: This implies that f is continuous on $D_r(z_0)$.)
$\Leftrightarrow \exists$ continuous function at z: $\forall z \in D_r(z_0)$ we have $f(w) = w(z) + (w-z)A_z(w)$, where A_z is continous at z, for all w in a neighbourhood of z.
($\Rightarrow$). Let $A_z(w) = \frac{f(w)-f(w)}{w-z}$ for $w \neq z$ and $f'(z)$ for $w = z$.
($\Leftarrow$). $\lim_{w\to z} \frac{f(w)-f(z)}{w-z} = \lim \frac{f(z)+(w-z)A_z(w)-f(z)}{w-z} = A_z(w)$.

Why do holomorphic functions have so many properties (theorems)? Bcause they satisfy a PDE. PDEs may endow solutions with special properties depending on the PDE.

Example 9.2 ("Boot strapping"). Let $\Delta := -\partial_x^2 - \partial_y^2$. A priori: $f \in C^2$. Then we have $\Delta f = \lambda f$ for $\lambda \in \mathbb{R}$, on a domain Ω where $f_{\partial\Omega} = 0$. $f \in C^2 \Rightarrow \Delta f = \lambda f \Rightarrow f \in C^4 \Rightarrow \dots \Rightarrow f \in C^\infty$. This is only a heuristic. The actual rigorous argument uses the Sobolev spaces. The solution f is à priori in H^2, but then Δf is also in H^2, which shows that $f \in H^4$. Continuing this we get that $f \in H^{2k}$ for all k which by the Sobolev Embedding Theorem shows that $f \in \mathcal{C}^\infty$.

Proposition 9.3. *f is holomorphic on $D_r(z_0) \Leftrightarrow f$ is $\mathbb{R}^2$ differentiable and $u = \Re(f)$, $v = \Im(f)$ satisfy $u_x = v_y$ and $u_y = -v_x$. These are the Cauchy-Riemann equations.*
$(\Leftrightarrow \bar{\partial} f) = 0$.

Proof:
Assume that f is holomorphic. Near z, $f(w) = f(z) + (w - z)A_z(w)$. For the coordinates $(z, \bar{z}) \in \mathbb{C} \cong \mathbb{R}^2$, we get $x = \frac{z+\bar{z}}{2}$, $y = \frac{z-\bar{z}}{2\imath}$ and $\frac{\partial f}{\partial \bar{z}} = 0$. Therefore we get

$$\frac{\partial f}{\partial \bar{z}} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \bar{z}} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \bar{z}} = \frac{1}{2}f_x - \frac{1}{2\imath}f_y = \frac{1}{2}(f_x - \imath f_y)$$
$$= \frac{1}{2}(u_x + \imath v_x + \imath(u_y + \imath v_y)) = \frac{1}{2}(u_x - v_y + \imath(v_x + u_y)) = 0$$

$\Leftrightarrow u_x = v_y$ and $u_y = -v_x$.
On the other hand, assume f is $\mathbb{R}^2$ differentiable and

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial z} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial z}.$$

Since $\bar{\partial} f = \frac{\partial f}{\partial \bar{z}} = 0$, we have near z_0:

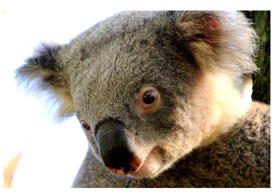
$$f(z) = f(z_0) + M \cdot \begin{bmatrix} z - z_0 \\ z - z_0 \end{bmatrix} + B(z)$$

where $\lim\limits_{z \to z_0} \left\| \frac{B(z)}{z - z_0} \right\| \to 0$.
Since $\bar{\partial} f = 0 \Rightarrow M = \begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix}$ and therefore

$$f(z) = f(z_0) + (z - z_0)\left((a + b) + \frac{B(z)}{z - z_0}\right)$$
$$= f(z_0) + (z - z_0)A(z), \quad A(z) = (a + b) + \frac{B(z)}{z - z_0},$$

and $A(z)$ is continuous because $\frac{B(z)}{z-z_0} \to 0$ as $z \to z_0$. Consequently

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} = a + b$$

exists. 

9.1. Properties of holomorphic functions.

(1) $f(z) = z$ and $f(z) \equiv c$ are holomorphic as in $\mathbb{R}$.
(2) f, g holomorphic $\Rightarrow fg, f + g, f/g (g \neq 0)$ also just as in $\mathbb{R}$.

(3) Not like in $\mathbb{R}$: Given $f\colon \mathbb{R} \to \mathbb{R}$ continuous. $\exists F\colon \mathbb{R} \to \mathbb{R}$ such that $F' = f$? Yes. $F(x) = \int_a^x f(t)dt$.
This is not necessarily true in $\mathbb{C}$.

Lemma 9.4. *If $f\colon \Omega \to \mathbb{C}$, where Ω is a domain, is continuous, and if $\exists F\colon \Omega \to \mathbb{C}$ such that $F' = f$ then $\int_\gamma f(z)dz = 0 \forall$ closed curve $\gamma \subset \Omega$.*

Example 9.5. $f(z) = \frac{1}{z}$ has no primitive since

$$\int_{\partial D_r} f(z)dz = \int_0^{2\pi} f(\gamma(t))\gamma'(t)dt = \int_0^{2\pi} \frac{1}{re^{it}} rie^{it} = 2\pi i \neq 0$$

Theorem 9.6 (Goursat)**.** *If f is holomorphic on Ω, then $\int_{\partial T} f = 0 \,\forall$ triangle $T \subset\subset \Omega$ where $\mathring{T} \subset \Omega$.*

Proof:

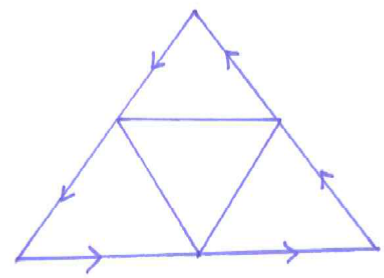


FIGURE 1.

First, we split the triangle into four triangles by joining the midpoints of each of the sides of T. Then integration along the interior edges cancel and so

$$|\int_{\partial T} f| \leq \sum_{i=1}^{4} |\int_{\partial T_i^1} f| \leq 4 \max_{1\leq i\leq 4} |\int_{\partial T_i^1} f|$$

We define T_1 to be any T_i^1 such that the integral achieves the maximum. We repeat this process with T_1, defining T_i^2 for $i = 1, 2, 3, 4$, such that the integral over the boundary of T_1 is equal to the sum of the integrals over the boundaries of the T_i^2. The triangle whose integral is maximal is defined as T_2. This triangle is again split into four, and so forth, defining a nested sequence of triangles

$$T \supset T_1 \supset T_2 \supset \dots$$

Note that the length of the boundary $|\partial T_1| = \frac{1}{2}|\partial T|$ and therefore $|\partial T_k| = 2^{-k}|\partial T|$. Furthermore, we have $\mathrm{diam}(T_1) = \frac{1}{2}\mathrm{diam}(T)$ and therefore $\mathrm{diam}(T_k) = 2^{-k}\mathrm{diam}(T)$.
Since the triangles are compact and nested, and their diameters converge to zero, the intersection

$$\bigcap T_k = \{z_0\} = \lim_{k\to\infty} T_k.$$

Since f is holomorphic at z_0 which is in the interior of Ω,

$$\begin{aligned} f(z) &= f(z_0) + (z - z_0)f'(z_0) + (z - z_0)(A(z) - A(z_0)) \\ &= f(z_0) + (z - z_0)(A(z)). \end{aligned}$$

Note that $B(z) := A(z) - A(z_0)$ is continuous at z_0 because A is, and that $B(z_0) = 0$. Since the function

$$f(z_0) + (z - z_0)f'(z_0)$$

has a primitive, namely

$$F(z) = z(f(z_0) - z_0 f'(z_0)) + \frac{z^2}{2} f'(z_0) \implies F'(z) = f(z_0) + (z - z_0)f'(z_0),$$

the integral

$$\int_{\partial T} (f(z_0) + (z - z_0)f'(z_0))dz = 0, \quad \int_{\partial T_k} (f(z_0) + (z - z_0)f'(z_0))dz = 0, \forall k.$$

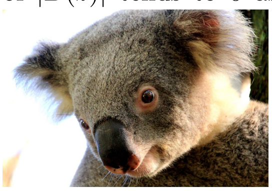
Consequently by linearity of the integral

$$\int_{\partial T_k} f(z)dz = \int_{\partial T_k} (z-z_0)B(z)dz$$

$$\Rightarrow |\int_{\partial T_k} f(z)dz| \leq |\partial T_k| \max_{\partial T_k} |z-z_0||B(z)| \leq |\partial T_k| \mathrm{diam}(T_k) \max_{\partial T_k} |B(z)| = 2^{-k}\mathrm{diam}(T) \max_{\partial T_k|} |B(z)| \cdot 2^{-k}|\partial T|$$

$$\Rightarrow |\int_{\partial T} f(z)dz| \leq 4^k \cdot 4^{-k}\mathrm{diam}(T)|\partial T| \max_{\partial T_k} |B(z)|$$

Since $T_k \to z_0$ and $B(z) \to B(z_0) = 0$ as $z \to z_0$, it follows that the maximum over ∂T_k of $|B(z)|$ tends to 0 as $k \to \infty$. Consequently the integral on the left above must vanish.



Recall that a domain is called *star-shaped* if there exists a point in the domain such that the line segment connecting this point and any other point of the domain lies entirely within the domain. This really looks like a star. Examples include all convex domains.

Proposition 9.7. *If Ω is star-shaped, f holomorphic, f has primitive $F(z) = \int_a^z f$, and $\int_\gamma f = 0 \, \forall$ closed γ.*

Proposition 9.8. *If f holomorphic on $G \setminus z_0$ and continous on G, we also get $\int_\gamma f = 0 \, \forall \gamma$ with $\gamma \cup \mathring{\gamma} \subset\subset G$.*

The converse is also true: If $\int\limits_{\partial T} f = 0 \, \forall T$ satisfying the hypothesis, then f is homolomorphic on G.

Remark 8. If f is holomorphic on $T \setminus z$, where z denotes a point, then $\int\limits_{\partial T} = 0$.

9.2. **Cauchy Integral Formula.** Let f be holomorphic on $D = D_r(z_0) \ni z$. Then

$$f(z) = \frac{1}{2\pi\imath} \int\limits_{\partial D} \frac{f(w)}{w-z} dw$$

Proof: [Sketch] Let

$$g(w) := \begin{cases} \frac{f(w)-f(z)}{w-z} & w \neq z \\ f'(z) & w = z \end{cases}$$

Then g is holomorphic on $D \setminus z$ and it is continous at z.
Therefore since D is convex and hence star-shaped

$$\int_{\partial D} g(w)dw = 0$$

$$\Rightarrow \int_{\partial D} \frac{f(w)}{w-z} dw = \int_{\partial D} \frac{f(z)}{w-z} dw = f(z) \int_{\partial D} \frac{dw}{w-z}$$

Compute $\int\limits_{\partial D} \frac{dw}{w-z_0} dw = 2\pi i$ and prove that the function $h(z) := \int\limits_{\partial D} \frac{dw}{w-z}$ is constant on D.

(1) Expanding $\frac{1}{w-z}$ in a geometric series one can prove that f has a power series expansion.
$$f(z) = \sum_{k\geq 0} a_k(z-z_0)^k.$$
(2) It follows from the above theorem and the Lebesgue Dominated Convergence Theorem that
$$f^{(k)}(z) = \frac{k!}{2\pi i}\int_{\partial D}\frac{f(w)}{(w-z)^{k+1}}dw.$$
(3) The coefficients in the power series expansion are therefore
$$a_k = \frac{f^{(k)}(z_0)}{k!}.$$
(4) **Super-Mega-Differentiability** The derivative of a holomorphic function is holomorphic as are *all* derivatives. Holomorphic functions are infinitely differentiable (and in fact much better than merely $\mathcal{C}^\infty$).

Another straightforward application of the Cauchy Integral Formula is the Maximum Principle.

Theorem 9.9 (Maximum Principle)**.** *$|f|$ has its maximum on the boundary. Otherwise, f is constant.*

Theorem 9.10 (Identity Theorem)**.** *TFAE*
1. $f \equiv g$
2. $f^k(z_0) = g^k(z_0)\,\forall k$ *and some* z_0
3. $f(z_n) = g(z_n)\,\forall n, z_n \neq z_0, z_n \to z_0 \in G.$

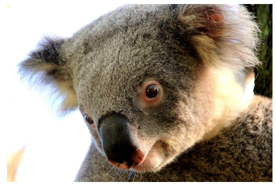
One way to prove the Identity Theorem is to show that 3 $\implies$ 2 by considering $h = f - g$ and the power series expansion at z_0. By continuity $h(z_0) = 0$. So, using the power series expansion of h at z_0, assume all coefficients up to a_j vanish (we know this is true for $j \geq 1$ some j, because $a_0 = h(z_0) = 0$. Then use the assumption to show that $a_j = 0$ also. By induction this shows 2. To show the first statement follows from 2, show that the set of points where $f = g$ is clopen (closed and open). Since the set is non-empty, this means that the set is the entire domain.

Theorem 9.11 (Open Mapping Theorem)**.** *Let $f\colon G \to \mathbb{C}$ be holomorphic and non-constant. Then f is an open map, i. e. $f(G)$ is a domain.*

Proof: Since G is connected and f is continuous, $f(G)$ is also connected.
Let $w_0 = f(z_0)$ and $r > 0$ such that $\overline{D_r(z_0)} \subset\subset G$ and
$$f|_{\overline{D_r(z_0)}\setminus z_0} \neq w_0 \tag{9.1}$$
This follows from the Identity Theorem. Otherwise, we'd have $\{z_n\} \to z_0$ with $z_n \neq z_0$. Then $f(z_n) = w_0 = f(z_0) \Rightarrow f \equiv f(z_0)$ ↯ since f is non-constant.
Let
$$\delta := \min_{\partial D_r(z_0)} |f(z) - w_0| > 0$$
Because $\partial D_r(z_0)$ is compact, the minimum is assumed at some point because of 9.1.
<u>Claim.</u> $D_{\frac{\delta}{2}}(w_0) \subset f(D_r(z_0))$. $\Rightarrow$ Every point in $f(G)$ has an open disk about it in $f(G)$ and therefore, $f(G)$ is open.
Fix w with $|w - w_0| < \frac{\delta}{2}$. Then we get for $z \in \partial D_r(z_0)$ the following equation:
$$|f(z) - w| \geq |f(z) - w_0| - |w - w_0|$$

Therefore, we have $|f(z) - w| > \delta - \frac{\delta}{2}$. The function $g(z) := f(z) - w$ satisfies $|g(z)| > \frac{\delta}{2}$ on $\partial D_r(z_0)$ and $|g(z_0)| < \frac{\delta}{2}$ because $|g(z_0)| = |w_0 - w| < \frac{\delta}{2}$ by assumption.
$\Rightarrow$ either g has a zero in $D_r(z_0)$ or if not $\frac{1}{g}$ is holomorphic on $D_r(z_0)$ and $|\frac{1}{g}| < \frac{2}{\delta}$ on $\partial D_r(z_0)$, but $|\frac{1}{g}(z_0)| > \frac{2}{\delta} \Rightarrow \frac{1}{g}$ has interior maximum and therefore it is constant.

$\Rightarrow f$ constant. $\lightning$. $\Rightarrow g$ must have a zero in $D_r(z_0)$.
$\Rightarrow \exists z \in D_r(z_0)$ such that $g(z) = 0 \Leftrightarrow f(z) = w \Rightarrow w \in f(D_r(z_0))$. Since w with $|w - w_0| < \frac{\delta}{2}$ was arbitrary, we get $D_{\frac{\delta}{2}}(w_0) \subset f(D_r(z_0))$.

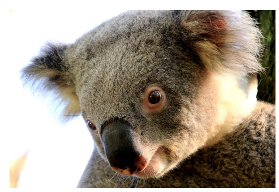


Corollary 9.12. *$f\colon G \to \Omega$ is biholomorphic $\iff$ $f'\big|_G \neq 0$, and f is 1:1.*

Proof:
$(\Rightarrow)$. f is biholomorphic $\Rightarrow f^{-1} \circ f = id$. $\Rightarrow$ Differentiate $(f^{-1})'(f(z))f'(z) = 1$ and therefore $f'\big|_G \neq 0$.
$(\Leftarrow)$. $f'\big|_G \neq 0$. Then f is not constant. Therefore $f(G) = \Omega$ is open. The inverse is continuous by the open mapping theorem. Then we compute

$$\lim_{w \to z_0 = f(\xi_0)} \frac{f^{-1}(w) - f^{-1}(z_0)}{w - z_0} = \lim_{\xi = f^{-1}(w)} \frac{\xi - \xi_0}{f(\xi) - f(\xi_0)} = \frac{1}{f'(\xi_0)}$$

This exists because f is holomorphic and $f'\big|_G \neq 0$.



Definition 9.13. If G, Ω are domains in $\mathbb{C}$ such that $\exists f\colon G \to \Omega$ biholomorphic, then G and Ω are biholomorphically equivalent. A map $f\colon G \to \mathbb{C}$ such that $f'\big|_G \neq 0$ is known as a conformal map and G is conformally equivalent to $f(G)$.

Remark 9. "Conformal" means angle-preserving.

Theorem 9.14 (Uniformization Theorem)**.** *Let $G \subset \mathbb{C}$ be simply connected. Then G is conformally equivalent to one of 1) $\mathbb{C}$ or 2) $\mathbb{D}$ of 3) $\hat{\mathbb{C}} = \mathbb{C} \cup \infty$.*
Moreover, the same holds for any simply connected Riemann surface (2-dimensional Riemannian manifold with biholomorphic coordinate charts $\to \mathbb{C}$).

Theorem 9.15 (Liouville)**.** *Let $f\colon \mathbb{C} \to \mathbb{C}$ be holomorphic. If f is bounded, then it is constant.*

Proof: Assume $|f| \leq M$ on $\mathbb{C}$. The Cauchy Ingegral Formula implies

$$f(z) = \frac{1}{2\pi i} \int\limits_{\partial D_R} \frac{f(w)}{w - z} dw$$

Therefore, we have

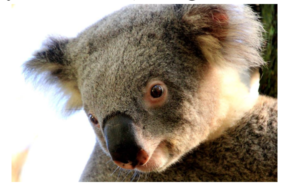
$$f'(z) = \frac{1}{2\pi i} \int_{\partial D_R(z_0)} \frac{f(w)}{(w - z)^2} dw$$

and

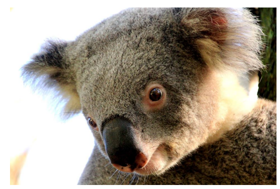
$$f^{(k)}(0) = \frac{k!}{2\pi i} \int\limits_{\partial D_R} \frac{f(w)}{w^{k+1}} dw$$

Therefore, we get the estimation

$$|f^{(k)}(0)| \leq \frac{k!}{2\pi} \frac{2\pi R M}{R^{k+1}} \forall R > 0$$

Letting $R \to \infty$, we get $f^{(k)}(0) = 0 \forall k \geq 1$. Using the Identity Theorem, we get since $f^{(k)}(0) = g^{(k)}(0) \forall k \geq 0$, $g(z) \equiv f(0) \Rightarrow f \equiv g \Rightarrow f \equiv f(0)$ is constant. 

Theorem 9.16 (Fundamental theorem of Algebra). *$p(z)$ is a polynomial with coefficients in $\mathbb{C}$, degree of p is $k \geq 1$. Then $\exists!$ (up to rearrangement) $\{r_j\}_{j=0}^{k}$ in $\mathbb{C}$ such that $p(z) = r_0 \prod_{j=0}^{k}(z - r_j)$.*

Proof: If degree of p is 1, then $p(z) = az + b$ and $a \neq 0 \Rightarrow r_0 = a$ and $r_1 = -\frac{b}{a}$. finish.
By induction on K. If $p\big|_{\mathbb{C}} \neq 0$ then $\frac{1}{p}$ is entire and $\to 0$ at ∞. $\Rightarrow$ bounded $\Rightarrow$ constant $\Rightarrow p$ constant $\lightning$
p has at least one zero $r_k \Rightarrow p$ is polynomial, $\frac{p(z)}{z - r_k}$ is a rational funcion without poles $\Rightarrow$ polynomial.
$p(z) = (z - r_k)q(z)$ where q has degree $k - 1 < k$. $\Rightarrow$ by induction $\exists! \{r_j\}_{j=0}^{k-1}$ such that $q(z) = r_0 \prod_{j=1}^{k-1}(z - r_j)$. $\Rightarrow p(z) = r_0 \prod_{j=0}^{k}(z - r_j)$. 

Definition 9.17. If f is holomorphic on $D_r(z_0) \setminus \{z_0\}$, then z_0 is an isolated singularity.

(i) Removable $\Leftrightarrow \exists!$ holomorphic extension to z_0.
(ii) $f(z) \to \infty$ as $z \to z_0 \Leftrightarrow \exists! g(z)$ holomorphic on $D_p(z_0)$ where $p \leq r$ such that $g(z_0) = 0$ and $f(z) = \frac{1}{g(z)}$ on $D_p(z_0) \setminus \{z_0\}$. z_0 is a pole.
(iii) Neither 1 nor 2. "Essential singularity". If f only has a finite set of singularities on $G \subset \mathbb{C}$ of type 1 and/or type 2, f is called "meromorphic".

Theorem 9.18 (Big Picard Theorem). *If f is holomorphic on $D_r(z_0) \setminus z_0$ and z_0 is an essential singularity, then $\forall \varepsilon \in (0, r)$, $\#\{\mathbb{C} \setminus f(D_\varepsilon(z_0) \setminus z_0)\} \leq 1$.*

Definition 9.19. If f is entire and $\lim_{z \to \infty} f(z) = \infty$, then f has a pole at ∞.

Corollary 9.20. *By Liouvilles Theorem if f is entire, then either 1) f is constant 2) f has a pole at ∞ or neither 1) nor 2) $\Rightarrow f$ has an essential singularity at ∞.*
f has a pole at $\infty \Leftrightarrow \frac{1}{f(\frac{1}{z})} =: g(z)$ is holomorphic near $z = 0$ and $g(0) = 0$.
f has an essential singulariy at $\infty \Leftrightarrow \frac{1}{f(\frac{1}{z})} =: h(z)$ has an essential singularity at 0.

Theorem 9.21 (Riemann's Removable Singularity Theorem). *Let $f\colon D_r(z_0) \setminus z_0 \to \mathbb{C}$ be holomorphic and bounded. Then z_0 is removable.*

Proof: $g(z) := (z - z_0)f(z)$, $z \neq z_0$. g is holomorphic on $D_r(z_0) \setminus z_0$ $\lim_{z \to z_0} g(z) = 0 \Rightarrow$ define $g(z_0) = 0 \Rightarrow g$ is continuous on $D_r(z_0)$. $\Rightarrow g$ is holomorphic on $D_r(z_0)$ and so $\lim_{z \to z_0} \frac{g(z) - g(z_0)}{z - z_0} = g'(z_0)$ exists, and $\lim_{z \to z_0} \frac{(z - z_0)f(z)}{z - z_0} =: f(z_0)$. Consequently this limit exists, is unique, and defining $f(z_0)$ by this limit is unique and makes f continuous at z_0. Moreover, any holomorphic function on a punctured disk which is continuous on the whole disk is in fact holomorphic, which follows from the fact that the integral of such a function over any triangle in the disk vanishes, hence the function has a well-defined primitive. By super-mega differentiability the original function, that is the derivative of the primitive, is also holomorphic. 

10. The Basics of Holomorphic Dynamics

Holomorphic Dynamics is the study of $\{f^n = f \circ \ldots \circ f\}$ and where $\{f^n\}$ converge or not. Consider $\{f^n\}$ where f is defined on a simply connected domain $G \subset \mathbb{C}$. Assume that $f : G \to G$. Let

$$\phi^{-1} : E \mapsto G,$$

be the conformal map given by the Uniformization Theorem, where $E = \mathcal{D}, \mathbb{C}$, or $\hat{\mathbb{C}}$. Then let

$$\tilde{f} := \phi \circ f \circ \phi^{-1} : E \to E.$$

Note that $\tilde{f^n} = \phi \circ f^n \circ \phi^{-1}$. Therefore dynamics of $\{f^n\}$ on G are the same as those of $\{\tilde{f^n}\}$ on E. Therefore the study of holomorphic dynamics on any simply connected domain is reduced, by the Uniformization Theorem, to the study of holomorphic dynamics on $\mathcal{D}$, $\mathbb{C}$, and $\hat{\mathbb{C}}$.
One of the pioneers of the field was Montel ("This is how we do it.")

Definition 10.1. A family of holomorphic functions $\mathcal{F}$ defined on a domain $G \subset \mathbb{C}$ is *normal* if for any sequence in $\mathcal{F}$, there exists a subsequence which converges locally uniformly (this means uniformly on compact subsets).

Theorem 10.2 (Montel's Little Theorem)**.** *If a family $\mathcal{F}$ is uniformly bounded then it is normal.*

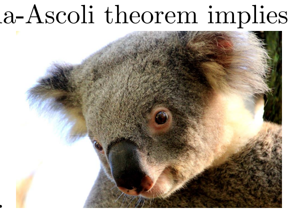
Proof: Let $M \geq ||f||_\infty$ for all $f \in \mathcal{F}$. Fix $z_0 \in G$ and $R > 0$ such that

$$D_R(z_0) \subset\subset G.$$

Then for any $z \in D_{R/2}(z_0)$ we have by the Cauchy Integral Formula for $f \in \mathcal{F}$,

$$f'(z) = \frac{1}{2\pi i} \int_{\partial D_R(z_0)} \frac{f(w)}{(w-z)^2} dw \implies$$

$$|f'(z)| \leq \frac{2\pi R}{2\pi} \frac{M}{(R - R/2)^2} =: c.$$

It follows that the family $\mathcal{F}$ is equicontinuous. Since it was already assumed to be bounded, the Arzela-Ascoli theorem implies that every sequence has a locally uniformly convergent subsequence. 

Remark 10. The locally uniform limit of holomorphic functions is again holomorphic.

By Montel's Theorem, if a function $f : \mathcal{D} \to \mathcal{D}$, then $\mathcal{F} := \{f^n\}$ is a normal family. So, we can already say something about the holomorphic dynamics on $\mathcal{D}$. Now let's consider holomorphic dynamics on $\mathbb{C}$. We want to exclude freaky behavior (i.e. essential singularities).

Theorem 10.3. *If $f : \mathbb{C} \to \mathbb{C}$ is entire and without essential singularity at infinity, then f is a polynomial.*

Proof: First note that if f is bounded, then it is constant, and hence a polynomial of degree 0. How interesting (not). Let us assume that f is non-constant and therefore unbounded, then we must have $|f(z)| \to \infty$ as $|z| \to \infty$. Consequently the function

$$\frac{1}{f(1/z)} = g(z)$$

is holomorphic on a disk about 0 with $g(0) = 0$. Since $f \not\equiv \infty$, we cannot have $g \equiv 0$, and therefore there exists $k \in \mathbb{N}$ such that

$$g(z) = \sum_{j \geq k} a_j z^j, \quad a_k \neq 0.$$

Consequently,

$$f(z) = \frac{1}{g(1/z)} = \frac{1}{a_k z^{-k} + \ldots} = \frac{z^k}{a_k + a_{k+1} z + \ldots} \sim z^k \text{ as } k \to \infty.$$

Next since $|f| \to \infty$ as $|z| \to \infty$, there exists $R > 0$ such that for all $|z| > R$, $|f(z)| > 100000$. In particular for all such z, $f \neq 0$. So, the set of zeros of f is contained in a compact set. Since we assumed that f is not constant, by the identity theorem f can only have a finite set of zeros (of finite order) because they are all contained in a compact set, and so any infinite set would accumulate there thus implying f vanishes identically (ID theorem) which it does not.
Let $\{z_k\}_1^n$ be the zeros of f of respective degrees d_k. Then consider

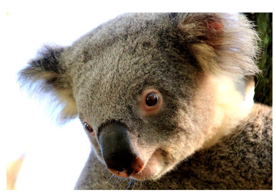
$$\frac{f(z)}{\prod_1^n (z - z_j)^{d_j}}.$$

We know that $|f(z)| \sim |z|^k$ as $|z| \to \infty$. If on the one hand $k < \sum d_j$, then this function tends to 0 at infinity and is entire, hence bounded, hence constant by Liouville's theorem. Since it tends to *zero* at infinity, this would imply the function is identically 0, hence so is f, which is a contradiction. So we must have $k \geq \sum d_j$. Now, on the other hand, we consider

$$\frac{\prod_1^n (z - z_j)^{d_j}}{f(z)}.$$

This function is also entire. If $k > \sum d_j$, then by the same argument we also get a contradiction. Hence $k = \sum d_j$, and so both of these functions are again bounded and entire, hence constant (and that constant cannot be zero), so there is $c \in \mathbb{C} \setminus \{0\}$ such that

$$\frac{f(z)}{\prod_1^n (z - z_j)^{d_j}} \equiv c \iff f(z) \equiv c \prod_1^n (z - z_j)^{d_j}$$

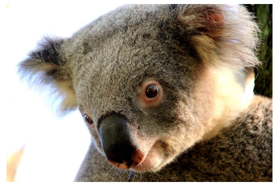
which is a polynomial. 

So, we now see that holomorphic dynamics for entire functions without essential singularity at ∞ is reduced to the study of iteration of polynomial functions.

Proposition 10.4. *Entire functions without essential singularity at infinity which are non-constant are surjective.*

Proof: By the theorem, such a function is a polynomial $p(z)$ of degree $d \geq 1$. Proceeding by contradiction we assume there is $q \in \mathbb{C}$ such that $p(z) \neq q$ for all $z\ in \mathbb{C}$. Then the function

$$\frac{1}{p(z) - q}$$

is entire. Moreover, since $|p(z)| \sim |z|^d$ as $|z| \to \infty$, it follows that this function tends to zero at infinity and hence is bounded. By Liouville the function is constant, which furthermore implies that p is constant which it is not. Therefore the assumption that $p(z) \neq q$ for all $z\ in \mathbb{C}$ must be false, and hence p is surjective. 

Next let's consider holomorphic dynamics for meromorphic functions on $\hat{\mathbb{C}}$.

Theorem 10.5. *Any meromorphic function on $\hat{\mathbb{C}}$ is a rational function. If it is non-constant, then it is surjective.*

Proof: Let's assume $f(z)$ is non-constant and meromorphic. Let $\{p_k\}_1^n$ be the poles of f with corresponding degrees d_k. Then

$$F(z) := f(z) \prod_1^n (z - p_k)^{d_k}$$

is entire, and has at worst a pole at ∞. Therefore this function is a polynomial $q(z)$ and hence

$$f(z) = \frac{q(z)}{\prod_1^n (z - p_k)^{d_k}}$$

is a rational function. To show surjectivity first note that a meromorphic function defined on $\hat{\mathbb{C}}$ without pole is constant by Liouville's theorem (it is entire and bounded!) Therefore the value ∞ is assumed at a pole. For $p \neq \infty$, for the sake of contradiction we assume $f(z) \neq p$ for all $z \in \hat{\mathbb{C}}$. The function $f(z) - p$ may have poles, but it has no zeros, so

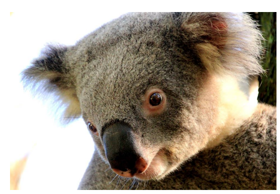
$$\frac{1}{f(z) - p} = g(z)$$

is entire. It has at worst a pole at infinity. If it has no pole at infinity, then it is constant and hence so is f which is a contradiction. So, this function has a pole at infinity and hence is a polynomial. Therefore

$$f(z) - p = \frac{1}{g(z)} \to 0 \text{ as } z \to \infty.$$

Since f is meromorphic, this shows that

$$f(z) \to p \text{ as } z \to \infty \implies f(\infty) = p.$$

Hence f does assume the value p since $\infty \in \hat{\mathbb{C}}$. 

So, holomorphic dynamics for meromorphic functions on $\hat{\mathbb{C}}$ is reduced to the study of iteration of rational functions.

The first two mathematicians to make big progress in holomorphic dynamics were Fatou and Julia, and consequently the two main sets one studies in holomorphic dynamics are named after them.

Definition 10.6. Let $f : G \to G$. The *Fatou set* of f is defined to be

$$\{z \in G : \exists r > 0 \text{ such that } \{f^n\} \text{ is a normal family on } D_r(z)\}.$$

The complement of this set is the *Julia set.*

What are some elements of the Fatou set? Fixed points are certainly a likely candidate, but they do not always belong to the Fatou set. To understand when they do and do not belong to the Fatou set, we classify fixed points into the following types depending upon the local behavior of the function f.

11. Fixed points

Definition 11.1. Let f be holomorphic in a neighborhood of z_0 and assume $f(z_0) = z_0$. The value $\lambda := f'(z_0)$ is known as the *multiplier* at the fixed point z_0.

(1) If $|\lambda| < 1$, then z_0 is an *attracting* fixed point. (hot) If $|\lambda| = 0$, then z_0 is a *super-attracting* fixed point. (super hot).
(2) If $|\lambda| > 1$, then z_0 is a *repelling* fixed point. (not hot)
(3) If there exists $n \in \mathbb{N}$ such that $\lambda^n = 1$, then z_0 is a *rationally neutral* fixed point. (boring).
(4) Otherwise z_0 is an *irrationally neutral* fixed point. (weird).

Note that near the fixed point

$$f(z) = z_0 + \lambda(z - z_0) + \dots, \quad \lambda \neq 0,$$

or presuming f is non-constant, then if $\lambda = 0$ there is some $p \in \mathbb{N}$ such that

$$f(z_0) = z_0 + a_p(z - z_0)^p + \dots.$$

Since the dynamics of f are the same as the dynamics of $\widetilde{f} = \phi^{-1} \circ f \circ \phi$ with

$$\phi(z) = z + z_0,$$

and $\widetilde{f(0)} = 0$, let's assume $z_0 = 0$. Then near the fixed point

$$f(z) = \lambda z + \dots, \quad \text{or} \quad a_p z^p + \dots.$$

So, roughly speaking f looks like either λz or $a_p z^p$. Let's call that function g (either $g(z) = \lambda z$ if $\lambda \neq 0$ or $g(z) = a_p z^p$ if $\lambda = 0$). These functions are significantly more simple than f. Schröder asked the question:

Question 2. *Does there exist a neighborhood of the fixed point and a holomorphic map ψ which conjugates f to g? In other words, does there exist a solution ψ to*

$$\psi \circ f = g \circ \psi?$$

This equation is known as Schröder's equation. Note that it immediately implies that ψ^{-1} is uniquely defined on $g(\psi(D_r))$ via $\psi^{-1}(g(\psi(x)) = f(g(\psi(x)))$ and hence any solution to Schröder's equation is a locally conformal map.

Proposition 11.2. *Let z_0 be an attracting fixed point for an holomorphic function f on $D_r(z_0)$. Then there exists $0 < p \leq r$ such that*

$$f^n(z) \to z_0$$

on $D_p(z_0)$.

Proof: Since f is holomorphic on $D_r(z_0)$, we can write it as a power series

$$f(z) = \sum_{k\geq 0} a_k(z - z_0)^k = \underbrace{a_0}_{=z_0} + \lambda(z - z_0) + (z - z_0)^1 \underbrace{\sum_{k\geq 0} a_{k+2}(z - z_0)^{k+1}}_{\text{this is a convergent power series on} D_r(z_0)}$$

For $\Lambda \in (|\lambda|, 1)$, note that

$$|f(z) - f(z_0)| = \left|\lambda(z - z_0) + (z - z_0)\sum_{k\geq 0} a_{k+2}(z - z_0)^{k+1}\right| \leq |\lambda||z - z_0| + |z - z_0||z - z_0| \underbrace{\left|\sum_{k\geq 0} a_{k+2}(z - z_0)^k\right|}_{\text{convergent}}$$

Because $\sum_{k\geq 0} a_{k+2}(z - z_0)^k$ converges in $D_r(z_0)$, it follows that there are $0 < p \leq r$ and $M > 0$ such that

(1) on $D_p(z_0)$, $\sum_{k\geq 0} a_{k+2}(z - z_0)^k \leq M$;
(2) on $D_p(z_0)$, $|z - z_0| < \frac{\Lambda - \lambda}{M}$, i.e. $p < \frac{\Lambda - \lambda}{M}$.

Therefore,

$$|f(z) - f(z_0)| \leq \underbrace{|\lambda||z - z_0| + \frac{\Lambda - \lambda}{M} M|z - z_0|}_{=\Lambda|z-z_0|}$$

on $D_p(z_0)$. Since $\Lambda < 1$ we have

$$|f(z) - f(z_0)| = |f(z) - z_0| \leq \Lambda|z - z_0| \leq |z - z_0|$$

which shows that

$$f(D_p(z_0)) \subset D_p(z_0).$$

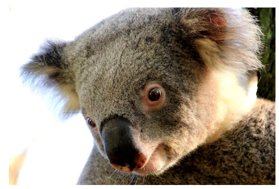
Hence we can apply our estimate to $f(f(z))$ since $f(z) \in D_p(z_0)$ presuming $z \in D_p(z_0)$, and we have

$$|f^2(z) - f^2(z_0)| \leq \Lambda|f(z) - f(z_0)| \leq \Lambda^2|z - z_0|$$

and in general

$$|f^n(z) - f^n(z_0)_{z_0}| = |f^n(z) - z_0| \leq \Lambda^n|z - z_0| \to 0$$

as $n \to \infty$ because $\Lambda < 1$.

This proves that $f^n(z) \to z_0$ for all $z \in D_p(z_0)$. 

Definition 11.3. For an attracting fixed point z_0, the *basin of attraction* of z_0 is

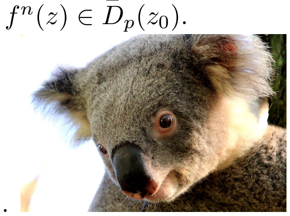
$$A(z_0) := \{z | f^n(z) \text{is defined for all z and} f^n(z) \to z_0\}.$$

We have proven that $D_p(z_0) \subset A(z_0)$.

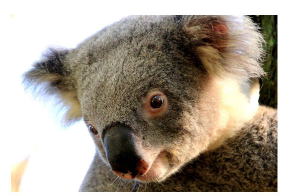
Proposition 11.4.

$$A(z_0) = \bigcup_{n \geq 1} f^{-n}(D_p(z_0)).$$

Proof: "$\subseteq$:" If $f^m(z) \to z_0$, then there exists $N \in \mathbb{N}$ such that for all $m \geq N$ $|f^m(z) - z_0| < p$. Thus $f^m(z) \in D_p(z_0)$ and therefore $z \in f^{-m}(f^m(z)) \in f^{-m}(D_p(z_0))$.
This means that $z \in f^{-m}(D_p(z_0))$, so that $A(z_0) \subseteq \bigcup_{n \geq 1} f^{-n}(D_p(z_0))$.
"$\supseteq$:" If $z \in f^{-n}(D_p(z_0))$ for some $n \geq 1$, then $f^n(z) \in D_p(z_0)$.

Thus $\underbrace{f^k(f^n(z))}_{f^{n+k}(z)} \to z_0$, and therefore $z \in A(z_0)$. 

Corollary 11.5. $A(z_0)$ *is open.*

Proof: Since f is continuous, $f^{-n}(D_p(z_0))$ is open for all $n \geq 1$. 

Definition 11.6. The connected component of $A(z_0)$ containing z_0 is the *immediate basin of attraction*, denoted $A^*(z_0)$.

Definition 11.7. $f : U \to U$ is *conformally conjugate* to $g : V \to V$, if there exists a conformal $\varphi : U \to V$ such that $g = \varphi \circ f \circ \varphi^{-1}$. (Schröder's equation)
(g and f are like the same, only in different coordinate systems).

Question 3. *Can we conjugate f to something simpler?*

- $f(z_0) = z_0$, $\widetilde{f} = T^{-1} \circ f \circ T$, $T(z) = z + z_0$, $\widetilde{f}(0) = 0$, $\widetilde{f}^n = T^{-1} \circ f^n \circ T$, *without loss of generality:* $z_0 = 0$.
- *So* $f(z) = \lambda z + \sum_{k \geq 2} a_k z^k$ *(we can write it loke this because f is holomorphic)*

Then we would have

$$f(z) \sim \begin{cases} \lambda z & \text{if } \lambda \neq 0 \\ a_p z^p & \text{if } \lambda = 0, p = \inf\{k \in \mathbb{N} | a_k \neq 0\} \end{cases}$$

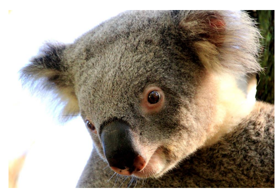
Can we conjugate f to one of these?

Note 1. *If $g = \varphi \circ f \circ \varphi^{-1}$, then z_0 is a fixed point for f if and only if $\varphi(z_0)$ is a fixed point for g.*

Proposition 11.8. *λ at a fixed point for f is the same for g. In words: The multiplier is* invariant *under conjugation by conformal maps.*

Proof: $f'(z_0) = \lambda$. $g = \varphi \circ f \circ \varphi^{-1}$ if and only if $g \circ \varphi = \varphi \circ f$. Therefore,

$$\underbrace{(g \circ \varphi)'(z_0)}_{g'(\varphi(z_0))\varphi'(z_0) = \lambda_g \varphi'(z_0)} = (\varphi \circ f)'(z_0) = \varphi'(\underbrace{f(z_0)}_{z_0})\lambda_f$$



With φ conformal it follows that $\varphi'(z_0) \neq 0$. Thus $\lambda_g = \lambda_f$.

Theorem 11.9. Koenigs
Let f have an attracting fixed point z_0 with $0 < |\lambda| < 1$. Then there exists a conformal mapping $\varphi(z)$ that maps a neighborhood of z_0 onto a neighborhood $D_r(0)$ of zero, such that

$$g(\varphi(z)) = \lambda\varphi(z) = \varphi(f(z))$$

φ is unique up to multiplication by $c \neq 0$.

Proof: Without loss of generality, let $z_0 = 0$. Let $\varphi_n(z) := \lambda^{-n} f^n(z)$. Then we claim that $\varphi_n(z) = z + \dots$, which we will show by induction. It is true for $n = 1$. Assume that it holds for n. Then (replacing f with its power series) we have

$$\varphi_{n+1}(z) = \lambda^{-1}\lambda^{-n} f(f^n(z)) = \lambda^{-1}\lambda^{-n}(\lambda f^n(z) + \dots) = \lambda^{-n} f^n(z) + \dots = z + \dots$$

by the induction assumption. Then

$$\varphi_n \circ f = \lambda^{-n} f^{n+1} = \lambda(\lambda^{-n-1} f^{n+1}) = \lambda\varphi_{n+1}$$

Consequently, if $\varphi_n \to \varphi$, then from $\varphi_n \circ f = \lambda\varphi_{n+1}$, $\varphi_n \circ f \to \varphi \circ f$ and $\lambda\varphi_{n+1} \to \lambda\varphi$ it follows that $\varphi \circ f = \lambda \circ \varphi$ and thus $\varphi \circ f(z) = \lambda\varphi(z)$.
Note that as $\{\varphi_n\}$ are holomorphic and converge locally uniformly towards φ, it follows that φ is holomorhpic too, and all that the derivatives converge, too!
Note that from $\varphi_n'(z) = 1$ for all n, we know that $\varphi'(z) = 1$ (in a neighborhood of $z = 0$ where the power series converge). This means that φ is conformal and

$$\varphi \circ f \circ \varphi^{-1}(\zeta) = \lambda\zeta$$

By the convergence of the power series of f, we know that there exist $c > 0$ fixed, and $\delta > 0$ such that

$$(*)\ |f(z) - \lambda z| \leq c|z|^2 \text{ for } |z| < \delta.$$

Thus

$$|f(z)| \leq |\lambda||z| + c|z|^2 \leq (|\lambda| + c\delta)|z|.$$

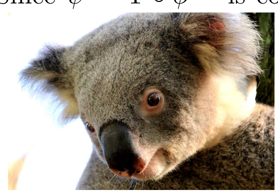
We can now choose δ small enough such that $(|\lambda| + c\delta) < 1$ (which is possible because $|\lambda| < 1$). Then

$$(**)\ |f^n(z)| \leq (|\lambda + c\delta|)^n |z| \text{ for } |z| \leq \delta.$$

Choose δ possibly smaller so that $(|\lambda + c\delta|)^2 \leq |\lambda|^2 + 2|\lambda|c\delta + c^2\delta^2 < |\lambda|$. Then

$$|\varphi_{n+1}(z) - \varphi_n(z)| = |\lambda^{-n-1} f^n(f(z)) - \lambda^{-n} f^n(z)| = \left|\frac{f(f^n(z)) - \lambda f^n(z)}{\lambda^{n+1}}\right|$$

$$\overset{(*)}{\leq} \frac{c|f^n(z)|^2}{|\lambda^{n+1}|} \overset{(**)}{\leq} \frac{c(|\lambda| + c\delta)^{2n}|z|^2}{|\lambda|^{n+1}} = \frac{c\rho^n|z|^2}{|\lambda|} \text{ where } \rho := \frac{(|\lambda| + c\delta)^2}{|\lambda|} < 1.$$

Thus $\{\varphi_n\}$ is Cauchy for $|z| < \delta$ and therefore converges uniformly.
Furthermore, we have to proof the uniqueness of ϕ: On $D_\delta(0)$ we have $\phi(f(z)) = \lambda\phi(z)$ if $\exists\Phi$ which also conjugates $\Phi(f(z)) = \lambda\Phi(z)$ where Φ is conformal. $\Rightarrow \phi(f(0)) = \phi(0) = \lambda\phi(0)$, $\lambda \neq 0 \Rightarrow \phi(0) = 0$. Similarly, we show $\Phi(0) = 0$.
$\Phi \circ \phi^{-1} \circ \lambda = \lambda \circ \Phi \circ \phi^{-1}$. Let $\psi = \Phi \circ \phi^{-1}$. Then we have $\psi \circ \lambda = \lambda \circ \psi$. Note that $\psi(0) = \Phi(\phi^{-1}(0)) = \Phi(0) = 0$. Therefore, $\psi(z) = cz + \ldots$ near $z = 0$ and $\psi(\lambda z) = \lambda\psi(z)$. Therefore $c\lambda z + a_2(\lambda z)^2 + \ldots = \lambda(cz + a_2 z^2 + \ldots)$. By the identity theorem, the coefficients are all identical and we get $a_k\lambda^k = \lambda a_k \forall k \geq 2$. By assumption $0 < |\lambda| < 1 \Rightarrow |\lambda^k| \neq |\lambda| \forall k \geq 2$. Therefore, $a_k = 0 \forall k \geq 2$. Since $\psi = \Phi \circ \phi^{-1}$ is conformal, $\psi(0) = c \neq 0$, and $\psi(z) = cz = \Phi \circ \phi^{-1}(z) \Rightarrow \Phi(z) = c\phi(z)$. 

Corollary 11.10. *If z_0 (WLOG $= 0$) is repelling, then $\exists!$ (up to $\star$ by $c \neq 0$) conformal ϕ conjugating $f(z)$ to λz.*

Proof: $f(z) = \lambda z + \ldots$ on $D_r(0)$. Since $|\lambda| > 1 > 0$, then $f'(0) \neq 0$ and WLOG we may assume $|f'(z)| \geq \frac{\lambda}{2}$ on $D_r(0)$. Therefore $f'|_{D_r(0)} \neq 0$ and f^{-1} is holomorphic on $f(D_r(0))$. Furthermore, we get $(f^{-1})'(0) = \lambda^{-1}$. Moreover $f^{-1}(0) = 0$ and $f^{-1}(z) = \lambda^{-1}z + \ldots$ on $D_\rho(0)$. Apply Koenig's Theorem to f^{-1}. $\exists!$ (up tp scale) ϕ conjugating f^{-1} to λ^{-1}. $\Rightarrow f^{-1} \circ \phi = \phi \circ \lambda^{-1}$ $\Rightarrow f \circ f^{-1} \circ \phi = f \circ \phi \circ \lambda^{-1}$. Therefore, $\phi = f \circ \phi \circ \lambda^{-1} \Rightarrow \phi \circ \lambda = f \circ \phi$. The uniqueness follows from the Theorem. 

Theorem 11.11. *If f has a super-attracting fixed point (WLOG at 0) then for $f \not\equiv 0$ $\exists!$ (up to $p-1$ root of unity) conformal ϕ such that $f \circ \phi = g \circ \phi$ and $g(z) = z^p$, where $f(z) = a_p z^p + \ldots$ on $D_r(0)$.*

Proof: Fix $c > 1$. Then $\exists\delta > 0$ such that $\forall\, |z| \leq \delta$, $|f(z)| \leq c|z|^p$. Then

$$|f(f(z))| \leq c|f(z)|^p \leq c(c|z|^p)^p.$$

Since we know $p \geq 2$, we get $cc^p \leq c^{p^2}$ and $|f(f(z))| \leq cc^p|z|^{p^2} \leq c^{p^2}|z|^{p^2}$.
Induction assume $|f^n(z)| \leq (c|z|)^{p^n}\ \forall z \in D_\delta(0)$ (we can choose $\delta > 0$ so small that $f(z) \in D_\delta(0)$ for $|z| \leq \delta$). Then

$$|f^{n+1}(z)| = |f(f^n(z))| \leq c|f^n(z)| \leq cc^{p^n}(|z|^{p^n})^p$$

Since $p \geq 2$ and $c > 1$, we get $c^{p^n+1} \leq c^{p^{n+1}}$. So induction shows that the statement is true $\forall n \in \mathbb{N}$.
So $f^n(z) \to 0$ super exponentially as $n \to \infty$ on $D_\delta(0)$. For $z \in D_\delta(0)$ define

$$b := \left(\frac{1}{|a_p|}\right)^{\frac{1}{p-1}} \cdot e^{\frac{i(2\pi-\theta)}{p-1}}$$

where $a_p := |a_p|e^{i\theta}$ and $\theta \in [0, 2\pi)$. Let $\phi(z) := bz$. Then $f \cong \tilde{f} = \phi^{-1} \circ f \circ \phi$ where

$$\tilde{f} = b^{-1}(a_p(bz)^p + \ldots) = a_p b^{p-1} z^p + \ldots = z^p + \ldots.$$

Since $f \cong \tilde{f}$, we may assume WLOG $f(z) = z^p + \ldots$ on $D_\delta(0)$. We're looking for $\phi(z) = z + \ldots$ such that $\phi(f(z)) = \phi(z)^p$. Define $\phi_n(z) := (f^n(z))^{p^{-n}}$. Then since $f(z) = z^p + \ldots$ we get $f(f(z)) = (f(z))^p + \ldots = (z^p)^p + \ldots = z^{p^2} + \ldots$, so

$$f^n(z) = z^{p^n} + \ldots \Rightarrow (f^n(z))^{p^{-n}} = (z^{p^n}(1+\ldots))^{p^{-n}} = z^{p^n p^{-n}}(1+\ldots)^{p^{-n}} = z(1+\ldots)^{p^{-n}}$$

Thus

$$\phi_{n-1} \circ f = (f^{n-1}(f(z)))^{p^{-n+1}} = (f^n(z))^{p^{-n} \cdot p} = (\phi_n(z))^p$$

If $\phi_n \to \phi$ as $n \to \infty$, then $\phi \circ f = \phi^p$.
First, assume $\{a_n\}_{n\geq 1}$ are positive. Then $\prod\limits_{n\geq 1} a_n$ converges and is positive $\Leftrightarrow \sum \log(a_n)$ converges.
We will prove that $\prod\limits_{n=1}^{N} \frac{\phi_{n+1}}{\phi_n}$ converges ($\Rightarrow \frac{\phi_{N+1}}{\phi_1}$ converges $\Rightarrow \phi_{N+1}$ converges) where

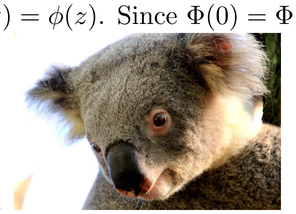
$$\frac{\phi_{n+1}}{\phi_n} = \frac{(f^{n+1})^{p^{-n-1}}}{(f^n)^{p^{-n}}} = \frac{(\phi_1(f^n))^{p^{-n}}}{(f^n)^{p^{-n}}}$$

since $\phi_1 \circ f^n = (f(f^n))^{p^{-1}} = (f^{n+1})^{p^{-1}}$. Therefore,

$$\begin{aligned}(\phi_1 \circ f^n)^{p^{-n}} &= \left((f^{n+1})^{p^{-1}}\right)^{p^{-n}} = (f^{n+1})^{p^{-1}\cdot p^{-n}} = \left(f^{n+1}\right)^{p^{-n-1}} \\ &= \left(\frac{f^n(1+f^n+\ldots)^{p^{-1}}}{f^n}\right)^{p^{-n}} = (1+\mathcal{O}(|f|^n))^{p^{-n}} \\ &= 1 + \mathcal{O}(p^{-n})\mathcal{O}(c^{p^n}|z|^{p^n}) = 1 + \mathcal{O}(p^{-n})\end{aligned}$$

if $|z| \leq c^{-1}$. The last estimate follows since for a sufficiently small x, expanding in a geometric series

$$\begin{aligned}(1+x)^{-y} &\cong \left(\frac{1}{1+x}\right)^y \cong \left(\sum(-1)^n x^n\right)^y \cong (1-x+\ldots)^y \\ &\cong 1 + cyx + \ldots \cong 1 + \mathcal{O}(yx)\end{aligned}$$

Since $\mathcal{O}(p^{-n}) = \mathcal{O}(-p^{-n})$, $\sum\limits_{n\geq 1} \log(1+\mathcal{O}(p^{-n})) \cong \sum\limits_{n\geq 1} p^{-n}$ converges since $p \geq 2$. Therefore, $\prod \frac{\Phi_{n+1}}{\Phi_n}$ converges and $\prod\limits_{n=1}^{N} \frac{\Phi_{n+1}}{\Phi_n} = \frac{\Phi_{N+1}}{\Phi_1}$. Therefore $\Phi_{n+1} \to \Phi$.
Uniqueness: If $\psi \circ f = \psi^p$, $\psi(f(0)) = \psi(0) = \psi(0)^p$. Let $\Phi := \phi \circ \psi^{-1}$. This satisfies $\Phi(z^p) = (\Phi(z))^p \Rightarrow \Phi(0) = \Phi(0)^p$ and $\Phi'(z^p)pz^{p-1} = p\Phi'(z)$. For $z = 0$ and $p \geq 2 \Rightarrow \Phi'(0) = 0$. Repeat this to show that $\Phi^{(k)} = 0 \, \forall k \geq 1$. Therefore, $\Phi \equiv \Phi(0) \Rightarrow \Phi(0) = \phi \circ \psi^{-1}$ and $\Phi(0)\psi(z) = \phi(z)$. Since $\Phi(0) = \Phi(0)^p \iff \Phi(0)^{p-1} = 1$ this shows equality up to a $p-1$ root of unity. 

Proposition 11.12. *Let $\lambda = e^{2\pi\imath\theta}$ where $\theta \notin \mathbb{Q}$. Then the solution h to $f(h(z)) = h(\lambda z), h'(0) = 0$, where f is holomorhic and has a fixed point at $z = 0$, is 1:1 in D_r for some $r > 0$.*

Proof: Assume for some z in D_r on which h is defined we have $h(z) = h(z')$. Then

$$f(h(z)) = f(h(z')) = h(\lambda z) = h(\lambda z').$$

We can repeat this and we obtain

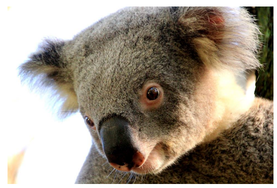
$$h(\lambda^n z) = h(\lambda_n z'), \quad \forall n \in \mathbb{N}.$$

Since the multiplier λ has $\theta \notin \mathbb{Q}$, the set $\{\lambda^n\}_{n\in\mathbb{N}}$ is dense in $\partial\mathcal{D}$. Therefore we have $h(wz) = h(wz')$ for every $w \in \partial\mathcal{D}$ by continuity of h. Considering the function

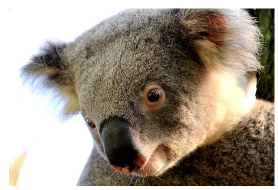
$$g(w) := h(wz) - h(wz'), \quad |w| \leq 1,$$

we see that $|g| \equiv 0$ on $\partial\mathcal{D}$. By the Maximum Principle it follows that $g \equiv 0$ on $\mathcal{D}$. Since $h'(0) = 1$, we have

$$g'(0) = zh'(0) - z'h'(0) = z - z' = 0,$$

which follows because g is constant, and we have used the chain rule (h is holomorphic). So, this shows that if $h(z) = h(z')$ then $z = z'$. 

Proposition 11.13. *A solution h exists iff $\{f^n\}$ is uniformly bounded on some $D_r(0)$.*

Proof: The proof is an exercise. 

Theorem 11.14. *There is $\lambda = \exp 2\pi i\phi$ so that the Schröder Equation has no solution for any polynomial f. (1917 - Pfeiffer).*

Definition 11.15. ϕ is *Diophantine* (badly approximable by rational numbers) if there exists $c > 0, \mu < \infty$ so that $|\phi - \frac{p}{q}| \geq \frac{c}{q^\mu}$ for all $p, q \in \mathbb{Z}, q \neq 0$. This is equivalent to $|\lambda^n - 1| \geq cn^{1-\mu}$ for all $n \geq 1$.

Remark 11. Almost all real numbers are Diophantine - but not all!

Theorem 11.16. *(Siegel, 1950s) If ϕ is Diophantine, $f(0) = 0$, $f'(0) = \exp 2\pi i\phi$, then there exists a solution h to Schröder's Equation.*

Remark 12. For $P(z) = \exp 2\pi i\phi z + z^2$, $\{\frac{p_n}{q_n}\} \to \phi$ continued fraction expansion, then there is a solution to the Schröder Equation if and only if $\sum \frac{\log q_{n+1}}{q_n} < \infty$.
(Sufficiency was shown by Brunjo in 1965, neccessity in 1988 by Yoccoz.)
(Every real number can be expressed as a limit $\lfloor x \rfloor + \frac{a}{b+s...}$.)

Definition 11.17. A simply connected component of the Fatou set such that f is conformally conjugate to an irrational rotation is a *Siegel disk*.

Theorem 11.18. *The Julia set contains all repelling fixed points and all neutral fixed points which do not correspond to Siegel disks.*
The Fatou set contains all attracting fixed points and all those neutral ones corresponding to Siegel disks.

Proof: Rationally neutral fixed points

(1) $\lambda = 1, p = 1$
(2) $\lambda = 1, p > 1$
(3) $\lambda^n = 1, \lambda \neq 1$

$\Rightarrow f(z) = \lambda z + az^{p+1} + \dots, a \neq 0$.
Case 1: Conjugate f by $\varphi(z) = az \to \tilde{f} = \varphi \circ f \circ \varphi^{-1} = a(f(\frac{z}{a})) = a(\frac{\lambda z}{a}) + a(\frac{z}{a})^2 + \dots) = \lambda z + z^2 + \dots \Rightarrow$ WLOG$a = 1$. Move 0 to ∞ by $z \to \frac{-1}{z} \to g(z) = z + 1 + \frac{b}{z} + \dots$. Fatou proved that φ conjugates g to $z \to z + 1$.
Case 2: Another conjugation..
Case 3: Reduce to case 1 or case 2 by considering f^n.
We conclude that at such a fixed point, there are both "repelling" and "attracting" directions. Thus all rationally neutral fixed points are in $\mathcal{J}$.

By definition, if a neutral fixed point is irrationally neutral and corresponds to a Siegel Disk: $\exists r > 0$ s.t. on $D_r(z_0)$,

$$f(z) = \varphi \circ \lambda \varphi^{-1}(z)$$

and by definition, the simply connected component of $\mathcal{F}$ containing $D_r(z_0)$ is in $\mathcal{F}$.

By the proposition $\{f^n\}$ is uniformly bounded on some $D_r(z_0)$. By Montel it follows that $\{f^n\}$ is normal there. Also by the proposition, $\{f^n\}$ is uniformally bounded on $D_r(z_0)$ if and only if f conjugates to an irrational rotation. Therefore $z_0 \in \mathcal{F}$ which is equivalent to f being conjugated to an irrational rotation, i.e. z_0 corresponds to a Siegel Disk.
We have proven that for all attracting fixed points z_0, $\exists r > 0$ s.t. $f^n(z) \to z \forall z \in D_r(z_0)$ it follows that $D_r(z_0) \subset \mathcal{F}$.

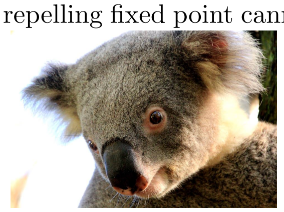
For repelling fixed points, WLOG $z_0 = 0$, we construct a contradiction:
Assume $D_r(0)$ is contained in the Fatou set. Assume the family $f^{n_k} \to g$ locally uniformly. Then $g(0) = 0$, and so for sufficiently small z, $|g(z)| < r/2$, so for large n_k, $|f^{n_k}(z)| < 3r/4$ hence $f^{n_k}(z) \in D_r(0)$. By definition of the conjugating map

$$\varphi \circ f = \lambda \varphi \implies \varphi \circ f \circ \varphi^{-1}(z) = \lambda z,$$

whenever the left side is defined. Taking r possibly smaller, since we have proven that the conjugating map also fixes the fixed point and is locally injective, we may assume $\varphi^{-1}(z) \in D_r(0)$ for $z \in D_r(0)$. Then f is defined there. Then taking $\delta < r$ such that $g(z) \in D_r(0)$ for $|z| < \delta$ which can be done by the above argument, we also have $f^{n_k}(z) \in D_r(0)$ for n_k sufficiently large. Taking ϵ sufficiently small, $|\varphi^{-1}(z)| < \delta$ for all $|z| < \epsilon$. So, for $|z| < \epsilon$, $f^{n_k}(z)$ is in $D_r(0)$ for all n_k large, hence $\varphi(f^{n_k}(\varphi^{-1}(z)))$ is defined for all $|z| < \epsilon$ and all n_k sufficiently large. Moreover, on this disk $\varphi \circ f^{n_k} \circ \varphi^{-1} \to \varphi \circ g \circ \varphi^{-1}$. So, we have for any $z \in D_\epsilon(0) \setminus \{0\}$

$$\varphi \circ f^{n_k} \circ \varphi^{-1}(z) = \lambda^{n_k} z \to \infty,$$

$$\varphi \circ f^{n_k} \circ \varphi^{-1}(z) \to \varphi \circ g \circ \varphi^{-1}(z).$$

This is a contradiction because the function on the right is holomorphic on $D_\epsilon(0)$. Consequently the repelling fixed point cannot be contained in the Fatou set and all such points lie in the Julia set. 

12. Iteration of rational functions

We know how to classify fixed points and whether they are in the Fatou or Julia set. If we are interested in iteration of meromorphic functions on $\hat{\mathbb{C}}$, we have proven that all such functions are rational functions. In this case we also know precisely how many fixed points such functions have.

Theorem 12.1. *A rational map of degree d has precisely $d+1$ fixed points, unless of course it is the identity.*

Proof: We assume that R can be conjugated so as not to fix infinity. For example, for $c \in \mathbb{C}$ let

$$\phi(z) := z^{-1} - c, \quad \phi^{-1}(z) = \frac{1}{z+c}.$$

Then

$$S := \phi^{-1} \circ R \circ \phi : \infty \to -c \to R(-c) \to \frac{1}{R(-c)+c}.$$

So choosing some $-c \neq R(-c)$ which is possible unless R is the identity (which we of course assume it is not) we have S does not fix ∞ and S is conformally conjugate to R. The fixed points of R are in bijection with the fixed points of S hence without loss of generality from now on we assume $R(\infty) \neq \infty$. Let $\zeta \neq \infty$ be a fixed point of $R = P/Q$. Then $Q(\zeta) \neq 0$ since $\zeta \neq \infty$ and $R(\zeta) = \zeta$. So it follows that the degree of the 0 of the function $R(z) - z$ at $z = \zeta$ is the same as the degree of the 0 of the function $P(z) - zQ(z)$ at $z = \zeta$ since

$$R(\zeta) = \zeta \iff R(\zeta) - \zeta = \frac{P(\zeta) - \zeta Q(\zeta)}{Q(\zeta)} = 0.$$

The number of fixed points of R is therefore equal to the number of solutions to $P(z) = zQ(z)$ counting multiplicity. Since R does not fix ∞, the degree of the numerator of R is less than or equal to the degree of the denominator, so the degree of P is less than or equal to the degree of Q and hence the degree of R is equal to the degree of Q. The degree of the polynomial

$$P(z) - zQ(z)$$

is therefore the degree of Q plus one which is equal to $d+1$. By the Fundamental Theorem of Algebra this polynomial has precisely $d+1$ zeros counting multiplicity. These are in bijection with the fixed points of R hence R has precisely $d+1$ fixed points counting multiplicity.

We will use the above result to prove the following.

Theorem 12.2. *The Julia set is not empty for rational functions with degree ≥ 2.*

Proof: Assume $\mathcal{J} = \emptyset$. Then the family of iterates of R is normal on $\hat{\mathbb{C}}$ which is compact, hence there exists a uniformly convergent subsequence. Passing to that subsequence we assume WLOG that $R^n \to f$. Since we have uniform convergence, it follows that f is a meromorphic function on $\hat{\mathbb{C}}$, and hence we have shown that all such functions are rational functions. So f has some degree D. Let d be the degree of R. Then we first claim the following.

Claim 4. *R^n has degree d^n.*

Proof: Write

$$R(z) = \frac{p(z)}{q(z)}, \quad p(z) = a\prod_1^n (z - r_k), \quad q(z) = b\prod_1^m (z - s_j).$$

Consider $R(R(z)) =$

$$\frac{a\prod(p/q - r_k)}{b\prod(p/q - s_j)} = \frac{aq(z)^{-n}\prod(p(z) - r_k q(z))}{bq(z)^{-m}\prod(p(z) - s_j q(z))}.$$

Assume $n = d$ is the degree of R, in which case $n \geq m$. Then

$$R(R(z)) = \frac{a\prod(p(z) - r_k q(z))}{bq(z)^{n-m}\prod(p(z) - s_j q(z))}.$$

Consider the numerator

$$a\prod(p(z) - r_k q(z)) = ap(z)^n + l.o.t. = a^{n+1}z^{n^2} + l.o.t.$$

The numerator has degree n^2. Consider the denominator

$$bq(z)^{n-m}\prod(p(z) - s_j q(z)) = bq(z)^{n-m}p(z)^n + l.o.t. = b^{n-m+1}a^n z^{m(n-m)+mn} + l.o.t,$$

where we use l.o.t. to denote lower order terms. This has degree

$$2mn - m^2 \leq n^2 \iff n^2 + m^2 - 2mn \geq 0,$$

which is true since

$$n^2 - 2mn + m^2 = (n-m)^2 \geq 0.$$

So it appears that the degree of $R(R(z))$ is $n^2 = d^2$, but what about cancellation? The numerator vanishes iff

$$p(z) = r_k q(z)$$

for some r_k and some z. The denominator vanishes iff

$$p(z) = s_j q(z)$$

for some s_j and some z. If (for the sake of contradiction) there is some z such that both numerator and denominator vanish, then there are some r_k and s_j such that

$$p(z) = r_k q(z) = s_j q(z).$$

By definition of a rational function, p and q have no common zeros. This means that $r_k \neq s_j$ for all k and j. In order for the equation above to be satisfied we would need either $r_k = s_j = 0$ which is a contradiction, or $p(z) = q(z) = 0$ which is also a contradiction. So, the numerator and denominator of $R(R(z))$ have no common zeros and hence the degree of $R(R(z))$ is indeed n^2.

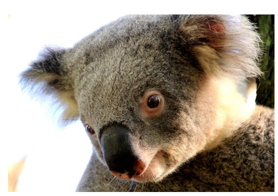
If instead the degree of q is greater than or equal to the degree of p, so that $m \geq n$, then we again look at the numerator and denominator. In the numerator we have

$$aq(z)^{m-n}\prod(p(z) - r_k q(z)) = a(\prod -r_k)q(z)^m + aq(z)^{m-n}p(z)^n + l.o.t.,$$

which has degree m^2 unless some $r_k = 0$. The denominator

$$b\prod(p(z) - s_j q(z)) = b(\prod -s_j)q(z)^m + l.o.t.,$$

which also has degree m^2 unless some $s_j = 0$. We cannot have some $r_k = 0$ *and* some $s_j = 0$ because then p and q would have a common root (0). So at least one of these is non-zero, and the degree of $R(R(z))$ is $m^2 = d^2$.



The same argument and a bit of induction shows that the degree of R^n is d^n.

Now we know that something must go amok. We have $R^n \to f$ uniformly on $\hat{\mathbb{C}}$ yet the degree of $R^n = d^n \to \infty$ since $d \geq 2$, but the degree of f is D which is fixed. We know that R^n has $d^n + 1$ fixed points, yet f has only $D + 1$ fixed points. Consider

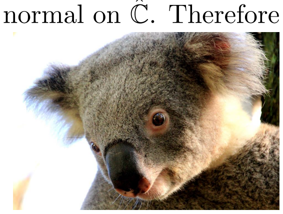
$$R^n(z) - z \to f(z) - z,$$

and the convergence is uniform on $\hat{\mathbb{C}}$. Let $\{f_k\}_{k=1}^{D+1}$ be the fixed points of f. Fix a small disk D_k about f_k which contains f_k but does not contain any pole of f (presuming f_k is not ∞). Then by the Argument Principle, the multiplicity of the fixed point at f_k is equal to

$$\int_{\partial D_k} \frac{f'(z) - 1}{f(z) - z} dz.$$

By the uniform convergence of R^n to f, the corresponding integral with R^n replacing f converges to this. By the Argument Principle this is equal to the number of zeros of $R^n(z) - z$ minus the number of poles. That is an *integer.* The only way a sequence of integers can converge to an integer is if it is eventually constant. So, for large n, the number of fixed points of R^n in D_k minus the number of poles in D_k is equal to the multiplicity m_k of the fixed point f_k. Note that for $f_k = \infty$ we can use the same argument with a conformal conjugation which moves the fixed point to a different point (0, for example), because we have seen that z is a fixed point for f iff $\phi^{-1}(z)$ is a fixed point for $\widetilde{f} = \phi^{-1} \circ f \circ \phi$.

So, for large n, the number of fixed points of R^n in D_k minus the number of poles in D_k is equal to m_k. If the number of fixed points of R^n in D_k stays bounded, then since the number of fixed points is $d^n + 1 \to \infty$, the fixed points must accumulate in $\hat{\mathbb{C}}$ outside of D_k. Let's say they accumulate at z_0 so z_n is a fixed point for R_n and $z_n \to z_0$. Then since $R^n \to f$ uniformly, z_0 must be a fixed point of f, which it is not. So, for at least one D_k, the number of fixed points of R^n in D_k is tending to infinity with n (passing to a subsequence if necessary). Since the number of fixed points is equal to m_k plus the number of poles, this means that the number of poles of R^n in D_k is tending to infinity. Hence they accumulate somewhere at say $z_0 \in D_k$. By the uniform convergence of $R^n \to f$, z_0 must be a pole of f. However, we assumed that D_k did not contain any poles of f. So, this too is a contradiction.

Consequently, this shows that it is impossible for $R^n \to f$ uniformly on $\hat{\mathbb{C}}$ and hence the family of iterates $\{R^n\}$ is not normal on $\hat{\mathbb{C}}$. Therefore at lest one point of $\hat{\mathbb{C}}$ is in the Julia set, which is therefore non-empty. 

13. The Fatou and Julia sets of rational functions

Definition 13.1. E is *completely invariant* if E and E^c are invariant under R (R rational), meaning $R(E) \subset E$ and $R(E^c) \subset E^c$.

Proposition 13.2. *This is the case if and only if $R(E) = E$.*

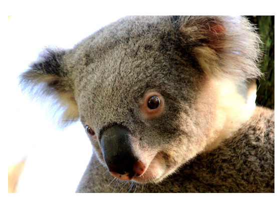
Proof: Assume E is completely invariant. In this case, since

$$R(\hat{\mathbb{C}}) = R(E \cup E^c) = \hat{\mathbb{C}} = E \cup E^c = R(E) \cup R(E^c) \subset E \cup E^c$$

if R is invariant, but since we have equality, it follows that $R(E) = E$, $R(E^c) = E^c$ and therefore $E = R^{-1}$ and $E^c = R^{-1}(E^c)$.

On the other hand, if $E = R^{-1}(E)$, it follows that $R(E) = E$ and $E^c = \hat{\mathbb{C}} \backslash R^{-1}(E)$. Thus

$$R(E^c) = R(\hat{\mathbb{C}} \backslash R(R^{-1}(E)) = R(\hat{\mathbb{C}}) \backslash E = \hat{\mathbb{C}} \backslash E$$

since $R(\hat{\mathbb{C}}) = \hat{\mathbb{C}}$ (R not a constant). 

Theorem 13.3. *The Julia set $\mathcal{J}$ is completely invariant.* ***Proof:*** *Let $z_0 \in D_r(z_0) \subset \mathcal{F}$. Then R^{n_k} converges uniformly on $\overline{D_{r/2}(z_0)}$.*

$\Rightarrow$ *R^{n_k+1} converges uniformly on $R^{-1}(\overline{D_{r/2}(z_0)})$ which is also compact.*

$\Rightarrow$ *$\{R^n\}$ is normal on $R^{-1}(D_{r/2}(z_0))$*

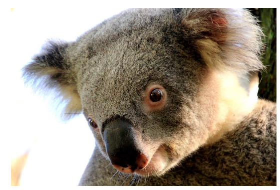
$\Rightarrow$ *$R^{-1}(D_{r/2}(z_0)) \subset \mathcal{F}$.*

This shows that $R^{-1}(\mathcal{F}) \subset \mathcal{F}$. We need to show equality. Suppose that $z_0 \in \mathcal{F}$ and $R^{n_{k+1}}$ converges uniformly on $\overline{D_{r/2}(z_0)}$. R is non-constant, holomorphic on $D_{r/2}(z_0) \Rightarrow$ an <u>open map</u>. So $R(D_{r/2}(z_0))$ is open and R^{n_k} converges local uniformly on $\underbrace{R(D_{r/2}(z_0))}_{\text{open neighborhood of } R(z_0)}$

$\Rightarrow R(z_0) \in \mathcal{F}$.

So, $R(\mathcal{F}) \subset \mathcal{F}$, and so $R^{-1}(R(\mathcal{F})) = \mathcal{F} \subset R^{-1}(\mathcal{F}) \subset \mathcal{F}$ hence equal.

So $\mathcal{F}$ is completely invariant by the Prop.



$\Rightarrow R(\mathcal{J}) = \mathcal{J}$ is also completely invariant since $\mathcal{J} = \mathcal{F}^c$.

Theorem 13.4. $\forall N \geq 1$, $\mathcal{J}(R) = \mathcal{J}(R^N)$ ***Proof:*** *If R is normal on $D_r(z_0)$*

$\Rightarrow$ *$\{R^{n_k}\}$ converges locally uniformly or rather uniformly on $D_{r/2}(z_0)$. So $R^{n_k} \overset{unif.}{\longrightarrow} f$.*

$\Rightarrow$ *$\underbrace{(R^{n_k})^N}_{=R^{Nn_k}} \longrightarrow f^N$ (fix N)*

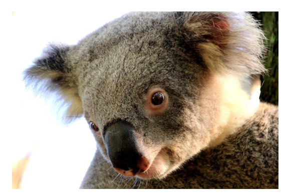
$\Rightarrow$ *The family $\{(R^N)^n\}$ is normal on $D_{r/2}(z_0)$ since $(R^N)^{n_k} = R^{Nn_k}$ converges uniformly.*

$\Rightarrow$ *$\mathcal{F}(R) \subset \mathcal{F}(R^N)$.*

Conversely, if $\{(R^N)^n\}$ is normal on $D_r(z_0)$ so $(R^N)^{n_k}$ converges uniformly on $\overline{D_{r/2}(z_0)}$

$\Rightarrow$ *$R^{Nn_k} = R^{m_k}$ converges uniformly on $\overline{D_{r/2}(z_0)}$*

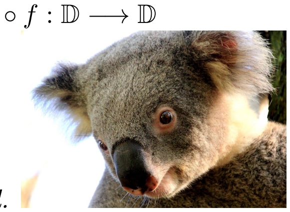
$\Rightarrow$ *$\{R^n\}$ normal there*



$\Rightarrow$ *$\mathcal{F}(R^N) \subset \mathcal{F}(R)$. Hence equal. Hence complements ($\mathcal{J}'s$) are also equal.*

Theorem 13.5 (Montel's Big Theorem (recall))**.** *If $\mathcal{F}$ meromorphic on domain G and $\exists z_1, z_2, z_3$ such that $f(G) \cap \{z_i\}_{i=1}^3 = \emptyset \ \ \forall f \in \mathcal{F}$ then $\mathcal{F}$ is normal.* ***Proof:*** *(Sketch) WLOG $z_1 = 0, z_2 = 1, z_3 = \infty S := \mathbb{C}\backslash\{0,1\}$.*

Uniformization Theorem $\Rightarrow S$ is conformal to $\mathbb{D}$. Let $\phi : S \longleftrightarrow \mathbb{D}$. Then for each $D_r \subset G$, D_r is also conformal to $\mathbb{D}$. WLOG however $D_r = \mathbb{D}$ so in fact we have $\phi \circ f : \mathbb{D} \longrightarrow \mathbb{D}$



$\Rightarrow$ This family is normal. This also (covering maps) $\Rightarrow$ $\mathcal{F}$ is normal.

In addition to fixed points another type of distinctive point is a critical point.

Definition 13.6. A critical point $z \in \hat{\mathbb{C}}$ is a critical point iff R is not injective on any open neighborhood of z iff $R'(z) = 0$ iff the multiplicity of the zero of the function $R(w) - R(z)$ for $w = z$ is greater than one.

Definition 13.7. The multiplicity of $z \in \mathbb{C}$ is the degree of the zero of the function $R(w) - z$ for $w = z$ and is denoted by $\text{mult}(z)$.

We will use but not prove the following theorem.

Theorem 13.8 (Riemann-Hurwitz). *Assume R is not constant and of degree d. Then*

$$\sum_{z \in \hat{\mathbb{C}}} \mathrm{mult}(z) - 1 = 2(d-1).$$

The proof of this theorem relies upon some rather deep results in topology concerning the Euler characteristic of Riemann surfaces. Similar to the proof of the fact that the Julia set of any rational function is non-empty for all rational functions of degree at least two, which relied on the number of fixed points, we can use the Riemann Hurwitz theorem to prove the following.

Theorem 13.9. *Any finite completely invariant set for R rational of degree at least two has at most 2 elements.*

Proof: Assume S is such a set. Then $R(S) = S$, and so R acts as a permutation on the elements of S. Assume S has n elements. Then R is uniquely identified with an element σ of the symmetric group S_n. This group has $n!$ elements hence the order of σ is finite. Let this order be k. This means that R^k acts as the identity element on S. We have already computed that the degree of R^k is d^k where d is the degree of R. Note that the multiplicity of the zero of $R^k(w) - z$ at $w = z$ is d^k. This is because the function $R^k(w) - z$ has precisely d^k zeros counting multiplicity by the Fundamental Theorem of Algebra. Perhaps that is not immediately apparent, but writing

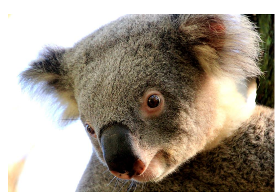
$$R^k(w) = \frac{p(w)}{q(w)}, \quad R^k(w) = z \iff g(w) := p(w) - zq(w) = 0.$$

The function $g(w) : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a polynomial of degree equal to the degree of R^k, which is d^k. Hence this function has precisely d^k zeros counting multiplicity by the Fundamental Theorem of Algebra, and $g(w) = 0$ iff $R^k(w) = z$. So, if one of these zeros were to be some $w \neq z$, then $R^{-k}(z) \ni w$ which shows that $w \in S$ because S is completely invariant. Then since R^k acts as the identity on S, this means that

$$R^k(w) = w \neq z = R^k(w).$$

This is a contradiction. So the only solutions to $R^k(w) - z = 0$ is $w = z$ and hence the multiplicity of z for R^k is d^k. This holds for each $z \in S$. So we have

$$\sum_{z \in S} \mathrm{mult}(z) - 1 = n(d^k - 1) \leq \sum_{z \in \hat{\mathbb{C}}} \mathrm{mult}(z) - 1 = 2(d^k - 1)$$

which shows that $n \leq 2$. 

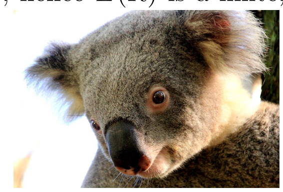
Definition 13.10. The orbit of a point $z \in \hat{\mathbb{C}}$ is

$$\mathcal{O}(z) := \{R^n(z)\}_{n \in \mathbb{Z}}.$$

Note that this includes both the forwards and backwards orbits. If the orbit of a point is finite, then we say that point is exceptional. The set of all such points is denoted by $E(R)$.

Proposition 13.11. *The exceptional set of a rational map of degree at least two has 0, 1, or 2 points.*

Proof: If $z \in E(R)$, then by definition the orbit of z has finitely many elements. Since the orbit of z is the same as the orbit of $R(z)$ as well as the same as the orbit of $R^{-1}(z)$, the orbit is completely invariant. By the preceding theorem the orbit of z has 1 or 2 elements. It has at least one element because it contains $z = R^0(z)$. If the orbit of z contains only z, then it is a fixed point. If the orbit of z also contains w so that $R(z) = w \neq z$, then we know that either $R(R(z)) = R(w) = w$ or $R(R(z)) = z$. Hence either w is a fixed point of R or z is a fixed point of R^2. Consequently the total number of exceptional points is at most twice the number of fixed points of R plus the number of fixed points of R^2. This is finite because R has precisely $d+1$ fixed points, and R^2 has precisely $2d+1$ fixed points. Since the orbit of any exceptional point is completely invariant, and the orbit of any point in the orbit of z is the same as the orbit of z, it follows that the orbit of each exceptional point is contained in $E(R)$. There are finitely many of these, they are each completely invariant, hence $E(R)$ is a finite, completely



invariant set. By the theorem it contains 0, 1, or 2 points.

Theorem 13.12. *The Julia set of any rational map of degree at least two is infinite, and the exceptional set is contained in the Fatou set.*

Proof: If the Julia set is finite, then because it is completely invariant, it contains at most 2 points. We know that the Julia set is not empty. So, first assume the Julia set contains one point. We can conjugate such that WLOG this point is ∞. Then since the Julia set is completely invariant,

$$R(\infty) \subset \mathcal{J} = \infty \implies R(\infty) = \infty,$$

and

$$R^{-1}(\infty) \subset \mathcal{J} = \infty \implies R^{-1}(\infty) = \infty.$$

Consequently, R has no poles in $\mathbb{C}$ and is an entire function. Since it has degree at least two, R is a polynomial. For any polynomial ∞ is a super-attracting fixed point, because 0 is a super-attracting fixed point for the function

$$\frac{1}{R(1/z)} = \phi^{-1} \circ R \circ \phi, \quad \phi(z) = \phi^{-1}(z) = 1/z,$$

and $\phi^{-1}(0) = \infty$. We have already seen that if two functions are conformally conjugate such as $\phi^{-1} \circ R \circ \phi = \widetilde{R}$, then R has a fixed point at ∞ if and only if $\widetilde{R}$ has a fixed point at $\phi^{-1}(\infty) = 0$. Moreover the multiplier at the fixed point is the same for R as for $\widetilde{R}$. Since the polynomial R is of degree $d \geq 2$, $1/R$ tends to 0 of order d as $z \to \infty$ hence $\widetilde{R}$ has a zero of order d at 0. By the Fundamental Theorem of Algebra, $\widetilde{R}$ has precisely d zeros counting multiplicity. Hence this function has only one zero of order d at zero so

$$\frac{1}{R(1/z)} = cz^d, \quad c \in \mathbb{C} \setminus \{0\} \implies R(z) = c^{-1}z^d.$$

Since 0 is a super-attracting fixed point for $\widetilde{R}$ it lies in the Fatou set for $\widetilde{R}$ and consequently $\phi^{-1}(0) = \infty$ also lies in the Fatou set of R. This is a contradiction because this point was assumed to be in the Julia set which is distinct from the Fatou set.

If the Julia set contains two points, we can again assume by conformal conjugation that these points are $\{0, \infty\}$. By the complete invariance of the Julia set we have a few possibilities. One possibility is that $R(0) = 0$ and $R(\infty) = \infty$, which reduces to the above argument which shows that in fact the points 0 and ∞ both lie in the Fatou set, a contradiction. The other possibility $R(0) = \infty$. In this case $R(z) = P(z)/Q(z)$ has a Laurent expansion about 0 of the form $c_j z^{-j} + ...$ with $c_j \neq 0$. Consequently when we consider long division of the polynomials P and Q it follows that the degree of Q is strictly larger than the degree of P. If we were to have $R(\infty) = \infty$, this requires the degree of P to be strictly larger than the degree of Q which it is not. Hence $R(\infty) = 0$ since $R(\infty)$ must be contained in the Julia set by the complete invariance of the Julia set. If there were any other point $p \in \mathbb{C}$ such that $R(p) = 0$, then again by the complete invariance of $\mathcal{J}$ such a point would necessarily be contained in $\mathcal{J}$ which it is not. Hence, the only zero of R is at infinity and this zero must therefore be of degree d which is the degree of R. Consequently $R(z) = cz^{-d}$. Then

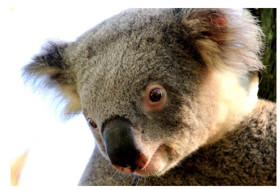
$$R^2(z) = R(R(z)) = c^{1-d} z^{d^2}$$

has a super-attracting fixed point at $z = 0$. It follows that 0 is in the Fatou set of R^2, and by one of our previous results, the Fatou set of R^N is the same as the Fatou set of R for any $N \in \mathbb{N}$. Hence 0 is in the Fatou set of R as well, which is a contradiction because 0 was assumed to be in the Julia set.

So, it is impossible for the Julia set to have 1 or 2 points, and this shows that it must have infinitely many points because it is not empty.

Next we consider the exceptional set. If it is just one point, by conformal conjugation we may assume that this point is ∞. Then the orbit of this point is ∞ and hence $R(\infty) = \infty = R^{-1}(\infty)$ and so R is a polynomial because it is an entire non-constant function with pole at infinity. As we have seen above ∞ is a super-attracting fixed point for any polynomial of degree at least two and hence lies in the Fatou set.

If the exceptional set contains two points, without loss of generality we assume these two points are 0 and ∞. Then we either have $R(0) = 0$, $R(\infty) = \infty$ which implies $R(z) = cz^d$, and both 0 and ∞ are in the Fatou set. By the above argument the other possibility is that $R(\infty) = 0$, $R(0) = \infty$. In this case we showed that $R(z) = cz^{-d}$, and again both 0 and ∞ lie in the Fatou set because this is true for R^2 (both 0 and ∞ are in the Fatou set of R^2 in this case).

So, in all cases the exceptional set lies in the Fatou set. 

Theorem 13.13. *Any completely invariant closed set A satisfies one of the following: either $A \subset E(R) \subset \mathcal{F}$ or $A \supset \mathcal{J}$.*

Proof: Assume A is such a set, and let $U := \hat{\mathbb{C}} \setminus A$. Then U is open and completely invariant. If A is finite, then it has at most two points. It follows that since A is completely invariant, the orbit of each element of A lies in A and hence is finite, so $A \subset E(R)$. If A is infinite, consider $\{R^n\}$ on U. Since U is completely invariant, for each $z \in U$, $R^n(z) \subset U \subset \hat{\mathbb{C}} \setminus A$ and hence the family $\{R^n\}$ on U omits all points of A, of which there are more than three! So, the family R^n is normal on U, and hence $U \subset \mathcal{F}$. The reverse inclusion therefore holds for their complements, so $U^c = A \supset \mathcal{F}^c = \mathcal{J}$. 

Theorem 13.14. *The Julia set is perfect.*

Proof: Let $\mathcal{J}'$ denote the set of accumulation points of the Julia set. Then since $\mathcal{J}$ is closed it follows that $\mathcal{J}' \subset \mathcal{J}$. Note that since $\mathcal{J}$ is infinite and is contained in $\hat{\mathbb{C}}$ which is compact, the Julia set has accumulation points, so $\mathcal{J}' \neq \emptyset$. The idea is thus to show that $\mathcal{J}'$ is completely invariant because then we have proven that any completely invariant closed set is either in the Fatou set or it contains the Julia set. Since $\mathcal{J}'$ is in the Julia set, it cannot be in the Fatou set! First let's show that $\mathcal{J}'$ is closed. If z is an accumulation point of J', then any open neighborhood U of z contains an element of J', which we can call z'. Since z' is in U which is open, and z' is an accumulation point of J, it follows that U also contains an element of J. Hence any open neighborhood U of z contains an element of J, and so z is an accumulation point of $\mathcal{J}$ and therefore $z \in \mathcal{J}'$. Hence, $\mathcal{J}'$ contains all its accumulation points and is therefore closed.

Next we show the complete invariance of $\mathcal{J}'$. Let $z \in \mathcal{J}'$. Then there is a sequence $\{z_n\} \subset \mathcal{J}$ which converges to z. The function R is continuous on $\hat{\mathbb{C}}$, and therefore $R(z_n) \to R(z)$. Since $R(z_n) \in \mathcal{J}$ for every n by the invariance of $\mathcal{J}$, we have a sequence in $\mathcal{J}$ converging to $R(z')$. Therefore $R(z')$ is also an accumulation point of $\mathcal{J}$ and so $R(z') \in \mathcal{J}'$. We have thereby shown the inclusion

$$R(\mathcal{J}') \subset \mathcal{J}' \implies \mathcal{J}' \subset R^{-1}(\mathcal{J}').$$

Next let $z \in R^{-1}(\mathcal{J}')$, and $w = R(z) \in \mathcal{J}'$. Then since R is open, and $w \in \mathcal{J}'$, for an open set U containing z, $R(U)$ is an open point containing w which is an accumulation point of $\mathcal{J}$, and so $R(U)$ has non-empty intersection with $\mathcal{J}$. Therefore

$$R^{-1}(R(U) \cap \mathcal{J}) = U \cap R^{-1}(\mathcal{J}) = U \cap \mathcal{J} \neq \emptyset.$$

So, for any open U containing z, $U \cap \mathcal{J} \neq \emptyset$. It follows that z is an accumulation point of $\mathcal{J}$ and so $z \in \mathcal{J}'$. This shows that

$$R^{-1}(\mathcal{J}') \subset \mathcal{J}' \implies \mathcal{J}' \subset R(\mathcal{J}') \subset \mathcal{J}'.$$

So

$$R(\mathcal{J}') = \mathcal{J}'$$

is completely invariant. Since it is a closed set, by the previous theorem it is either contained in $\mathcal{F}$ or it contains $\mathcal{J}$. Since $\mathcal{J}' \subset \mathcal{J}$ which is disjoint from $\mathcal{F}$, we cannot have $\mathcal{J}' \subset \mathcal{F}$, and so we must have

$$\mathcal{J}' \supset \mathcal{J} \supset \mathcal{J}' \implies \mathcal{J}' = \mathcal{J}.$$

Hence every point of $\mathcal{J}$ is an accumulation point of $\mathcal{J}$ which is the definition of being perfect.

Theorem 13.15. *The Julia set of a rational map R of degree at least two is either $\hat{\mathbb{C}}$ or has empty interior.*

Proof: Let us decompose $\hat{\mathbb{C}}$ as a disjoint union

$$\hat{\mathbb{C}} = \partial\mathcal{J} \cup \mathring{\mathcal{J}} \cup \mathcal{F}.$$

Let us also assume that $z \in \mathring{\mathcal{J}}$, so the interior of $\mathcal{J}$ is not empty. Then there exists $r > 0$ such that $D_r(z) \subset \mathring{\mathcal{J}} \subset \mathcal{J}$. Applying R, by the Open Mapping Theorem, $R(D_r(z)) \ni R(z)$ is an open set. By the complete invariance of $\mathcal{J}$ this set lies in $\mathcal{J}$. Hence there is an open neighborhood of $R(z)$ in $\mathcal{J}$, so $R(z) \in \mathring{\mathcal{J}}$. This shows that

$$R(\mathring{\mathcal{J}}) \subset \mathring{\mathcal{J}}.$$

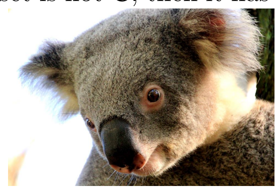
For the reverse inclusion we use continuity, because $R^{-1}(D_r(z))$ is an open set contained in $\mathcal{J}$ hence contained in $\mathring{\mathcal{J}}$ so

$$R^{-1}(\mathring{\mathcal{J}}) \subset \mathring{\mathcal{J}},$$

and we see that $\mathring{\mathcal{J}}$ is completely invariant. Since the Fatou set is also completely invariant, we have the following

$$R(\mathring{\mathcal{J}} \cup \mathcal{F}) = \mathring{\mathcal{J}} \cup \mathcal{F} \implies R(\partial\mathcal{J}) = \partial\mathcal{J},$$

so the boundary of $\mathcal{J}$ is also completely invariant. It is closed since its complement is by definition open. By a preceding result, since the intersection of the Julia set, which is closed and hence contains its boundary, with the Fatou set is empty, either the boundary of the Julia set contains the Julia set, or the boundary of the Julia set is empty. By assumption the Julia set has non-empty interior, so if it has non-empty boundary, then it cannot be contained in its boundary. It follows that the boundary of the Julia set is empty. This means that the Julia set is open as well as closed, and hence is the entire $\hat{\mathbb{C}}$. This shows that if the Julia set has non-empty interior, then it is $\hat{\mathbb{C}}$. On the other hand, if the Julia set is not $\hat{\mathbb{C}}$, by the contrapositive, it cannot have non-empty interior, so if the Julia set is not $\hat{\mathbb{C}}$, then it has empty interior. These are the only two mutually exclusive possibilities.



The following proposition will allow us to prove the self-similarity property of $\mathcal{J}$.

Proposition 13.16. *Let R be a rational map of degree at least two, U a non-empty open set such that $U \cap \mathcal{J} \neq \emptyset$. Then (1)*

$$\cup_{n \geq 0} R^n(U) \supset \hat{\mathbb{C}} \setminus E(R) \supset \mathcal{J},$$

and moreover (2) there exists $N \in \mathbb{N}$ such that

$$R^n(U) \supset \mathcal{J}$$

for all $n \geq N$.

Proof: Well, it makes sense to prove (1) first, because we will likely need it to prove (2) which is a stronger statement. Define

$$U_0 := \cup_{n \geq 0} R^n(U).$$

Define

$$V := \hat{\mathbb{C}} \setminus U_0.$$

If $V = \emptyset$ then we are done. If V has three or more points, we are led to a contradiction because this would mean that the family $\{R^n\}_{n \geq 1}$ on the set U is normal. Then we would have $U \subset \mathcal{F}$ which contradicts the fact that $U \cap \mathcal{J} \neq \emptyset$. So, V has at most 2 points. For the sake of contradiction we assume there is some $z_0 \in V \setminus E(R)$. Then it must have an infinite orbit. We will show that a point has an infinite orbit iff the backwards orbit is infinite. Assume that the backwards orbit is finite,

$$\mathcal{O}^-(z_0) = K = \{z_0, \ldots, z_k\}.$$

Then consider R^{-1} on K. $R^{-1}(z_j)$ is a set of one or more points in K. If two points z_j and z_l have a common pre-image meaning the sets

$$R^{-1}(z_j) \cap R^{-1}(z_l) \neq \emptyset,$$

then applying R to a common point in this pre-image we get that $z_j = z_l$. Hence, for each $j = 0, \ldots k$,

$$R^{-1}(z_j) \subset K$$

is distinct. Each of these sets contains at least one point. Since K is a finite set, this means that each of these pre-images contains *exactly* one point, and so $R^{-1} : K \to K$ is a bijection. It can therefore be identified with a permutation, an element of the group S_{k+1}. This is a group of finite order, so there exists $n \in \mathbb{N}$ such that $(R^{-1})^n = R^{-n}$ acts as the identity on K. Now we consider the forward orbit. For each $z_j \in K$ we have

$$R^{-n}(z_j) = z_j \implies z_j = R^n(z_j)$$

for all $j = 0, 1, ..., n$. In particular $R^n(z_0) = z_0$. Hence

$$R^{n+k}(z_0) = R^k(z_0), \quad \forall k \in \mathbb{N}.$$

Consequently, the forward orbit $O^+(z_0)$ can have at most $n+1$ elements. This shows that if the backward orbit is finite, then the whole orbit is finite. Consequently, if the whole orbit is infinite, then the backwards orbit is infinite. Of course the reverse statement is also true: if the backwards orbit is infinite, then the whole orbit is infinite (because it contains the backward orbit!). So, we have shown the equivalence

$$\#\mathcal{O}^-(z) = \infty \iff \#\mathcal{O}(z) = \infty,$$

where in this statement z is arbitrary.
In our particular case of concern here, we have z_0 not in $E(R)$ hence it has infinite orbit, hence the backwards orbit is infinite. We will use this to achieve a contradiction. First, if some $R^{-m}(z_0) \in U_0$, for some $m \in \mathbb{N}$ then there is some $k \in \mathbb{N} \cup \{0\}$ such that

$$R^{-m}(z_0) \in R^k(U) \implies R^{-m}(z_0) = R^k(w), \quad w \in U.$$

Then applying R^m to both sides,

$$z_0 = R^{m+k}(w) \in R^{m+k}(U) \subset U_0.$$

This contradicts $z_0 \in V = \hat{\mathbb{C}} \backslash U_0$. So, this shows that we must have $R^{-m}(z_0) \ni U_0$ for all $m \in \mathbb{N}$. Since the backwards orbit of z_0 is infinite, there are infinitely many points $R^{-m}(z_0) \in \hat{\mathbb{C}} \setminus U_0$. By definition of U_0, the family of iterates R^n on U omits *all* these points, and there are not just three but infinitely many! By Montel's Theorem the family of iterates is therefore normal on U, so $U \subset \mathcal{F}$ which we have already seen is a contradiction since $U \cap \mathcal{J} \neq \emptyset$.
So, the assumption of a point $z_0 \in V \setminus E(R)$ leads in all cases to a contradiction, hence there can be no such problematic point! This shows that $V \subset E(R)$ and taking complements reverses the inclusion,

$$\hat{\mathbb{C}} \setminus V = U_0 \supset \hat{\mathbb{C}} \setminus E(R) \supset \mathcal{J}.$$

The second statement is rather ingenious. Since we know that the Julia set is infinite and perfect, the intersection $U \cap \mathcal{J}$ is not only empty, but must contain infinitely many distinct points. Choose three distinct points. Since they are all in U which is open, let's call the points for instance z_1, z_2, z_3, and there exist $\epsilon_i > 0$ for $i \in I = \{1, 2, 3\}$ such that $D_{\epsilon_i}(z_i) \subset U$. Moreover we can choose

$$\epsilon = \frac{1}{2} \min\{\epsilon_1, \epsilon_2, \epsilon_3, |z_i - z_j| i \neq j \in I\}.$$

Then $D_\epsilon(z_i) := U_i$ are at a positive distance from each other, have non-empty intersection with $\mathcal{J}$, and are open sets contained in U.

Claim 5. *For each $i \in I$ there exists $j \in I$ and $n \in \mathbb{N}$ such that*

$$U_j \subset R^n(U_i)$$

Proof: By contradiction we assume not. Then there exists an $i \in I$ such that for each $j \in I$ and every $n \in \mathbb{N}$

$$U_j \not\subset R^n(U_i).$$

Hence

$$U_j \not\subset \cup_{n \geq 1} R^n(U_i), \quad j = 1, 2, 3.$$

Since these three sets are disjoint, there exist points in U_j which are not in $\cup_{n\geq 1} R^n(U_i)$, and which are distinct. Hence R^n on U_i omits these three points and is therefore normal. This is again a contradiction because it would imply $U_i \subset \mathcal{F}$ which it is not because $U_i \cap \mathcal{J} \neq \emptyset$.

Claim 6. *There exists $n \in \mathbb{N}$ and $i \in I$ such that*

$$U_i \subset R^n(U_i).$$

Proof: We have shown that there is some $j \in I$ such that

$$U_j \subset R^{n_1}(U_1).$$

If $j = 1$, then the claim is proven. Otherwise, without loss of generality (we can change their names) assume $U_j = U_2$. Then by the previous claim once more, we have some $k \in I$ and $n_2 \in \mathbb{N}$ such that

$$U_k \subset R^{n_2}(U_2).$$

If $k = 2$, the claim is proven. Otherwise, if $k = 1$, then

$$U_1 \subset R^{n_2}(U_2) \subset R^{n_2}(R^{n_1}(U_1)) = R^{n_2+n_1}(U_1),$$

and so in this case the claim is also proven. So, the remaining case is that $k = 3$. Then by the previous claim, there is $l \in I$ and $n_3 \in \mathbb{N}$ such that

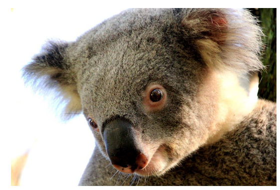
$$U_l \subset R^{n_3}(U_3).$$

If $l = 3$, then the claim is proven. If $l = 2$, then

$$U_2 \subset R^{n_3}(U_3) \subset R^{n_3}(R^{n_2}(U_2)) = R^{n_3+n_2}(U_2),$$

and so the claim is proven. If $l = 2$, then

$$U_1 \subset R^{n_3}(U_3) \subset R^{n_3}(R^{n_2}(U_2)) \subset R^{n_3}(R^{n_2}(R^{n_1}(U_1)))$$
$$= R^{n_3+n_2+n_1}(U_1).$$



So in this case the claim is also proven, and we have proven it in every possible case!
Now we can complete the proof of the proposition, which given the amount of work perhaps ought to be a theorem. For $U_i \subset R^n(U_i)$ as in the claim, let

$$S := R^n.$$

Then S is also a rational map of degree at least two. Since

$$U_i \subset S(U_i) \implies S(U_i) \subset S^2(U_i)$$

we have an increasing sequence

$$U_i \subset S(U_i) \subset \dots S^k(U_i) \subset S^{k+1}(U_i).$$

We have proven that the Julia set of R and any of its iterates R^n are identical. So the Julia set of R is the same as that of S, and we write both as $\mathcal{J}$. By definition of U_i,

$$U_i \cap \mathcal{J} \neq \emptyset,$$

and U_i is open, so by part (1) applied to U_i with respect to S,

$$\mathcal{J} \subset \cup_{n \geq 0} S^n(U_i).$$

On the right side we have an open cover by the open mapping theorem. The Julia set is a closed subset of $\mathbb{C}$ which is compact, hence $\mathcal{J}$ is also compact. Therefore any open cover admits a finite sub-cover and so there is $M \in \mathbb{N}$ such that

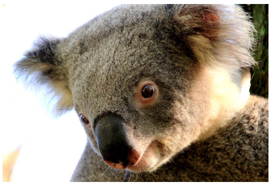
$$\mathcal{J} \subset \cup_{n=0}^{M} S^n(U_i) = S^M(U_i),$$

since $S^n(U_i) \subset S^M(U_i)$ for all $n \leq M$, $n \geq 0$. Note that $S^M = R^{nM}$. So, we have by complete invariance of $\mathcal{J}$ for any $s \in \mathbb{N}$

$$\mathcal{J} = R^s(\mathcal{J}) \subset R^s(R^{nM}(U_i)) = R^{nM+s}(U_i) \subset R^{nM+s}(U),$$

where the last statement follows since $U_i \subset U$. Hence for any $m \geq N := nM$ we have

$$\mathcal{J} \subset R^m(U).$$



The last lovely result to prove is the following.

Theorem 13.17. *The Julia set is self-similar in the sense that for any $z \in \mathcal{J}$,*

$$\mathcal{J} = \overline{\{R^{-n}(z)\}_{n\geq 1}}.$$

Proof: Let $z \in \mathcal{J}$. Then $z \notin E(R) \subset \mathcal{F}$, so the backwards orbit of z is infinite. Let $\epsilon > 0$ and $z_0 \in \mathcal{J}$. Consider $U := D_\epsilon(z_0)$. By the proposition there is $N \in \mathbb{N}$ such that

$$\mathcal{J} \subset R^N(U).$$

Moreover the Julia set is completely invariant which means that $R^{-n}(z) \in \mathcal{J} \subset R^N(U)$. So there exists $w \in U$ such that $R^{-n}(z) = R^N(w)$ and hence $w \in R^{-n-N}(z)$. By definition of $U \ni w$

$$|w - z_0| < \epsilon.$$

This shows that for each $z_0 \in \mathcal{J}$ and $\epsilon > 0$, there is an element of $\mathcal{O}^-(z) = \{R^{-n}(z)\}_{n\geq 1}$ which is at a distance less than ϵ from z_0. Hence $\mathcal{O}^-(z)$ is dense in $\mathcal{J}$. Therefore the closure of this set contains the closure of $\mathcal{J}$ which is equal to $\mathcal{J}$ because $\mathcal{J}$ is closed.

This last result as well as our previous result shows the connection between Julia sets and sets of non-integer Hausdorff dimension. Julia sets have an invariance property, a self-similarity property, and either have empty interior or are the whole space.

References

[1] H. Federer, *Geometric Measure Theory,* Springer, (1969).
[2] G. Folland, *Real Analysis* Second Edition, Wiley-Interscience, (1999).
[3] L. Simon, *Lectures on Geometric Measure Theory,* Proceedings of the Center for Mathematical Analysis, Australian National University, Vol. 3, (1983).