TRAVELING SALESMAN THEOREMS AND THE CAUCHY TRANSFORM

TUOMAS ORPONEN

ABSTRACT. These are the lecture notes for the course Geometric measure theory and singular integrals, given in Spring 2017 at the University of Helsinki. They contain the constructive part of P. Jones’ $L^\infty$ traveling salesman theorem in the plane, following the book of Bishop and Peres, and two proofs of an $L^1$ traveling salesman theorem for doubling measures, due to Badger and Schul (the first proof assumes a quantitative form of non-atomicity from the measure, and follows an argument of Tolsa). As an application of the traveling salesman theorems, the notes contain a proof of the Mattila-Melnikov-Verdera theorem from 1996, on the Cauchy transform and uniform rectifiability. Several additional topics are also discussed:

• David’s theorem, stating that non-atomic measures with bounded Cauchy transform have linear growth,
• the Denjoy conjecture (aka Calderón’s theorem), stating that positive-length subsets of rectifiable curves are non-removable for bounded analytic functions,
• and finally the fact that sufficiently irregular sets, including the four-corners Cantor set, are removable for bounded analytic functions.

CONTENTS

1. Introduction 2
2. Uniform rectifiability and $\beta$-numbers 4
2.1. Various $\beta$-numbers 4
2.2. An overview of traveling salesman theorems 5
3. Preliminaries on compact and (mostly) connected sets 7
4. The $L^\infty$ traveling salesman theorem of P. Jones 11
4.1. Construction and connectedness 12
4.2. Length estimates 15
5. The $L^1$-traveling salesman theorem 18
5.1. Tolsa’s observation 19
5.2. Proof of the $L^1$ traveling salesman theorem for smooth doubling measures 23
5.3. Proof of the $L^1$ traveling salesman theorem for general doubling measures 24
6. Rectifiability of sets and measures, and the Cauchy transform 26
6.1. The theorem of Mattila-Melnikov-Verdera 26
6.2. David’s theorem 29
6.3. The Denjoy conjecture (aka Calderón’s theorem) 31
6.4. Removability of the four corners Cantor set 36
7. The geometric construction of M. Badger and R. Schul 42
7.1. Construction of the continuums 43
8. The connectedness of $\Gamma_k$ 46
9. Length estimates 48
9.1. Virtual credit 50
1. INTRODUCTION

Remark 1.1. The material for the lecture notes has been gathered and combined from various sources, see the list of references. The main sources are the books of Bishop-Peres [2], Falconer [9], Mattila [12], and Tolsa [17], and the article of Badger-Schul [1].

These lecture notes have two main goals. First, to describe and prove various “travelling salesman theorems”, see Section 2.2 for an overview. Second, to explore the connections between rectifiability, the Cauchy transform, and removability. In particular, we prove the following theorem of P. Mattila, M. Melnikov and J. Verdera [13] from 1996:

**Theorem 1.2.** Let $E \subset \mathbb{C}$ be a 1-AD regular set such that the Cauchy transform associated to $H^1|_E$ is bounded on $L^2(H^1|_E)$. Then, the set $E$ is uniformly 1-rectifiable.

The Cauchy integral operator associated to a Radon measure $\mu$ is, formally speaking, the object

$$C_\mu f(z) = \int \frac{f(w)}{z-w} d\mu w.$$  

For more details and results, see Section 6. A "1-AD-regular set" is short for a 1-Ahlfors-David regular set, defined below:

**Definition 1.3** (AD regularity). Let $0 \leq s \leq n$. A Borel set $E \subset \mathbb{R}^n$ is called $s$-Ahlfors-David regular, or $s$-AD regular in short, if there is a constant $M \geq 1$ such that

$$\frac{r^s}{M} \leq H^s(E \cap B(x,r)) \leq Mr^s, \quad x \in E, \ 0 < r \leq \text{diam}(E).$$

More generally, a Borel measure $\mu$ is called $s$-AD regular, if $r^s/M \leq \mu(B(x,r)) \leq Mr^s$ for $x \in \text{spt} \mu$ and $0 < r \leq \text{diam} \text{(spt} \mu)$; thus, a Borel set $E$ is $s$-AD regular, if and only if $H^s|_E$ is $s$-AD regular.

The distinction between AD regular sets and measures is mostly semantic: if $\mu$ is $s$-AD regular with $0 \leq s \leq n$, then $\mu = H^s|_E$ for an $s$-AD regular set $E$. Since only 1-AD regular sets and measures will be considered in these lecture notes, I will abbreviate

AD-regular = 1-AD-regular.

Here is one possible definition of uniform 1-rectifiability:

**Definition 1.4** (Uniform rectifiability). Let $n \geq 2$. A set $E \subset \mathbb{R}^n$ is called uniformly 1-rectifiable, if there is a constant $C \geq 1$ with the following property: for every ball $B \subset \mathbb{R}^n$, the intersection $E \cap B$ can be covered by a continuum $\Gamma_B$ satisfying $H^1(\Gamma_B) \leq C \text{diam}(B)$. A Radon measure $\mu$ is called uniformly 1-rectifiable, if $\text{spt} \mu$ is uniformly 1-rectifiable.

Remark 1.5. The basic example of a uniformly 1-rectifiable set is an AD regular continuum. In fact, if $E$ is 1-AD regular to begin with (as in Theorem 1.2), then it is known (see [7], the discussion at the end of p. 14) that $E$ is uniformly 1-rectifiable, if and only if $E$ is contained in an AD regular continuum. In the plane, and for compact sets, this equivalence is quite easy to prove, even without the a priori AD regularity assumption.
Exercise 1.6. Let $E \subset \mathbb{R}^2$ a uniformly 1-rectifiable compact set. Prove that there exists an AD regular continuum $\Gamma \supset E$ with $\text{diam}(\Gamma) \sim \text{diam}(E)$, where all the implicit constants only depend on $C$. Hint: read Section 3 first.

Remark 1.7. At least two essentially different proofs of Theorem 1.2 are now available: the original from [13], based on curvature, and then a more recent one based on the notion of reflectionless measures, due to B. Jaye and F. Nazarov [10]. In these lecture notes, I take the original route; that said, many of the details are gathered and pieced together from resources more recent than [13].

Remark 1.8. Theorem 1.2 says that if $\mu$ is a priori AD regular, then the $L^2$-boundedness of $C_\mu$ on $L^2(\mu)$ implies that $\mu$ is uniformly 1-rectifiable. It is fair to ask, whether the a priori regularity assumption is sensible. G. David [5] has shown that if $\mu$ is non-atomic to begin with, and $C_\mu$ is bounded on $L^2(\mu),^1$ then

$$\mu(B(x,r)) \leq Cr, \quad x \in \mathbb{R}^2, \ r > 0,$$

where $C \geq 1$ depends on the constants in the $L^2$-boundedness; on the other hand, $C_{\delta_0}$ is nearly trivially bounded on $L^2(\delta_0)$. I postpone the proof of David’s result to Section 6, see Proposition 6.9.

The inequality $\mu(B(x,r)) \leq Cr$ from David’s result is, of course, the “upper” inequality required for AD regularity. The lower regularity is definitely not necessary for the $L^2(\mu)$-boundedness of $C_\mu$. For instance, fix any $s$-AD regular set $E \subset \mathbb{R}^2$ with $s > 1$ and $0 < \mathcal{H}^s(E) < \infty$, and let $\mu = \mathcal{H}^s|_E$. A simple computation shows that

$$\int \frac{d\mu_w}{|z-w|^{q}} \leq C, \quad z \in \mathbb{R}^2,$$

for any $0 \leq q < s$. Now, fix some such $q < s$, and let $p < \infty$ be the dual exponent. Then, cheating a little bit (to be precise, you should do the following for the $\epsilon$-truncations)

$$\|C_\mu(f)\|_{L^p(\mu)} = \int |C_\mu(f)|^p \,d\mu \leq \int \left( \int \frac{|f(w)|}{|z-w|} \,d\mu_w \right)^p \,d\mu_z \leq \int \int |f(w)|^p \,d\mu_w \left( \int \frac{d\mu_w}{|z-w|^{q}} \right)^{p/q} \,d\mu_z \lesssim_E \int |f(w)|^p \,d\mu_w.$$

So, $C_\mu$ is bounded on $L^p(\mu)$. It is also easy to see that $C_\mu$ is bounded on $L^q(\mu)$: if $f \in L^q(\mu)$ and $g \in L^p(\mu)$, then

$$\left| \int C_\mu(f) \cdot g \,d\mu \right| \leq \|f\|_{L^q(\mu)} \|C_\mu(g)\|_{L^p(\mu)} \lesssim \|f\|_{L^q(\mu)} \|g\|_{L^p(\mu)}$$

by the previous computation, and now $\|C_\mu\|_{L^q(\mu) \to L^p(\mu)} < \infty$ by duality. Finally, standard Marcinkiewicz interpolation gives $\|C_\mu\|_{L^2(\mu) \to L^2(\mu)} < \infty$.

So, in conclusion, the upper AD regularity condition $\mu(B(x,r)) \lesssim r$ is necessary in Theorem 1.2, for non-atomic measures, whereas the lower inequality $\mu(B(x,r)) \gtrsim r$ is there to make life interesting. For measures decaying much more rapidly than $O(r)$, the problem is too easy.

---

1 In the the usual sense that all $\epsilon$-truncations are uniformly bounded on $L^2(\mu)$, see Section 6.
2. Uniform Rectifiability and β-Numbers

To prove Theorem 1.2, one needs to develop the theory of uniformly 1-rectifiable sets. One of the seminal results in this theory was the Analyst's traveling salesman theorem of P. Jones from 1990 [11], which provided a useful "multi-scale" characterisation of uniform 1-rectifiability (the terminology "uniformly rectifiable" was coined shortly afterwards by G. David and S. Semmes in another seminal paper [7]). Jones' characterisation is formulated in terms of β-numbers, which I will now discuss.

2.1. Various β-numbers. Let μ be a Radon measure on \( \mathbb{R}^n \). In the typical application, \( \mu = \mathcal{H}^1|_E \) for a set \( E \) with positive and σ-finite 1-dimensional measure. For \( p \in [1, \infty) \), a compact set \( B \) (which will always be a ball or a cube) and a straight line \( \ell \subset \mathbb{R}^n \), write

\[
\beta_{\mu,p}(B, \ell) := \left[ \int_B \left( \frac{\text{dist}(x, \ell)}{\text{diam}(B)} \right)^p \frac{d\mu x}{\mu(B)} \right]^{1/p},
\]

where we agree that \( \beta_{\mu,p}(B, \ell) = 0 \), if \( \mu(B) = 0 \). Then, define

\[
\beta_{\mu,p}(B) := \inf_{\text{lines } \ell} \beta_{\mu,p}(B, \ell).
\]

It is clear from Hölder's inequality that

\[
\beta_{\mu,p_1}(B) \leq \beta_{\mu,p_2}(B), \quad 1 \leq p_1 \leq p_2 < \infty.
\]

(2.1)

How about \( p = \infty \)? The definition practically writes itself: again, for a line \( \ell \subset \mathbb{R}^n \), define the auxiliary number

\[
\tilde{\beta}_{\mu,\infty}(B, \ell) := \mu - \text{ess sup}_{x \in B} \frac{\text{dist}(x, \ell)}{\text{diam}(B)},
\]

and then set

\[
\tilde{\beta}_{\mu,\infty}(B) := \inf_{\text{lines } \ell} \tilde{\beta}_{\mu,\infty}(B, \ell).
\]

It is clear from (2.1) that the numbers \( \beta_{p,\mu}(B) \) tend to a limit as \( p \to \infty \), and the limit is bounded by \( \tilde{\beta}_{\mu,\infty}(B) \).

Exercise 2.2. Prove or disprove:

\[
\tilde{\beta}_{\mu,\infty}(B) = \lim_{p \to \infty} \beta_{\mu,p}(B).
\]

There is another fairly natural definition for \( \tilde{\beta}_{\mu,\infty} \), which will be used more in in these lecture notes (both for historical reasons, and for convenience). For any set \( E \subset \mathbb{R}^n \) (such as \( E = \text{spt } \mu \)), write

\[
\beta_{E,\infty}(B, \ell) := \sup_{x \in B \cap E} \frac{\text{dist}(x, \ell)}{\text{diam}(B)}.
\]

The number \( \beta_{E,\infty}(B) \) is then defined in the obvious way. Thanks to the following inequalities, it makes little practical difference, which convention for \( \beta_\infty \) is used:

\[
\tilde{\beta}_{\mu,\infty}(B) \leq \beta_{\text{spt } \mu,\infty}(B) \lesssim_\lambda \tilde{\beta}_{\mu,\infty}(\lambda B), \quad \lambda > 1,
\]

(2.3)

where

\[
\lambda B = \{ x : \text{dist}_\infty(x, B) \leq (\lambda - 1) \text{diam}_\infty(B) \}, \quad \lambda \geq 1.
\]

(2.4)

Here \( \text{dist}_\infty \) and \( \text{diam}_\infty \) refer to distance and diameter in the \( L_\infty \)-distance \( \|x - y\|_\infty = \max\{|x_i - y_i| : 1 \leq i \leq n\} \). This detail will be convenient in the sequel, where the
notation is mostly applied to cubes: with the current definition, $\lambda Q$ remains a cube for all $\lambda \geq 1$, see Figure 1.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{The cubes $Q, 2Q$ and $3Q$ with our convention of "$\lambda E$".}
\end{figure}

2.2. An overview of traveling salesman theorems.

2.2.1. Jones’ traveling salesman theorem for $\beta_\infty$-numbers. As mentioned above, this is where it all started in 1990:

**Theorem 2.5 (Jones).** Let $\mathcal{D}$ be the family of closed\footnote{This distinction makes no different in this theorem, but it will be useful later, and I want to keep the same notation everywhere.} dyadic cubes in $\mathbb{R}^n$, and let $E \subset \mathbb{R}^n$ be a compact set satisfying

$$\beta_{E,\infty}^2(2Q)\ell(Q) < \infty.$$  \hfill (2.6)

Then, for any $\delta > 0$, there exists a compact connected set $\Gamma \subset \mathbb{R}^n$ such that $E \subset \Gamma$, and

$$\mathcal{H}^1(\Gamma) \leq (1 + \delta) \text{diam}(E) + C_\delta \beta_{\infty}^2(E).$$

The proof of Jones’ theorem in the plane is contained in Section 4. In fact, Jones also proved the converse for $n = 2$, using complex analysis. The result was generalised to higher dimensions, with a different, geometric proof, by K. Okikiolu \cite{Okikiolu} a bit later: for rectifiable curves $\Gamma$ of finite length, the sum $\beta_{\infty}^2(\Gamma)$ is bounded by $\lesssim \mathcal{H}^1(\Gamma)$. Jones also observed in \cite{Jones} that if, in place of (2.6), the $\beta_\infty$-numbers satisfy the following Carleson condition,

$$\sum_{Q \subset R} \beta_{E,\infty}^2(2Q)\ell(Q) \lesssim \ell(R), \quad R \in \mathcal{D},$$  \hfill (2.7)

with implicit constants independent of $R$ of course, then $E$ can be covered by an AD regular continuum: in particular, $E$ is uniformly 1-rectifiable. Note that the Carleson condition (2.7) can sometimes hold, even if the full sum in (2.6) diverges: this is, for instance, the case for unbounded AD regular curves (unless they happen to be lines, or otherwise sufficiently flat at infinity).
2.2.2. Traveling salesman theorems for \( \beta_p \)-numbers. The \( L^p \)-versions of the \( \beta \)-numbers do not make sense for sets, unless there is some canonical measure supported on the set. For measures \( \mu \), however, it is reasonable to ask whether a condition for the \( \beta_{\mu,p} \)-numbers \( p \in [1, \infty] \), analogous to either (2.6) or (2.7) gives some geometric information about the support of \( \mu \). The answer is positive, to a certain extent, and this was one of the main results in David and Semmes’ paper [7], where the notion of uniform rectifiability was first introduced.

**Theorem 2.8 (David-Semmes).** Assume that \( E \subset \mathbb{R}^n \) is a 1-AD regular set, let \( \mu := \mathcal{H}^1|_E \), and let \( p \in [1, \infty] \). Assume that the \( \beta_{\mu,p} \)-numbers satisfy the Carleson condition

\[
\sum_{Q \subset R} \beta_{\mu,p}^2(2Q) \ell(Q) \lesssim \ell(R), \quad R \in \mathcal{D}.
\]  

(2.9)

Then \( E \) is uniformly 1-rectifiable (in fact, \( \text{spt } \mu \) can be covered by a 1-AD regular continuum).

Using the inequalities (2.3), the case \( p = \infty \) reduces easily to the previous result of Jones, but the cases \( p \in [1, \infty) \) are \textit{a priori} harder, because a "cube-wise” comparison \( \beta_{\mu,p}(Q) \sim \beta_{\mu,\infty}(Q) \) is not true for \( p < \infty \).\(^3\) However, for 1-AD regular measures (and more generally "smooth” measures, to be introduced in Section 5), these cases can also be reduced to the case \( p = \infty \) via the following trick (Theorem 7.52 in Tolsa’s book [17]):

**Theorem 2.10 (Tolsa).** Let \( \mu \) be a smooth Radon measure in \( \mathbb{R}^n \). Then

\[
\sum_{Q \subset R} \beta_{\text{spt } \mu, \infty}(2Q) \ell(Q) \lesssim \sum_{Q \subset 2R} \beta_{\mu,p}^2(3Q) \ell(Q)
\]

for any cube \( R \in \mathcal{D} \), and any \( p \in [1, \infty) \).

Recall that a Radon measure \( \mu \) on \( \mathbb{R}^n \) is called \textit{doubling}, if there exists a constant \( D_\mu \geq 1 \) such that

\[\mu(B(x, 2r)) \leq D_\mu \mu(B(x, r)), \quad x \in \text{spt } \mu, \quad r > 0.\]

The constant \( D_\mu \) is called the \textit{doubling constant} of \( \mu \). It turns out that Theorem 2.9 holds for all doubling measures,\(^4\) and the following results are some of the main topics of these lecture notes. They are due to M. Badger and R. Schul [1] from 2016.

**Theorem 2.11 (Badger-Schul).** Let \( \mu \) be a doubling measure on \( \mathbb{R}^n \) with compact support \( E := \text{spt } \mu \), let \( p \in [1, \infty) \), and assume that the numbers \( \beta_{\mu,p} \) satisfy

\[
\beta_p^2(\mu) := \sum_{Q \in \mathcal{D}_E \cap Q \subset \lambda E} \beta_{\mu,p}^2(2Q) \ell(Q) < \infty,
\]

where \( \lambda = \lambda_n \geq 1 \) is a sufficiently large constant, and \( \mathcal{D}_E = \{ Q \in \mathcal{D} : Q \cap E \neq \emptyset \} \). Then \( E \) can be covered by a continuum \( \Gamma \subset \mathbb{R}^n \) with

\[\mathcal{H}^1(\Gamma) \lesssim \text{diam}(E) + \beta_p^2(\mu).\]

\(^3\)One can check that if \( E \) is AD regular, and \( \mu = \mathcal{H}^1|_E \), then \( \beta_{E,\infty}(2Q) \lesssim \beta_{\mu,1}(\lambda Q)^{1/2} \). This bound is sometimes quite useful, but not good enough for direct application to the traveling salesman problem.

\(^4\)It is an open research topic, which are the “minimal” \textit{a priori} assumptions on a measure, so that summability of the \( \beta \)-numbers implies rectifiability, as in Theorem 2.9. Without any \textit{a priori} assumptions, the situation does not look very promising at the moment [14].
Corollary 2.12. Assume that $\mu$ is a doubling measure on $\mathbb{R}^n$ and $p \in [1, \infty]$. If the numbers $\beta_{\mu,p}$ satisfy the Carleson condition
\[
\sum_{Q \in \mathcal{D}_{\text{opt},\mu}} \beta_{\mu,p}^2(2Q)\ell(Q) \lesssim \ell(R), \quad R \in \mathcal{D},
\]
then $\mu$ is uniformly 1-rectifiable.

Proof. The case $p = \infty$ already follows from the work of Jones and does not require doubling from $\mu$. The case $p \in [1, \infty)$ uses Theorem 2.11 as follows. Fix a ball $B \subset \mathbb{R}^n$. By Theorem 2.11, the intersection $B \cap (\text{spt } \mu)$ can be covered by a continuum $\Gamma_B$ of length
\[
\mathcal{H}^1(\Gamma_B) \lesssim \text{diam}(B \cap (\text{spt } \mu)) + \sum_{Q \in \mathcal{D}_{\text{opt},\mu}} \beta_{\mu,p}^2(2Q)\ell(Q) \lesssim \text{diam}(B).
\]

The last inequality follows from the Carleson condition, and the fact that $\lambda B$ can be covered by $\lesssim 1$ cubes $R \in \mathcal{D}$ with $\ell(R) \sim \text{diam}(B)$. \qed

Section 5 contains two proofs for Theorem 2.11: The first one only works for "smooth" measures (which means "doubling + quantitatively non-atomic"), in which case the result can be reduced to Jones’ $L^\infty$ traveling salesman theorem via Tolsa’s trick, Theorem 2.10. The second one is the original proof by Badger and Schul, and works for all doubling measures.

3. PRELIMINARIES ON COMPACT AND (MOSTLY) CONNECTED SETS

This section is not strictly necessary for the sequel, but it would be odd to read about traveling salesman theorems without knowing the material here. I take the following result for granted (see [8, Exercise 6.3.12]):

Theorem 3.1. A connected set $\Gamma \subset \mathbb{R}^n$ with finite length is arcwise connected: for every pair $x, y \in \Gamma$ there exists an injective curve $\psi([0, 1]) \subset \Gamma$ such that $\psi(0) = x$ and $\psi(1) = 1$.

The rest of the material from this section is from Falconer’s book [9]. Let’s recall the Hausdorff metric on non-empty compact subsets of $\mathbb{R}^n$. For $K \subset \mathbb{R}^n$ and $\delta > 0$, write $K(\delta) := \{x \in \mathbb{R}^n : \text{dist}(x, K) < \delta\}$. Then, for non-empty compact sets $K_1, K_2 \subset \mathbb{R}^n$, let
\[
d_H(K_1, K_2) = \inf\{\delta > 0 : K_1 \subset K_2(\delta) \text{ and } K_2 \subset K_1(\delta)\}.
\]
Then (exercise, if this is news) $d_H$ is a metric on non-empty compact subsets of $\mathbb{R}^n$. A very useful result is that "the set of compact sets is a compact space" (at least almost):

Theorem 3.2 (Blaschke selection theorem). Let $\mathcal{F}$ be an infinite family of non-empty compact sets in $\mathbb{R}^n$, all lying in some fixed closed ball $B \subset \mathbb{R}^n$. Then, there exists a sequence of distinct sets $\{K_j\}_{j \in \mathbb{N}} \subset \mathcal{F}$, and a non-empty compact set $K \subset B$ such that $K_j \to K$ in the Hausdorff metric.

Remark 3.3. The conclusion would be pretty obvious without the requirement that that the sets $K_j$ be distinct. Also, the family $\mathcal{F} = \{\{k\} \subset \mathbb{R} : k \in \mathbb{N}\}$ shows that the hypothesis involving $B$ is necessary.
Proof. Induction: let \( \{ K^1_j \} \subset \mathcal{F} \) be any sequence of distinct sets, and assume that the sequence \( \{ K^m_j \} \) has already been defined for some \( m \geq 1 \). Define a subsequence \( \{ K^{m+1}_j \} \subset \{ K^m_j \} \) as follows. Cover \( B \) by a finite number \( B^m \) of balls of diameter \( 1/m \). Then, every set \( K^m_j \) intersects every ball in some finite sub-collection \( B^m_j \); there are only finitely many different sub-collections, and infinitely many distinct sets, so there must be a fixed sub-collection \( B^m \) such that \( B^m = B^m_j \) for infinitely many indices \( j \). Define the subsequence \( \{ K^{m+1}_j \} \) by picking only those indices \( j \). Then, it is clear that

\[
d_H(K^m_i, K^{m+1}_j) \leq \frac{2}{m}, \quad i, j \in \mathbb{N},
\]

because \( K^{m+1}_i \) and \( K^{m+1}_j \) both intersect all the balls in \( B^m \), and are covered by them.

Because \( \{ K^{m+p}_j \} \) is always a subsequence of \( \{ K^m_j \} \), it follows that

\[
d_H(K^{m+p}_j, K^m_j) \leq \frac{2}{m}
\]

for all \( i, j \) and \( p \geq 0 \). Now, let \( K_m := K^m_m \). From (3.4), one sees that \( d_H(K_m, K_{m+p}) \leq 2/m \), which implies that \( \{ K_m \}_{m \in \mathbb{N}} \) is a Cauchy sequence in \( d_H \). So, it remains to show that \( d_H \) is a complete metric.

Define

\[
K := \bigcap_{m \geq 0} \bigcup_{k \geq m} K_k =: \bigcap_{m \geq 0} E_m.
\]

This is clearly a non-empty compact set. Further, since \( K_{m+p} \subset K_m(2/m) \) for all \( p \geq 0 \), it follows that \( E_{m+p} \subset K_m(2/m) \) for \( p \geq 0 \), and consequently

\[
K \subset K_m(2/m), \quad m \geq 0.
\]

If the converse inclusion were also true, then \( d_H(K, K_m) \leq 3/m \), and the proof would be complete. So, it suffices to check that the converse inclusion is true.

Pick \( x \in K_m \). Then \( x \in K_{m+p}(2/m) \) for all \( p \geq 0 \), so also \( x \in E_{m+p}(2/m) \). Now, it suffices to choose a sequence \( \{ y_{m+p} \} \) with \( y_{m+p} \in E_{m+p} \) with \( |x - y_{m+p}| \leq 2/m \). The sequence has a subsequence convergent to a point \( y \in K \) satisfying \( |x - y| \leq 2/m \). This proves that \( K_m \subset K(2/m) \).

A tree \( T \) is a continuum without loops: that is, for every pair \( x, y \in T \), there is a unique path \( \gamma([0, 1]) \subset \Gamma \) with \( \gamma(0) = x \) and \( \gamma(1) = y \). For \( x, y \in T \), let \( d_\gamma(x, y) \) be the path distance between \( x, y \); \( d_\gamma(x, y) = H^1(\gamma_{x,y}) \), where \( \gamma_{x,y} \) is the unique path joining \( x \) to \( y \) in \( T \). This is a metric on \( T \), and satisfies \( d_\gamma(x, y) \geq |x - y| \). The diameter of a set \( K \subset T \) in the \( d_\gamma \) metric is denoted by \( \text{diam}_\gamma(K) \).

For the purposes below, the key feature of trees is the following: they can be easily chopped into pieces of smaller diameter preserving connectedness. That would not be so easy for arbitrary continua.

Lemma 3.5 (Tree-chopping lemma). Let \( T \subset \mathbb{R}^n \) be a tree with finite length. Then, given \( \delta > 0 \), we can express \( T \) as the union of \( H^1 \)-essentially disjoint sub-trees \( T_1, \ldots, T_m \) with the following properties:

(a) \( \text{diam}(T_k) \leq \min\{\delta, H^1(T_k)\} \) for \( 1 \leq k \leq m \),

(b) \( m \leq H^1(T)/\delta + 1 \).
Remark 3.6. The inequality $\text{diam}(T_k) \leq \mathcal{H}^1(T_k)$ follows simply from the fact that $T_k$ is connected. Indeed, consider any connected set $\Gamma \subset \mathbb{R}^n$, and fix $x, y \in \Gamma$. Then consider the map $\pi_x(z) = |z - x|$. It is clear that $\pi_x$ is 1-Lipschitz, and from the connectedness of $\Gamma$ it follows easily that $\pi_x(\Gamma) \supset [0, |x - y|]$. Hence,

$$|x - y| = \mathcal{H}^1([0, |x - y|]) \leq \mathcal{H}^1(\pi_x(\Gamma)) \leq \mathcal{H}^1(\Gamma),$$

and now the inequality $\text{diam}(\Gamma) \leq \mathcal{H}^1(\Gamma)$ follows by taking a sup on the left hand side.

Proof of Lemma 3.5. If $\text{diam}(T) \leq \delta$, there is nothing to prove. So, assume $\text{diam}_\gamma(T) \geq \text{diam}(T) > \delta$ (the first inequality follows by an argument similar to the one in Remark 3.6). Fix any point $y_0 \in T$, and let $M_0 := \sup\{d_\gamma(z, y) : z \in T\}$. Note that

$$\frac{\delta}{2} < M_0 \leq \text{diam}_\gamma(T),$$

because otherwise $\text{diam}_\gamma(T) \leq 2M_0 \leq \delta$. Fix any point $z_0 \in T$ with $d_\gamma(y_0, z_0) \geq M_0 - \frac{\delta}{6}$, and finally fix $x \in \gamma_{y_0, z_0}$ with $d_\gamma(y_0, x) = M_0 - \frac{\delta}{6}$.

Next, consider the equivalence relation $\sim$ on $T \setminus \{x\}$:

$$v \sim w \iff x \notin \gamma_{v, w}.$$

The equivalence class of $v \in T \setminus \{x\}$ is denoted by $[v]$. Let

$$T^x := \{x\} \cup \{v \in T \setminus \{x\} : v \not\sim y\}.$$

Then $T^x$ is a tree. To see this, consider a pair of points $v, w \in T^x$. Then $\gamma_{x, v} \setminus \{x\} \subset [v]$ and $\gamma_{x, w} \setminus \{x\} \subset [w]$, because clearly every pair of points in $\gamma_{x, v} \setminus \{x\}$ (resp. $\gamma_{x, w} \setminus \{w\}$) can be joined to $v$ (resp. $w$) without passing through $x$. It follows that

$$\gamma_{v, x} \cup \{x\} \cup \gamma_{x, w} \subset [v] \cup \{x\} \cup [w] \subset T^x,$$

which implies that $T^x$ is path connected, hence a tree (uniqueness is inherited from $T$).

I claim that $\text{diam}(T^x) \leq \text{diam}_\gamma(T^x) \leq \delta$. To see this, note that if $v \in T^x$, then $v \not\sim y_0$, which implies that $x \in \gamma_{y_0, y_0}$, and consequently

$$M_0 \geq d_\gamma(x, y_0) = d_\gamma(y_0, x) + d_\gamma(x, v) = (M_0 - \frac{\delta}{6}) + d_\gamma(x, v).$$

This gives $d_\gamma(x, v) \leq \frac{\delta}{2}$, and so $\text{diam}_\gamma(T^x) \leq \delta$. Since obviously $\text{diam}_\gamma(T^x) \leq \mathcal{H}^1(T^x)$, we see that $T^x$ is a tree satisfying (a). Furthermore, $\mathcal{H}^1(T^x) \geq \frac{\delta}{3}$. Indeed, since $x \in \gamma_{y_0, y_0}$, one has $z_0 \not\sim y_0$, and consequently $\gamma_{x, z_0} \subset T^x$. This implies that

$$\mathcal{H}^1(T^x) \geq d_\gamma(x, z_0) = d_\gamma(y_0, z_0) - d_\gamma(y_0, x) \geq (M_0 - \frac{\delta}{6}) - (M_0 - \frac{\delta}{6}) = \frac{\delta}{3}. \quad (3.7)$$

Finally, observing that $(T \setminus T^x) \cup \{x\} = [y_0] \cup \{x\}$ is also a tree, one just needs to iterate the construction, chopping off another tree from $T \setminus T^x$. Since the part removed always has measure $\geq \frac{\delta}{3}$ by (3.7), the number of iterations is bounded by $\leq \mathcal{H}^1(T)/\delta + 1$, as claimed in (b). The proof is complete.

Now, for the main result of the section:

**Theorem 3.8 (Lower semicontinuity of length of continuums).** Let $\{\Gamma_k\}_{k \in \mathbb{N}}$ be a sequence of compact continua in $\mathbb{R}^n$, convergent in the Hausdorff metric to a compact space $\Gamma$. Then $\Gamma$ is a continuum, and

$$\mathcal{H}^1(\Gamma) \leq \liminf_{k \to \infty} \mathcal{H}^1(\Gamma_k). \quad (3.9)$$
Proof. If $\Gamma$ were not connected, there would be open disjoint sets $U_1, U_2$ with $\Gamma \subset U_1 \cup U_2$. Then, it follows from compactness, and the definition of Hausdorff convergence, that $\Gamma_k \subset U_1 \cup U_2$ for sufficiently large $k$, a contradiction.

To prove (3.9), one may assume that $\mathcal{H}^1(\Gamma_k) \leq C < \infty$ for all $k \in \mathbb{N}$, and also that the full sequence of numbers $\mathcal{H}^1(\Gamma_k)$ converges to the value on the right hand side of (3.9), say $L \leq C$ (because the following considerations are valid for any subsequence). For each $k \in \mathbb{N}$, choose a finite subset $S_k \subset \Gamma_k$ so that
d_{H}(S_k, \Gamma) \to 0
as $k \to \infty$. Since $\Gamma_k$ is arcwise connected, there exist trees $T_k$ such that $S_k \subset T_k \subset \Gamma_k$. To see this, declare any singleton $\{s_0\} \subset S_k$ as an initial tree $T_k^0$. Then, assume that $T_k^j$ has been constructed for some $j$, consisting of finitely many arcs, and assume that at least one $s_{j+1} \in S_k$ yet lies outside $T_k^j$. Connect $s_{j+1}$ to any point of $T_k^j$ by an arc $\gamma = \gamma([0,1]) \subset \Gamma_k$, with $\gamma(0) = s_{j+1}$. Then there exists a smallest number $t \in (0,1]$ such that $\gamma(t) \in T_k^j$, and now $T_k^{j+1} := T_k^j \cup \gamma([0,t])$ is a tree containing $s_{j+1}$.

It is clear that
d_{H}(T_k, \Gamma) \to 0
as $k \to \infty$. Fix $\delta > 0$, and decompose every tree $T_k$ as a $\mathcal{H}^1$-essentially disjoint union

$$T_k = \bigcup_{j=1}^{m_k} T_{k,j},$$

where $\text{diam}(T_{k,j}) \leq \min\{\delta, \mathcal{H}^1(T_{k,j})\}$ and $m_k \leq C/\delta + 1$. Without loss of generality, one may assume that $m_k = m$ for all $k$ (there are only finitely many choices for $m_k$, so this is anyway true after passing to a subsequence).

By the Blaschke selection theorem, every sequence $\{T_{k,j}\}_{k \in \mathbb{N}}$, $1 \leq j \leq m$, has a subsequence convergent in the Hausdorff metric to a non-empty compact continuum $\Gamma^j \subset \Gamma$; by re-indexing appropriately, and possibly finding subsequences inside subsequences, one may assume that the full sequences converge. It is clear that $\text{diam}(\Gamma^j) \leq \delta$ and $\Gamma \subset \bigcup \Gamma^j$. It follows from the definition of $\mathcal{H}^1_{\delta}$, and the $\mathcal{H}^1$ essential disjointness of the trees $T_{k,j}$, $1 \leq j \leq m$, that

$$\mathcal{H}^1_{\delta}(\Gamma) \leq \sum_{j=1}^{m} \text{diam}(\Gamma_j) = \limsup_{k \to \infty} \sum_{j=1}^{m} \text{diam}(T_{k,j}) \leq \limsup_{k \to \infty} \sum_{j=1}^{k} \mathcal{H}^1(T_{k,j}) \leq \lim_{k \to \infty} \mathcal{H}^1(\Gamma_k) = L.$$

Letting $\delta \to 0$ proves the theorem. \qed

Remark 3.10. Note how the connectedness of the trees $T_{k,j}$ was used in the second-to-last inequality. This estimate would not work, if the sets $\Gamma_j$ were chopped up into smaller pieces with some dumber procedure.

The following corollary will be useful in the sequel:

**Corollary 3.11** (Existence of a shortest covering continuum). Assume that $K \subset \mathbb{R}^n$ is a compact set, and there exists a continuum $\Gamma_0 \supset K$ of finite length. Then, there exists a continuum $\Gamma \supset K$ of finite and minimal length.

**Proof.** Choose a sequence of continuums $(\Gamma_k)_{k \in \mathbb{N}}$ with $\Gamma_k \supset E$, and such that

$$\mathcal{H}^1(\Gamma_k) \to \inf\{\mathcal{H}^1(\Gamma) \colon \Gamma \text{ is a continuum with } \Gamma \supset K\} =: m_K < \infty.$$
One may clearly assume that all the continuums are contained in a sufficiently large ball containing \( K \), so the Blaschke selection theorem is applicable: there exists a subsequence \((k_j)_{j \in \mathbb{N}}\) and a compact set \( \Gamma \) with \( d_H(\Gamma_{k_j}, \Gamma) \to 0 \). It is now easy to check that \( K \subset \Gamma \). By Theorem 3.8, moreover, the set \( \Gamma \) is a continuum, and satisfies
\[
m_K \leq \mathcal{H}^1(\Gamma) \leq \liminf_{k \to \infty} \mathcal{H}^1(\Gamma_k) = m_K.
\]
This proves the corollary.

4. THE \( L^\infty \) TRAVELING SALESMAN THEOREM OF P. JONES

This section contains the proof of Peter Jones’ original traveling salesman theorem for the numbers \( \beta_{\infty} \), but only in the plane. Recall the statement:

**Theorem 4.1 (Jones).** Let \( \mathcal{D} \) be family of closed dyadic cubes in \( \mathbb{R}^n \), and let \( E \subset \mathbb{R}^n \) be a compact set satisfying
\[
\beta_{\infty}^2(E) := \sum_{Q \in \mathcal{D}_E, Q \subset 3E} \beta_{E, \infty}^2(2Q)\ell(Q) < \infty,
\]
where \( \mathcal{D}_E = \{Q \in \mathcal{D} : Q \cap E \neq \emptyset \} \). Then, for any \( \delta > 0 \), there exists a compact connected set \( \Gamma \subset \mathbb{R}^n \) such that \( E \subset \Gamma \), and
\[
\mathcal{H}^1(\Gamma) \leq (1 + \delta) \text{diam}(E) + C_\delta \beta_{\infty}^2(E). \tag{4.2}
\]

**Exercise 4.3.** Is it possible to eliminate the \( \delta > 0 \) altogether?

**Definition 4.4 (Convex hulls and extreme points).** The convex hull of a bounded set \( K \subset \mathbb{R}^2 \), denoted by \( \text{conv}(K) \), is the minimal (relative to inclusion) convex set \( R \subset \mathbb{R}^2 \) with \( K \subset R \). Such a set exists, and in fact
\[
\text{conv}(K) = \bigcap_{K \subset R, R \text{ convex}} R.
\]
It is useful to note that \( \text{conv}(K) \) can also be expressed as the set of all (finite) convex combinations of points in \( K \). That is,
\[
\text{conv}(K) = \left\{ \sum_{k=1}^m \lambda_k x_k : m \in \mathbb{N}, \lambda_j \in [0, 1], x_j \in K \text{ and } \sum_{k=1}^m \lambda_k = 1 \right\}. \tag{4.5}
\]
To see this, simply note that the set on the left hand side is convex and contains \( K \), and is clearly contained in any set with these properties.

A convex combination as on the right hand side of (4.5) is called non-trivial, if \( \lambda_j < 1 \) for all \( 1 \leq j \leq m \). The set of extreme points of \( K \subset \mathbb{R}^2 \), denoted by \( \text{Ex}(K) \), are those points in \( K \), which cannot be expressed as non-trivial convex combinations of elements in \( K \). In other words \( x \in \text{Ex}(K) \), if and only if \( x \in K \), and the following holds: if \( x \) has a representation
\[
x = \sum_{k=1}^m \lambda_k x_k, \quad x_j \in K,
\]
with \( 0 \leq \lambda_j \leq 1 \) and \( \sum_{k=1}^m \lambda_k = 1 \), then \( \lambda_j = 0 \) for all but one index \( j = j_0 \), and \( x_{j_0} = x \).

It is slightly, but not very, non-trivial but true that the convex hull of a compact set is a compact set:
Lemma 4.6. If $K \subset \mathbb{R}^2$ is compact, then $\text{conv}(K)$ is compact.

Proof. The details are contained in [16, Theorem 3.25]. The main idea is the following: if $K \subset \mathbb{R}^2$ is bounded, then $\text{conv}(K)$ can be expressed as

$$\text{conv}(K) = f(S \times K \times K \times K),$$

where $S \subset \mathbb{R}^3$ is the (compact) simplex $S = \{ (\lambda_1, \lambda_2, \lambda_3) : \lambda_i \in [0, 1] \text{ and } \sum \lambda_i = 1 \}$, and

$$f(\lambda_1, \lambda_2, \lambda_3, x_1, x_2, x_3) = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3.$$  

Since $f$ is continuous and $K$ is assumed compact, the right hand side of (4.7) is clearly compact, and it suffices to prove (4.7). This takes a bit of linear algebra: if $x$ is any convex combination of $m$ points in $K$, then it is actually the convex combination of three points in $K$. In $\mathbb{R}^n$ this the same is generally true with "three" replaced by "$(n+1)$" (so the lemma remains valid in $\mathbb{R}^n$). For the remaining details, see [16]. □

Lemma 4.8. $\text{Ex}(\text{conv}(K)) \subset K$ for all bounded sets $K \subset \mathbb{R}^2$.

Proof. If $x \in \text{Ex}(\text{conv}(K))$, then $x \in \text{conv}(K)$ by definition of "Ex", and hence $x$ can be represented as a convex combination of points in $K \subset \text{conv}(K)$. If $x \notin K$, then the combination is necessarily non-trivial, and hence $x \notin \text{Ex}(\text{conv}(K))$ contrary to assumption. □

Now, we start the proof of Theorem 4.1. The argument is copied nearly verbatim from the book of Bishop and Peres, see [2, Theorem 10.5.1]. For a closed convex set $R \subset \mathbb{R}^2$, define the following variant of the $\beta_{\infty}$-number:

$$\beta(R) := \max_L \sup_{x \in R} \frac{\text{dist}(x, L)}{\text{diam}(R)},$$

where the "max$_L$" is taken over all chords of $R$ of length $\mathcal{H}^1(L) = \text{diam}(R)$. A chord is a line segment with both endpoints on $\partial R$. A simple compactness argument shows that the first max is well-defined, and attained for some chord $L_R$, and $L_R$ will be called the diameter of $R$. Note that $\text{dist}(z, L_R) \leq \beta(R) \text{ diam}(R)$ for all $x \in R$, whence $R$ can be covered by a rectangle of dimensions $\text{diam}(R) \times 2\beta(R)$ parallel to $L_R$.

4.1. Construction and connectedness. The curve $\Gamma$, covering $E$, will be obtained as the intersection of a nested sequence of compact connected sets $\Gamma_n$, each containing $E$. Every set $\Gamma_n$ has the form

$$\Gamma_n = \bigcup_{R \in \mathcal{R}_n} R \cup \bigcup_{k=0}^n \bigcup_{B \in \mathcal{B}_k} B,$$

where the family $\mathcal{R}_n$ consists of interior-disjoint closed convex sets $\mathcal{R}_n$, and the families $\mathcal{B}_k$ consist of closed line segments, called bridges. Note that bridges are never deleted: $\Gamma_{n+1}$ contains all the bridges contained in $\Gamma_n$.

The following properties will be maintained throughout the construction, and they will guarantee that each set $\Gamma_n$ is connected:

Property 1 (Connectivity). For $n \in \mathbb{N}$, let

$$\mathcal{F}_n := \mathcal{R}_n \cup \bigcup_{k=0}^n \mathcal{B}_k$$
be the family of sets such that $\Gamma_n = \bigcup \{ F : F \in \mathcal{F}_n \}$. First, the extreme points $\text{Ex}(R)$ of every set $R \in \mathcal{F}_n$ lie in $E$. Second, if $K_1, K_2 \in \mathcal{F}_n$, then then $K_1$ and $K_2$ can be joined by an extreme point tour: there exist sets

$$K_1 = E_1, E_2, \ldots, E_{m-1}, E_m = K_2 \in \mathcal{F}_n$$

such that $\text{Ex}(E_j) \cap \text{Ex}(E_{j+1}) \neq \emptyset$ for $1 \leq j \leq m - 1$. In particular, $\Gamma_n$ is connected.

Now, the construction begins. Set

$$R_0 := \{ \text{conv}(E) \} \quad \text{and} \quad B_0 = \emptyset.$$ 

Then Property 1 is satisfied by Lemma 4.8. Assume that $R_n, B_n$ have already been defined for some $n \geq 0$. Then, $R_{n+1}$ will be defined by replacing every set $R \in R_n$ by two further interior-disjoint closed convex sets, called the children of $R$. The children of $R$ may be connected by a line segment, which is added to $B_{n+1}$. In particular, no sets are ever deleted "later" from $B_n$.

Fix $R \in R_n$. The construction now divides to two cases.

**Case (NB).** Here "NB" stands for "no bridge". Let $S_{\text{mid}}$ be the closed middle third of the diameter chord $L_R$, see Figure 2. Let $\pi$ be the orthogonal projection to $L_R$. In the case (NB), assume that $E \cap R \cap \pi^{-1}(S_{\text{mid}}) \neq \emptyset$, and pick any point $x \in E \cap R \cap \pi^{-1}(S_{\text{mid}})$. Then, divide $L_R$ into two closed sub-segments $L_1$ and $L_2$, with a common endpoint at $\pi(x)$. Define $R_1$ and $R_2$ to be the convex hulls of the sets

$$E \cap R \cap \pi^{-1}(L_1) \quad \text{and} \quad E \cap R \cap \pi^{-1}(L_2),$$

respectively. It is easy to see that that $\pi^{-1}(\pi(x))$ contains a point in $\text{Ex}(R_1) \cap \text{Ex}(R_2)$, so in particular

$$\text{Ex}(R_1) \cap \text{Ex}(R_2) \neq \emptyset.$$ 

In fact, the set $R_1 \cap R_2$ is a (possibly degenerate) line segment, whose endpoints lie in $\text{Ex}(R_1) \cap \text{Ex}(R_2)$.

To check that Property 1 remains valid, first note that $\text{Ex}(R_1), \text{Ex}(R_2) \subset E$: indeed, $R_1, R_2$ are the convex hulls of certain compact subsets of $E$, and one can just apply Lemma 4.8. So, it remains to check that the "extreme points tour" property remains valid. To this end, first note that

$$\text{Ex}(R) \subset \text{Ex}(R_1) \cup \text{Ex}(R_2).$$  

(4.9)
This is because if \( x \in \text{Ex}(R) \), then \( x \in E \) by Property 1, and so evidently \( x \in R_1 \cup R_2 \). Assume, for instance, that \( x \in R_1 \). Now, if \( x \) could be expressed as a non-trivial convex combination of elements in \( R_1 \), then it could certainly be expressed as such a combination of elements in \( R \), which would violate \( x \in \text{Ex}(R) \). It follows that \( x \in \text{Ex}(R_1) \), which proves (4.9).

Finally, let \( K_1, K_2 \in F_{n+1} \) be sets as in Property 1. The task is to find an extreme point tour in \( F_{n+1} \), connecting \( K_1 \) to \( K_2 \). For \( j \in \{1, 2\} \), write

\[
\hat{K}_j := \begin{cases} K_j, & \text{if } K_j \notin \{R_1, R_2\}, \\ R, & \text{if } K_j \in R_1, R_2. \end{cases}
\]

Then \( \hat{K}_1, \hat{K}_2 \in F_n \), and there exists an extreme point tour

\[
\hat{K}_1 = E_1, E_2, \ldots, E_{m-1}, E_m = \hat{K}_2 \in F_n
\]

If \( E_j \neq R \) for all \( 1 \leq j \leq m \), then \( E_1, \ldots, E_m \) is also an extreme point tour in \( F_{n+1} \), connecting \( \hat{K}_1 \) to \( \hat{K}_2 \). Otherwise, if \( E_j = R \) for some \( 1 \leq j \leq m \), the tour does not lie in \( F_{n+1} \), and one needs to modify it. Assume first that \( 1 < j < m \), and \( E_j = R \). By definition of the tour, \( \text{Ex}(E_{j-1}) \cap \text{Ex}(E_j) \neq \emptyset \) and \( \text{Ex}(E_j) \cap \text{Ex}(E_{j+1}) \neq \emptyset \). Since \( \text{Ex}(E_j) = \text{Ex}(R) \subset \text{Ex}(R_1) \cup \text{Ex}(R_2) \) by (4.9), one has either

\[
\text{Ex}(E_{j-1}) \cap \text{Ex}(R_1) \neq \emptyset \quad \text{or} \quad \text{Ex}(E_{j-1}) \cap \text{Ex}(R_2) \neq \emptyset.
\]

For instance, assume that \( \text{Ex}(E_{j-1}) \cap \text{Ex}(R_1) \neq \emptyset \). Similarly, either

\[
\text{Ex}(E_{j+1}) \cap \text{Ex}(R_1) \neq \emptyset \quad \text{or} \quad \text{Ex}(E_{j+1}) \cap \text{Ex}(R_2) \neq \emptyset.
\]

Assume for instance that \( \text{Ex}(E_{j+1}) \cap \text{Ex}(R_2) \neq \emptyset \). Because also \( \text{Ex}(R_1) \cap \text{Ex}(R_2) \neq \emptyset \), the set \( E_j = R \) can be replaced by \( E_1^j = R_1 \) and \( E_2^j = R_2 \), and \( E_1, \ldots, E_1^j, E_2^j, \ldots, E_m \) remains an extreme point tour connecting \( \hat{K}_1 \) to \( \hat{K}_2 \). Once all occurrences of \( E_j = R \), \( 1 < j < m \), have been replaced in this manner, then we have a tour connecting \( \hat{K}_1 \) to \( \hat{K}_2 \) in \( F_{n+1} \), apart possibly from the endpoints. If one of the endpoints does not lie in \( F_n \), say \( \hat{K}_1 \notin F_n \), this means precisely that \( \hat{K}_1 = R \), and \( K_1 \in \{R_1, R_2\} \). Then, repeating the argument from above, \( E_1 = K_1 = R \) can be replaced by either, or both of, the sets \( R_1, R_2 \), while maintaining the extreme point tour property. This gives a tour connecting \( K_1 \) to \( K_2 \) in \( F_{n+1} \), as desired.

\[\text{Figure 3. The case, where a new bridge } B_R \text{ is added to } B_{n+1}.\]
**Case (B).** Here "B" stands for "bridge". Recall the notation from the previous case, and this time assume that

\[ E \cap R \cap \pi^{-1}(S_{\text{mid}}) = \emptyset. \]

Thus, if \( L_R = S_{\text{left}} \cup S_{\text{mid}} \cup S_{\text{right}} \), then

\[ E \cap R \subset \pi^{-1}(S_{\text{left}}) \cup \pi^{-1}(S_{\text{right}}). \]

It follows that there are two minimal intervals \( L_1 \subset S_{\text{left}} \) and \( L_2 \subset S_{\text{right}} \) such that

\[ E \cap R \subset \pi^{-1}(L_1) \cup \pi^{-1}(L_2), \]

see Figure 3. As in the previous case, define \( R_1 := \text{conv}[E \cap R \cap \pi^{-1}(L_1)] \) and \( R_2 := \text{conv}[E \cap R \cap \pi^{-1}(L_2)] \). Property 1 remains valid by similar considerations as in the previous case.

Now, the construction of the sets \( \Gamma_n \) is complete, and \( \Gamma \) is defined by

\[ \Gamma := \bigcap_{n \geq 0} \Gamma_n. \]

It is clear that \( \Gamma \supset E \). Also, \( \Gamma \) is connected: if \( \Gamma \subset U_1 \cup U_2 \) with \( U_1, U_2 \) disjoint open sets, then \( \Gamma_n \subset U_1 \cup U_2 \) for sufficiently large \( n \) (otherwise \( (\Gamma_n \setminus [U_1 \cup U_2])_{n \in \mathbb{N}} \) would be a sequence of nested non-empty compact sets with empty intersection). But this is impossible, since \( \Gamma_n \) is connected for all \( n \).

**4.2. Length estimates.** It remains to prove the length bound (4.2). To this end, the following inequality will be first verified:

\[ \sum_{R \in \mathcal{R}_n} \text{diam}(R) + \sum_{k=0}^{n} \sum_{B \in \mathcal{B}_k} H^1(B) \leq (1 + \delta) \text{diam}(E) + C_\delta \sum_{k=0}^{n-1} \sum_{R \in \mathcal{R}_n} \beta^2(R) \text{diam}(R). \quad (4.10) \]

where \( \delta > 0 \), and \( C_\delta \geq 1 \) only depends on \( \delta \). For line segments, such as \( B \in \mathcal{B}_n \) or \( L_R \), I will abbreviate \( H^1(\cdot) =: |\cdot| \) in the sequel.

**Lemma 4.11.** Assume that \( R \in \mathcal{R}_n \) is replaced by \( R_1, R_2 \in \mathcal{R}_{n+1} \) and \( B_R \in \mathcal{B}_{n+1} \). Then

\[ \text{diam}(R_1) + \text{diam}(R_2) + \frac{|B_R|}{1 + \delta} \leq \text{diam}(R) + C_\delta \beta^2(R) \text{diam}(R). \]

**Proof.** Consider Case (NB) first, and recall the two line segments \( L_1, L_2 \subset L_R \) with common endpoint \( \pi(x) \). Note that

\[ \min\{|L_1|, |L_2|\} \geq \frac{|L_R|}{3} = \frac{\text{diam}(R)}{3}. \quad (4.12) \]

Since \( R_j \) is contained in a rectangle with dimensions \( |L_j| \times 2\beta(R) \text{diam}(R) \) parallel to \( L_R \), one has, using (4.12), that\(^5\)

\[ \text{diam}(R_j) \leq \sqrt{|L_j|^2 + 4\beta^2(R) \text{diam}(R)^2} \leq |L_j| + 6\beta^2(R) \text{diam}(R). \quad (4.13) \]

This should be contrasted with the "trivial estimate"

\[ \text{diam}(R_j) \leq \sqrt{|L_j|^2 + 4\beta^2(R) \text{diam}(R)^2} \leq |L_j| + 2\beta(R) \text{diam}(R), \quad (4.14) \]

---

\(^5\)You may first wish to check, abstractly, that \( a \geq c/3 \) implies \( \sqrt{a^2 + 4bc^2} \leq a + 6bc \) for \( a, b, c \geq 0 \).
which holds without any assumptions on $|L_j|$. Since $|L_1| + |L_2| \leq \text{diam}(R)$, the claim of the lemma then follows from (4.13) with $C_\delta = 12$.

Next, consider Case (B). Let $J_R \subset L_R$ be the segment $J_R = L_R \setminus (L_1 \cup L_2)$, where $L_1$ and $L_2$ are defined as in Case (B). Then

$$|B_R| \leq \sqrt{|J_R|^2 + 4\beta^2(R) \text{diam}(R)^2} \leq |J_R| + 2\beta(R) \text{diam}(R). \quad (4.15)$$

Now, there are two subcases. First, if $\beta(R) \geq \theta := \delta/100$, then

$$\text{diam}(R_1) + \text{diam}(R_2) + \frac{|B_R|}{1+\delta} \leq 3 \text{diam}(R) \leq \text{diam}(R) + \left(\frac{2}{\theta^2}\right) \beta^2(R) \text{diam}(R).$$

If $\beta(R) < \theta$, then, combining (4.15) with the "trivial estimate" (4.13) gives

$$\text{diam}(R_1) + \text{diam}(R_2) + \frac{|B_R|}{1+\delta} \leq |L_1| + |L_2| + 4\theta \text{diam}(R) + \frac{|J_R| + 2\theta \text{diam}(R)}{1+\delta} = \frac{1}{1+\delta} \left[1 + 2\delta/3 + 4\theta(1+\delta) + 2\theta \text{diam}(R)\right] \quad (4.16)$$

noting that $|L_1| + |L_2| + |J_R| = \text{diam}(R)$, and $|L_1| + |L_2| \leq (2/3) \text{diam}(R)$. Since $\theta = \delta/100$, the factor of $\text{diam}(R)$ on line (4.16) is $\leq 1$, and the lemma follows. □

Now the time is ripe to prove (4.10). For $R \in \mathcal{R}_n$, write $\text{ch}(R) := \{R_1, R_2\} \subset \mathcal{R}_{n+1}$ for the children of $R$, and write $B_R \in B_{n+1}$ for the bridge associated to $R$; in Case (NB), let $B_R := \emptyset$. Then, for any $n \geq 1$, using Lemma 4.11,

$$\sum_{R \in \mathcal{R}_n} \text{diam}(R) + \frac{1}{1+\delta} \sum_{k=0}^n \sum_{B \in B_k} |B| \leq \sum_{R \in \mathcal{R}_{n-1}} \left[\sum_{R' \in \text{ch}(R)} \text{diam}(R') + \frac{|B_R|}{1+\delta}\right] + \frac{1}{1+\delta} \sum_{k=0}^{n-1} \sum_{B \in B_k} |B| \leq \sum_{R \in \mathcal{R}_{n-1}} \text{diam}(R) + \frac{1}{1+\delta} \sum_{k=0}^{n-1} \sum_{B \in B_k} |B| + C_\delta \sum_{R \in \mathcal{R}_{n-1}} \beta^2(R) \text{diam}(R). \quad L.4.11$$

The term in brackets on the last line is of the same form as the term on the first line, so the estimate can be iterated, and $n$ repetitions give

$$\sum_{R \in \mathcal{R}_n} \text{diam}(R) + \frac{1}{1+\delta} \sum_{k=0}^n \sum_{B \in B_k} |B| \leq \text{diam}(R_0) + C_\delta \sum_{k=0}^{n-1} \beta^2(R) \text{diam}(R).$$

Multiplying both sides by $(1+\delta)$ gives (4.10), recalling that $\text{diam}(R_0) = \text{diam}(E)$.

The next task will be to prove

$$\sum_{k=0}^{\infty} \sum_{R \in \mathcal{R}_k} \beta^2(R) \text{diam}(R) \lesssim \beta^2_{\omega\varepsilon}(E). \quad (4.17)$$

Let’s start with an easy observation: for any $R \in \mathcal{R}_n$,

$$\max\{\mathcal{L}^2(R_1), \mathcal{L}^2(R_2)\} \leq \tau \mathcal{L}^2(R) \quad (4.18)$$
for some absolute constant $\tau < 1$. To see this, let $\Delta \subset R$ be the triangle spanned by the endpoints of $L_R$, and some point $x \in \partial R$ with $\text{dist}(x, L_R) = \beta(R) \, \text{diam}(R)$. Then

$$L^2(R) \sim L^2(\Delta) = \frac{\beta(R) \, \text{diam}(R)^2}{2} \quad (4.19)$$

and

$$L^2(\Delta \cap \pi^{-1}(S_{\text{left}})) \sim L^2(\Delta) \sim L^2(\Delta \cap \pi^{-1}(S_{\text{right}}))$$

with absolute constants. It follows that

$$L^2(R_1) \leq L^2(R) - L^2(\Delta \cap \pi^{-1}(S_{\text{right}})) \leq \tau L^2(R),$$

and a similar estimate holds for $L^2(R_2)$. This proves (4.18). Next, a similar estimate is desired for the diameters of the convex sets:

**Lemma 4.20.** There is an absolute constant $N \in \mathbb{N}$ with the following property. If $R \in \mathcal{R}_n$ and $R' \in \mathcal{R}_{n+N}$ with $R' \subset R$, then

$$\text{diam}(R') \leq \frac{3 \, \text{diam}(R)}{4}.$$  

**Proof.** If the number $\beta(R)$ is small enough, say $\beta(R) \leq \theta$, then

$$\max\{\text{diam}(R_1), \text{diam}(R_2)\} \leq 3 \, \text{diam}(R)/4.$$

(Just have a look at the estimates for $\text{diam}(R_1)$ within Lemma 4.11, if you are unsure.) So, if

$$R = R^{(0)} \supset R^{(1)} \supset \ldots \supset R^{(N)} = R'$$

is a sequence with $R^{(j)} \in \mathcal{R}_{n+j}$ and if $\beta(R^{(N-1)}) \leq \theta$, then

$$\text{diam}(R') \leq 3 \, \text{diam}(R^{(N-1)})/4 \leq 3 \, \text{diam}(R)/4,$$

and the proof is complete. Otherwise $\beta(R^{(N-1)}) \geq \theta$, and by (4.18)-(4.19),

$$\theta \, \text{diam}(R')^2 \leq \theta \, \text{diam}(R^{(N-1)})^2 \leq L^2(R^{(N-1)}) \leq \tau^{N-1} \, L^2(R) \leq \tau^{N-1} \, \text{diam}(R)^2.$$

This proves that, in every case, $\text{diam}(R') \leq 3 \, \text{diam}(R)/2$, if $N$ is large enough. \hfill \Box

Now, the table is set for (4.17). For a dyadic square $Q$, let

$$\mathcal{R}(Q) := \bigcup_{n \geq 0} \{ R \in \mathcal{R}_n : R \cap Q \neq \emptyset \text{ and } \ell(Q)/2 < \text{diam}(R) \leq \ell(Q) \}.$$

I claim that the collection $\mathcal{R}(Q)$ has bounded "interior overlap", that is,

$$\sum_{R \in \mathcal{R}(Q)} \chi_{\text{int}R} \lesssim 1. \quad (4.21)$$

To see this, note that two sets $R_1 \in \mathcal{R}_m$ and $R_2 \in \mathcal{R}_n$, $m, n \geq 0$, can have shared interior, only if either $R_1$ is a descendant of $R_2$, or vice versa. So, if $x \in \text{int} R_i$ for $R_i \in \mathcal{R}(Q)$, $1 \leq i \leq M$, then the sets $R_1, \ldots, R_M$ are nested. However, by Lemma 4.20, the diameters of the convex sets decay rapidly in long nested sequences; since $\ell(Q)/2 < \text{diam}(R_i) \leq \ell(Q)$ for all $1 \leq i \leq M$, this sets an upper bound for $M$ and proves (4.21).

Another observation about $\mathcal{R}(Q)$ is the following: if $R \in \mathcal{R}(Q)$, then $R \subset 2Q$, because $R \cap Q \neq \emptyset$ and $\text{diam}(R) \leq \ell(Q)$. In particular, $E \cap R \subset E \cap 2Q$ is contained in $W \cap 2Q$, where $W$ is a strip of width $2 \beta_{E,\infty}(2Q)/\ell(2Q)$. Since $W \cap 2Q$ is convex, it follows that

$$R = \text{conv}(E \cap R) \subset W \cap 2Q, \quad R \in \mathcal{R}(Q). \quad (4.22)$$
Next, use (4.19), (4.21) and (4.22) to make the following estimate:

$$
\sum_{R \in \mathcal{R}(Q)} \beta(R) \, \text{diam}(R)^2 \sim \sum_{R \in \mathcal{R}(Q)} \mathcal{L}^2(\text{int } R)
$$

$$
= \int_{W \cap 2Q} \sum_{R \in \mathcal{R}(Q)} \chi_{\text{int } R} \, d\mathcal{L}^2
\lesssim \mathcal{L}^2(W \cap 2Q) \lesssim \beta_{E,\infty}(2Q) \ell(Q)^2.
$$

Since \( \text{diam}(R) \sim \ell(Q) \) for all \( R \in \mathcal{R}(Q) \), one concludes that

$$
\sum_{R \in \mathcal{R}(Q)} \beta(R) \, \text{diam}(R)^2 \lesssim \beta_{E,\infty}^2(2Q).
$$

Finally, observe that every convex set \( R \in \mathcal{R}_n \), for any \( n \geq 0 \) is contained in at least one of the sets \( \mathcal{R}(Q), Q \in \mathcal{D}_E \), with \( \ell(Q) \leq 2 \text{diam}(E) \) (because \( E \cap R \neq \emptyset \), and any point \( x \in E \cap R \) is contained in a dyadic square \( Q \in \mathcal{D}_E \) with \( \ell(Q)/2 < \text{diam}(R) \leq \ell(Q) \)). Consequently,

$$
\sum_{k=0}^{n} \sum_{R \in \mathcal{R}_k} \beta(R) \, \text{diam}(R) \lesssim \sum_{Q \in \mathcal{D}_E} \ell(Q) \sum_{R \in \mathcal{R}(Q)} \beta^2(R) \lesssim \sum_{Q \in \mathcal{D}_E} \sum_{Q \subset 3E} \beta_{E,\infty}^2(2Q) \ell(Q) = \beta_{E,\infty}^2(E),
$$

which proves (4.17).

Combining (4.10) and (4.17), it has now been established that

$$
\sum_{R \in \mathcal{R}_n} \text{diam}(R) + \sum_{k=0}^{n} \sum_{B \in \mathcal{B}_k} \mathcal{H}^1(B) \leq (1 + \delta) \text{diam}(E) + \beta_{E,\infty}^2(E)
$$

uniformly for \( n \geq 0 \). Since

$$
\Gamma \subset \Gamma_n = \bigcup_{R \in \mathcal{R}_n} R \cup \bigcup_{k=0}^{n} \bigcup_{B \in \mathcal{B}_k} B, \quad n \geq 0,
$$

the desired bound for \( \mathcal{H}^1(\Gamma) \) now follows from the very definition of \( \mathcal{H}^1 \) (and the fact that the diameters of the sets in \( \mathcal{R}_n \) tend to zero uniformly, as \( n \to \infty \)).

5. The \( L^1 \)-Traveling Salesman Theorem

This section contains two proofs of the \( L^1 \)-traveling salesman theorem for doubling measures, Theorem 2.11, which is reproduced as Theorem 5.2 below. The result first appeared in a paper [1] of M. Badger and R. Schul from 2016; there it was obtained via a new "geometric traveling salesman theorem", which is the main topic in Section 7. However, an observation of X. Tolsa from [17, Section 7] allows one to reduce Theorem 5.2 to Jones’ \( L^\infty \) traveling salesman theorem, if one assume that the doubling measure \( \mu \) is smooth in the following sense:

**Definition 5.1** (Smooth measures). A Radon measure \( \mu \) on \( \mathbb{R}^n \) is smooth, if it is doubling, and there is a constant \( \theta > 0 \) such that

$$
\mu(B(x, tr)) \leq \frac{\mu(B(x, r))}{2}, \quad x \in \text{spt } \mu, \; 0 < r \leq \text{diam}(\text{spt } \mu).
$$
The reduction to Jones’ theorem gives a (fairly) short proof of Theorem 5.2 in the plane, assuming smoothness. These considerations constitute the first half of the section. The second half contains the proof of Theorem 5.2 in full generality, following the argument of Badger and Schul.

I now recall the result:

**Theorem 5.2 (Badger-Schul).** Let $\mu$ be a doubling measure on $\mathbb{R}^n$ with compact support $E = \text{spt} \, \mu$, and assume that the numbers $\beta_{\mu, 1}$ satisfy

$$\beta_1^2(\mu) := \sum_{Q \in \mathcal{D}_E} \beta_{\mu, 1}^2(2Q) \ell(Q) < \infty,$$

where $\lambda \geq 1$ is a sufficiently large constant depending only on $n$, and, as always, $\mathcal{D}_E = \{Q \in \mathcal{D} : Q \cap E \neq \emptyset\}$. Then $\text{spt} \, \mu$ can be covered by a continuum $\Gamma \subset \mathbb{R}^n$ with

$$\mathcal{H}^1(\Gamma) \lesssim D_{\mu, n} \text{diam}(E) + \beta_1^2(\mu).$$

If the $\beta_{\mu, 1}$-numbers satisfy a Carleson condition, then $\Gamma$ can be taken to be 1-Ahlfors-David regular, at least in the plane:

**Theorem 5.3 (Badger-Schul, Carleson version).** Same assumptions as in Theorem 5.2, except that $n = 2$, and the finiteness of $\beta_1^2(\mu)$ is replaced by the Carleson condition

$$\sum_{Q \in \mathcal{D}_E} \beta_{\mu, 1}^2(2Q) \ell(Q) \lesssim \ell(R), \quad R \in \mathcal{D}_E.$$

Then $E = \text{spt} \, \mu$ can be covered by an AD regular continuum.

**Remark 5.4.** Note that, in both theorems above, using the $\beta_{\mu, 1}$-numbers gives the strongest possible result (as contrasted to the numbers $\beta_{\mu, p}$ for any $1 \leq p \leq \infty$).

5.1. **Tolsa’s observation.** Here is the main result of this subsection: it allows us to transfer assumptions on the $\beta_{\mu, 1}$-numbers to those on $\beta_{E, \infty}$-numbers.

**Theorem 5.5 (Tolsa).** Let $\mu$ smooth measure with $E = \text{spt} \, \mu \subset \mathbb{R}^2$. Then

$$\sum_{Q \in \mathcal{D}_E} \beta_{E, \infty}^2(2Q) \ell(Q) \lesssim \sum_{Q \in \mathcal{D}_E} \beta_{\mu, 1}^2(3Q) \ell(Q)$$

for any square $R \in \mathcal{D}$ with $\ell(R) \leq 10 \text{diam}(E)$.

Key to the proof is the following geometric lemma:

**Lemma 5.6.** Let $\mu$ be a smooth measure with $E := \text{spt} \, \mu$, and let $Q \in \mathcal{D}_E$ with $\ell(Q) \leq 10 \text{diam}(E)$. Let $\ell_Q$ be a line with $\beta_{\mu, 1}(3Q, \ell_Q) = \beta_{\mu, 1}(3Q)$. Then, for any $x \in E \cap 2Q$,

$$\text{dist}(x, \ell_Q) \lesssim \sum_{P \in \mathcal{D}_E} \beta_{\mu, 1}(3P) \ell(P),$$

where the implicit constants only depend on the doubling constant of $\mu$.

The proof of the lemma needs two further lemmas. The first is an observation of G. Lerman, while the second is folklore.
Lemma 5.7. Let \( n \geq 2 \) and \( 1 \leq p < \infty \). Let \( \mu \) be a Radon measure, and let \( B \subset \mathbb{R}^n \) be a set with \( 0 < \text{diam}(B) < \infty \) and \( 0 < \mu(B) < \infty \). Let \( c_B \) be the centre of mass of \( \mu \) in \( B \), namely

\[
c_B := \frac{1}{\mu(B)} \int_B x \, d\mu.
\]

Then

\[
dist(c_B, \ell) \leq \beta_{\mu,p}(B, \ell) \text{diam}(B)
\]

for every straight line \( \ell \subset \mathbb{R}^n \).

Proof. Fix a straight line \( \ell \subset \mathbb{R}^n \) and note that the function \( x \mapsto dist(x, \ell)^p \) is convex for \( p \in [1, \infty) \). So, by a vector-valued version of Jensen’s inequality, see lemma below,

\[
dist(c_B, \ell)^p := \left( \frac{1}{\mu(B)} \int_B x \, d\mu, \ell \right)^p \leq \frac{1}{\mu(B)} \int_B dist(x, \ell)^p \, d\mu = \text{diam}(B)^p \beta_{\mu,p}(B, \ell)^p.
\]

\( \square \)

Lemma 5.8 (Jensen’s inequality). Let \((X_1, \ldots, X_n)\) be a random vector on some probability space \((\Omega, \mathbb{P})\), and let \( \varphi : \mathbb{R}^n \to \mathbb{R} \) be a convex function. Then,

\[
\varphi(\mathbb{E}[X_1, \ldots, X_n]) \leq \mathbb{E}[\varphi(X_1, \ldots, X_n)].
\]

Proof. We cheat by assuming that \( \varphi \) is differentiable everywhere, because this suffices for the application in Lemma 5.7 (at least for \( p > 1 \), and the case \( p = 1 \) can be obtained by taking limits). A consequence of convexity is the following inequality:

\[
\varphi(x) - \varphi(y) \geq \nabla \varphi(y) \cdot (x - y), \quad x, y \in \mathbb{R}^n.
\]

Indeed, writing \( v := (x - y)/|x - y| \), we have

\[
\nabla \varphi(y) \cdot (x - y) = |\nabla \varphi(y) \cdot v| \cdot |x - y| = \left| \lim_{h \to 0^+} \frac{\varphi(y + hv) - \varphi(y)}{h} \right| \cdot |x - y|
\]

\[
= \left| \lim_{h \to 0^+} \varphi((h/|x - y|)x + (1 - h/|x - y|)y) - \varphi(y) \right| \cdot |x - y|
\]

\[
\leq \left| \limsup_{h \to 0^+} \frac{(h/|x - y|) \varphi(x) + (1 - h/|x - y|) \varphi(y) - \varphi(y)}{h} \right| \cdot |x - y|
\]

\[
= \varphi(x) - \varphi(y),
\]

using convexity in the inequality. In particular, setting \( x = \bar{X} := (X_1, \ldots, X_n) \) and \( y = \mathbb{E}[\bar{X}] \) in (5.9), and taking expectations on both sides,

\[
\mathbb{E}[\varphi(\bar{X})] - \varphi(\mathbb{E}[\bar{X}]) \geq \mathbb{E}[\nabla \varphi(\mathbb{E}[\bar{X}])] \cdot (\bar{X} - \mathbb{E}[\bar{X}]) = 0.
\]

This proves the lemma. \( \square \)

Lemma 5.10. Let \( \mu \) be a smooth measure, and let \( P, R \) be (non-dyadic) cubes with \( \ell(P) \sim \ell(\mu) \leq \text{diam}(\text{spt} \mu) \), and let \( \ell_P, \ell_R \) be lines, which minimise \( \beta_{\mu,1}(P) \) and \( \beta_{\mu,1}(R) \), respectively. Assume that

\[
\tau P \cap \tau R
\]

contains a point of \text{spt} \( \mu \) for some \( \tau < 1 \). Then, \( \ell_P \) and \( \ell_R \) are very close in the following sense:

\[
dist(z, \ell_R) \lesssim_{\tau} \min\{\beta_{\mu,1}(P)\ell(P), \beta_{\mu,1}(R)\ell(R)\}, \quad z \in \ell_P \cap P.
\]

By symmetry, the same holds for \( \dist(z, \ell_P) \), for \( z \in \ell_R \cap R \).
Proof. Start by finding two points $x_0, y_0 \in P \cap R$ with $|x_0 - y_0| \sim \ell(P) \sim \ell(R)$ with the property that
\[
\max \{ \text{dist}(x_0, \ell_P), \text{dist}(y_0, \ell_P) \} \lesssim \beta_{\mu,1}(P) \ell(P)
\]
and
\[
\max \{ \text{dist}(x_0, \ell_R), \text{dist}(y_0, \ell_R) \} \lesssim \beta_{\mu,1}(R) \ell(R).
\]
Then, show that both $\ell_P$ and $\ell_R$ are close to the line spanned by the segment $[x_0, y_0]$, hence close to each other. The details are an exercise. □

The lemma above is the only place, where the smoothness of the measure $\mu$ is required.

Now, for the proof of Lemma 5.6:

![Diagram of cubes in the proof of Lemma 5.6](image)

**Figure 4.** Possible cubes in the proof of Lemma 5.6

**Proof of Lemma 5.6.** Let $Q = Q_0, Q_1, Q_2, \ldots$ be a sequence of dyadic cubes such that $x \in Q_m \subset 2Q$ and $\ell(Q_m) = 2^{-m}\ell(Q)$ for all $m \geq 1$ (Note that the requirement $x \in Q_m$ may not be possible for $m = 0$, in case $x \in 2Q \setminus Q = 2Q_0 \setminus Q_0$.) One can now easily check, see Figure 4, that the cubes $P_m := 3Q_m$, $m \geq 0$, are nested, and
\[
E \cap \frac{9}{10} P_m \cap \frac{9}{10} P_{m+1} \neq \emptyset, \quad m \geq 0.
\] (5.11)

For each $m \geq 0$, let $\ell_m$ be a line, which minimises $\beta_{\mu,1}(P_m, \ell)$ (so that $\ell_0 = \ell_Q$). Fix a constant $\theta > 0$ (whose value will only depend on the doubling of $\mu$). Let $N \geq 0$ be the smallest number such that $\beta_{\mu,1}(P_N) \geq \theta$. If no such number exists, just let $N$ be any (large) number. Let $a_N = c_{P_N} \in P_N$ be the $\mu$-centre of mass in $P_N$ (which evidently has positive measure by (5.11)). For $0 \leq m \leq N - 1$, define $a_m$ recursively to be the orthogonal projection of $a_{m+1}$ to the line $\ell_m$, so that $|a_{m+1} - a_m| = \text{dist}(a_{m+1}, \ell_m)$. Then
\[
\text{dist}(x, \ell_Q) \leq |x - a_N| + \text{dist}(a_N, \ell_Q)
\] (5.12)
If $N = 0$, then $\beta_{\mu,1}(P_0) \geq \theta$, and the claim of the lemma is clear. So, in the sequel, assume that $N \geq 1$. Then, the term on the right hand side can be further estimated as follows:

$$\text{dist}(a_N, \ell_Q) \leq |a_N - a_{N-1}| + \text{dist}(a_{N-1}, \ell_Q)$$

$$= \text{dist}(a_N, \ell_{N-1}) + \text{dist}(a_{N-1}, \ell_Q) \leq \ldots \leq \sum_{m=1}^{N} \text{dist}(a_m, \ell_{m-1}),$$

recalling that $\ell_0 = \ell_Q$. The next task is to prove

$$\text{dist}(a_m, \ell_{m-1}) \lesssim \beta_{\mu,1}(P_{m-1})\ell(P_{m-1}), \quad 1 \leq m \leq N. \quad (5.14)$$

For $m = N$, this is simple. Since $a_N = cP_N$, Lemma 5.7 with $p = 1$ and $\ell = \ell_{N-1}$ says that

$$|a_{N-1} - a_N| = \text{dist}(a_N, \ell_{N-1})$$

$$\lesssim \beta_{\mu,1}(P_N, \ell_{N-1})\ell(P_N) \quad (5.15)$$

$$\lesssim \beta_{\mu,1}(P_{N-1}, \ell_{N-1})\ell(P_{N-1})$$

$$= \beta_{\mu,1}(P_{N-1})\ell(P_{N-1}) \leq \theta\ell(P_{N-1}).$$

using also the doubling of $\mu$ (and (5.11)) in the second inequality, the definition of $\ell_{N-1}$ in the third, and the minimality assumption on $N$ in the last. If $\theta > 0$ was chosen small enough, this implies that $a_{N-1}$ is very close to $a_N$, and in particular

$$a_{N-1} \in \ell_{N-1} \cap P_{N-1}.$$  

Next, en route to (5.14), the plan is to verify by backward induction that

$$a_m \in \ell_m \cap P_m, \quad 0 \leq m \leq N - 1, \quad (5.16)$$

which was just seen to be true for $m = N - 1$. Suppose the claim is true for some $1 \leq m \leq N - 1$. Then, by definition, $a_{m-1}$ is the projection of $a_m \in \ell_m \cap P_m$ to the line $\ell_{m-1}$, which minimises $\beta_{\mu,1}(P_{m-1})$. Applying Lemma 5.10 with $P = P_m$, $R = P_{m-1}$, and $z = a_m \in \ell_m \cap P_m$ (the lemma can be used because of (5.11)) gives

$$|a_m - a_{m-1}| = \text{dist}(a_m, \ell_{m-1}) \lesssim \beta_{\mu,1}(P_{m-1})\ell(P_{m-1}) \leq \theta\ell(P_{m-1}). \quad (5.17)$$

Again, if $\theta > 0$ is small enough, this implies that $a_{m-1}$ lies very close to $a_m$, and in particular inside $P_{m-1}$.

Now that (5.16) has been verified for all $0 \leq m \leq N - 1$, the middle inequality in (5.17) is at our disposal for all $0 \leq m \leq N - 1$ (this may feel a bit complicated, but we really need the information $a_m \in P_m$ to invoke Lemma 5.10, so we had to check that first). Combining this with (5.15) gives (5.14).

Combining (5.14) further with (5.13)-(5.15) yields

$$\text{dist}(x, \ell_Q) \leq |x - a_N| + C \sum_{m=1}^{N} \beta_{\mu,1}(P_{m-1})\ell(P_{m-1}),$$

This is almost what we wanted. If $\beta(P_N) \geq \theta$, the term $|x - a_N|$ can be finally estimated by

$$|x - a_N| \lesssim \ell(P_N) \lesssim \frac{\beta(P_N)}{\theta}\ell(P_N),$$
Otherwise, if \( \beta(P_m) < \theta \) for all \( m \in \mathbb{N} \), one can just let \( N \to \infty \), and the term \( |x - a_N| \) vanishes. \( \square \)

Armed with the lemma, the proof of Theorem 5.5 is quite short:

**Proof of Theorem 5.5.** Fix \( Q \in \mathcal{D}_E \) with \( Q \subset R \), so that in particular \( \text{diam}(Q) \leq 10 \text{diam}(E) \). Find a line \( \ell_Q \), which minimises \( \beta_{\mu,1}^2(3Q, \ell) \), and then a point \( x \in E \cap 2Q \), which maximises \( \text{dist}(x, \ell_Q) \). Then

\[
\beta_{E,\infty}(2Q)\ell(Q) \leq \text{dist}(x, \ell_Q) \lesssim \sum_{P \in \mathcal{D}, x \in P \subset 2Q} \beta_{\mu,1}(3P)\ell(P)
\]

by Lemma 5.6. Taking squares and using Cauchy-Schwarz leads to

\[
\beta_{E,\infty}^2(2Q)\ell(Q)^2 \leq \left( \sum_{P \in \mathcal{D}_E, x \in P \subset 2Q} \beta_{\mu,1}^2(3P)\ell(P)^{3/2} \right) \left( \sum_{P \in \mathcal{D}_E, x \in P \subset 2Q} \ell(P)^{1/2} \right)
\]

\[
\lesssim \sum_{P \in \mathcal{D}_E, x \in P \subset 2Q} \beta_{\mu,1}^2(3P)\ell(P)^{3/2}\ell(Q)^{1/2}.
\]

Then, dividing by \( \ell(Q) \) and simply dropping the condition \( x \in P \) gives

\[
\beta_{E,\infty}^2(2Q)\ell(Q) \lesssim \sum_{P \in \mathcal{D}_E, P \subset 2Q} \beta_{\mu,1}^2(3P)\ell(P)^{3/2}\ell(Q)^{1/2} \leq \sum_{P \in \mathcal{D}_E, P \subset 2Q} \beta_{\mu,1}^2(3P)\ell(P)^{3/2}\ell(Q)^{1/2}.
\]

Next, sum over \( Q \in \mathcal{D}_E \) with \( Q \subset R \):

\[
\sum_{Q \in \mathcal{D}_E, Q \subset R} \beta_{E,\infty}^2(2Q)\ell(Q) \lesssim \sum_{Q \in \mathcal{D}_E, Q \subset R} \sum_{P \in \mathcal{D}_E, P \subset 2Q} \beta_{\mu,1}^2(3P)\ell(P)^{3/2}\ell(Q)^{1/2}
\]

\[
\leq \sum_{P \in \mathcal{D}_E, P \subset 3R} \beta_{\mu,1}^2(3P)\ell(P)^{3/2} \sum_{Q \in \mathcal{D}_E, 2Q \supset P} \frac{1}{\ell(Q)^{1/2}}
\]

\[
\lesssim \sum_{P \in \mathcal{D}_E, P \subset 3R} \beta_{\mu,1}^2(3P)\ell(P).
\]

This proves Theorem 5.5. \( \square \)

5.2. **Proof of the \( L^1 \) traveling salesman theorem for smooth doubling measures.** The proofs of Theorems 5.2 and 5.3, for smooth measures, are now straightforward applications of Jones’ \( L^\infty \) traveling salesman theorem, and the preceding machinery.

**Proof of Theorem 5.2.** Write \( E := \text{spt} \mu \) and let \( R \in \mathcal{D} \) be the smallest dyadic cube such that \( 2R \) contains \( 3E \). Then \( 2R \) is the union of 9 dyadic squares \( R_1, \ldots, R_9 \), and for every \( j \), it holds that \( 2R_j \subset 10E \). Then

\[
\sum_{Q \in \mathcal{D}_E, Q \subset 3E} \beta_{E,\infty}^2(2Q)\ell(Q) = \sum_{j=1}^9 \sum_{Q \in \mathcal{D}_E, Q \subset R_j} \beta_{E,\infty}^2(2Q)\ell(Q) \lesssim \sum_{P \in \mathcal{D}_E, P \subset 10E} \beta_{\mu,1}^2(3P)\ell(P)
\]
The "3" is easy to replace by "2". By the doubling hypothesis on $\mu$, if $Q \in D_E$ is the smallest cube such that $3P \subset 2Q$, then $\beta^2_{\mu,1}(3P) \lesssim \beta^2_{\mu,1}(2Q)$, and of course $2Q$ is still contained in, say, $20E$. So, the right hand side is finite by assumption, and now the existence of $\Gamma$ follows from Jones’ $L^\infty$ traveling salesman Theorem 4.1. \hfill \square

To prove Theorem 5.3, we recall an earlier exercise:

**Exercise 5.18.** Let $E \subset \mathbb{R}^2$ a uniformly 1-rectifiable compact set: for every ball $B$, the intersection $B \cap E$ can be covered by a continuum $\Gamma_B$ of length $\mathcal{H}^1(\Gamma_B) \leq C \operatorname{diam}(B)$. Prove that there exists an AD regular continuum $\Gamma \supset E$ with $\operatorname{diam}(\Gamma) \sim \operatorname{diam}(E)$, where the implicit constant only depends on $C$.

**Proof of Theorem 5.3.** By the exercise, it suffices to prove that $E = \operatorname{spt} \mu$ is uniformly rectifiable: for every disc $B \subset \mathbb{R}^2$ there exists a continuum $\Gamma \supset B \cap E$ with $\mathcal{H}^1(\Gamma_B) \lesssim \operatorname{diam}(B)$. Here the implicit constants will only depend on the constant "C" in the assumed Carleson condition

$$\sum_{Q \in D_E, Q \subset R} \beta^2_{\mu,1}(2Q) \ell(Q) \leq C \ell(R), \quad R \in D_E. \quad (5.19)$$

The plan is to apply Jones’ traveling salesman theorem to the set $E_B := \Gamma \cap E$_covered by a continuum $\Gamma \supset E$ with $\mathcal{H}^1(\Gamma_B) \lesssim \operatorname{diam}(B)$.\hfill \square

5.3. **Proof of the $L^1$ traveling salesman theorem for general doubling measures.** In this section, I discuss the proof of Theorem 5.2, as given in the original paper [1], and without the "smoothness" assumption. The main difference to the previous proof is that there is no need for a reduction to Jones’ $L^\infty$ traveling salesman theorem: Badger and Schul use the $\beta_{\mu,1}$-numbers directly. The argument below (taking into account Section 7) may seem more complicated than the one above, but note that it works in all dimensions (and gives a better result). A caveat of the Badger-Schul approach seems to be that the 1-Ahlfors-David regularity of the curve is not so easy to prove (assuming the Carleson condition for the $\beta_{\mu,1}$-numbers).

The main component in the proof of Badger and Schul is the following "geometric traveling salesman theorem":

**Theorem 5.20.** Let $n \geq 2, A > 1, x_0 \in \mathbb{R}^n$ and $r_0 > 0$. Let $(V_k)_{k \in \mathbb{N}}$ be a sequence of non-empty finite subsets of $B(x_0, Ar_0)$ such that the following conditions are satisfied:

- $(V_{sep})$ The distance between distinct points in $V_k$ is at least $2^{-k} r_0$.
- $(V^1)$ For all $v \in V_k$, there exists $v^+ \in V_{k+1}$ with $|v - v^+| < A 2^{-(k+1)} r_0$.
- $(V^2)$ For all $v \in V_{k+1}$, there exists $v^+ \in V_k$ with $|v - v^+| < A 2^{-k} r_0$.  

Further, assume that for all $k \geq 1$ and for all $v \in V_k$ there is a line $\ell_v = \ell_{k,v} \subset \mathbb{R}^n$ and a number $\alpha_v = \alpha_{k,v} \geq 0$ such that
\[
\sup_{x \in (V_{k-1} \cup V_k) \cap B(v, 65A2^{-k}r_0)} \text{dist}(x, \ell_v) \leq \alpha_v 2^{-k}r_0.
\] (5.21)
Then the sets $V_k$ converge in the Hausdorff metric to a compact set $V \subset \overline{B(x_0, Ar_0)}$, and there exists a compact, connected set $\Gamma \subset \overline{B(x_0, Ar_0)}$ such that $\Gamma \supset V$, and
\[
\mathcal{H}^1(\Gamma) \lesssim_{A,n} r_0 + \sum_{k \in \mathbb{N}} \sum_{v \in V_k} \alpha_v^2 2^{-k}r_0.
\] (5.22)

The proof of Theorem 5.20 is given in Section 7.

**Proof of Theorem 5.2, assuming Theorem 5.20.** Without loss of generality, we may assume that $\text{spt} \mu$ is contained in a single dyadic cube $Q_0 \in \mathcal{D}$ with $\ell(Q_0) \lesssim \text{diam}(\text{spt} \mu)$. In any case, at most $2^n$ cubes with this property are needed, and we can construct $2^n$ separate curves inside each of those; in the end, to get a single curve, the resulting $2^n$ curves are simply joined with line segments of length $\lesssim \text{diam}(\text{spt} \mu)$.

Let $\mathcal{T} := \mathcal{D}_{\text{spt} \mu}$ be the collection of dyadic cubes intersecting the support of $\mu$, namely
\[
\mathcal{T} := \{Q \subset Q_0 : Q \cap \text{spt} \mu \neq \emptyset\}.
\]
Then certainly $\mu(2Q) > 0$ for all $Q \in \mathcal{T}$, and we may define $c_{2Q}$ as the centre of mass of $\mu$ in $2Q$:
\[
c_{2Q} := \frac{1}{\mu(2Q)} \int_{2Q} x \, d\mu x \in 2Q
\]
Write $r_0 := 2^{-k_0} := \ell(Q_0)$. For $k \geq 0$, let $V_k$ be a maximal $2^{-(k+k_0)}$-separated set in
\[
\{c_{2Q} : Q \in \mathcal{T} \cap \mathcal{D}_{k+k_0}\}.
\]
It is then clear that $V_k$ satisfies the separation condition $(V_{\text{sep}})$ of Theorem 5.20. The properties $(V^\uparrow)$ and $(V_\downarrow)$ are also fairly clear. To check $(V_\uparrow)$, for instance, fix $v = c_{2Q} \in V_k$. Then $Q$ clearly has a child $Q' \in \mathcal{T} \cap \mathcal{D}_{k+1}$, so either $v' = c_{2Q'} \in V_{k+1} \cap 2Q'$, or then there is some other point $v'' \in V_{k+1}$ at distance $\leq 2^{-(k+k_0)+1}$ from $v'$. In both cases, $v$ is at distance $\lesssim 2^{-(k+k_0)}$ some point in $V_{k+1}$, which can then be designated as $v^\downarrow$. The proof of condition $(V^\uparrow)$ is similar, and left for the reader.

Now we would like to define the lines $\ell_v$ and the numbers $\alpha_v$ for $v \in V_k$, $k \geq 1$. So, fix $v = c_{2Q} \in V_k$, $k \geq 1$. By the estimate (5.21), the line $\ell_v$ ought to be chosen so that $\text{dist}(x, \ell_v)$ is nicely under control for all $x \in V_{k-1} \cap V_k$ nearby $v$. To do this, we consider a certain cube $Q \supset Q$, which is so large that $2Q$ contains not only $2Q$, but also $2Q'$ for all $Q' \in \mathcal{T} \cap [\mathcal{D}_k \cup \mathcal{D}_{k-1}]$ with
\[
c_{2Q'} \in [V_{k-1} \cup V_k] \cap B(v, 65A2^{-(k+k_0)})
\]
Here $A \geq 1$ is any (dimensional) constant so that $(V^\downarrow)$ and $(V^\uparrow)$ hold. While $\hat{Q}$ needs to be significantly larger than $Q$, it is clear that it can be chosen so that
\[
\ell(\hat{Q}) \leq A\ell(Q) = A\ell v 2^{-(k+k_0)}
\] (5.23)
for some (dimensional) constant $A' \geq 1$. Now we can define $\ell_v$. Let $\ell_v$ be any line $\ell$ such that
\[
\beta_{\mu,p}(2\hat{Q}, \ell) \leq 2\beta_{\mu,p}(2\hat{Q}) =: \alpha_v
\] (5.24)
Now, to estimate \(\text{dist}(x, \ell_v)\) for \(x \in [V_{k-1} \cup V_k] \cap B(v, 65.42^{-(k+k_0)})\), we use the "centre of mass lemma", Lemma 5.7, which also played a role in the proof of Lemma 5.6. Fix \(x = c_2Q' \in [V_{k-1} \cup V_k] \cap B(v, 65.42^{-(k+k_0)})\), so that \(2Q' \subset 2\tilde{Q}\), and first observe that

\[
\text{dist}(x, \ell_v) \leq \beta_{\mu,p}(2Q', \ell_v) \text{ diam}(2Q')
\]

\[
= \frac{1}{\mu(2Q')} \int_{2Q'} \text{dist}(x, \ell_v)^p \, d\mu x
\]

\[
\leq \frac{1}{\mu(2Q')} \int_{2\tilde{Q}} \text{dist}(x, \ell_v)^p \, d\mu x.
\]

Next, since \(Q' \in \mathcal{T}\) contains a point of \(\text{spt} \mu\), we have

\[
\mu(2Q') \gtrsim D_{\mu,n} \mu(2\tilde{Q}),
\]

where the implicit constants depend on the doubling of \(\mu\) and \(A'\) from (5.23). It follows that

\[
\text{dist}(x, \ell_v)^p \lesssim D_{\mu,n} \frac{1}{\mu(2\tilde{Q})} \int_{2\tilde{Q}} \text{dist}(x, \ell_v)^p \, d\mu x = \beta_{\mu,p}(2\tilde{Q}, \ell_v)^p \text{ diam}(2\tilde{Q})^p.
\]

Recalling (5.24),

\[
\text{dist}(x, \ell_v) \lesssim \alpha_v \cdot \text{diam}(2\tilde{Q}) \lesssim D_{\mu,n} \alpha_v \cdot 2^{-(k+k_0)}.
\]

Now, it remains to apply Theorem 5.20. Since \(\text{spt} \mu\) was assumed compact, it is evident that the sets \(V_k\) converge to \(\text{spt} \mu\) in the Hausdorff metric. So, according to Theorem 5.20, the support of \(\mu\) can be covered by a single curve \(\Gamma\) of length

\[
\mathcal{H}^1(\Gamma) \lesssim \beta_0 + \sum_{k \in \mathbb{N}} \sum_{v \in V_k} \alpha_v^2 2^{-(k+k_0)} \lesssim D_{\mu,n} \text{ diam}(\text{spt} \mu) + \sum_{k \in \mathbb{N}} \sum_{v \in V_k} \beta_{\mu,p}^2(2\tilde{Q}_v) 2^{-(k+k_0)}.
\]

Here \(\tilde{Q}_v\) is, of course, the cube \(\tilde{Q}\) associated with \(Q\), if \(v = c_2Q\). To complete the proof, it suffices to note that (1) all the cubes \(\tilde{Q}_v\) arising this way are contained in \(\lambda [\text{spt} \mu]\) from some dimensional constant \(\lambda \geq 1\), and (2) every cube \(\tilde{Q}_v\) is only repeated a bounded number of times in the sum above (where "bounded" depends only on \(n\)). Hence,

\[
\mathcal{H}^1(\Gamma) \lesssim D_{\mu,n} \text{ diam}(\text{spt} \mu) + \sum_{Q \subset [\lambda [\text{spt} \mu]]} \beta_{\mu,p}^2(2Q) \ell(Q),
\]

and the proof of Theorem 2.11 is complete, except for the geometric part in Theorem 5.20 (where the main work lies, of course). \(\square\)

6. Rectifiability of sets and measures, and the Cauchy transform

In this section, we discuss various connections between rectifiability, measures, and the Cauchy transform.

6.1. The theorem of Mattila-Melnikov-Verdera. The first goal is to prove the theorem of Mattila, Melnikov and Verdera [13] from 1996:

**Theorem 6.1.** Let \(E \subset \mathbb{C}\) be a 1-AD regular set such that the Cauchy transform associated to \(\mathcal{H}^1|_E\) is bounded on \(L^2(\mathcal{H}^1|_E)\). Then, the set \(E\) is uniformly 1-rectifiable.
Other items on the menu are a theorem of G. David in Section 6.2, and the "Denjoy conjecture" (now a theorem) in Section 6.3.

To begin with, we briefly recall, what the "$L^2$-boundedness of the Cauchy transform" means. Let $\mu$ be a Radon measure on $\mathbb{C}$ For $\delta > 0$, let $C_{\mu, \delta}$ be the operator formally defined by

$$C_{\mu, \delta}f(z) := \int_{|z-w| \geq \delta} \frac{f(w) \, d\mu w}{z - w}.$$  

In this section, we only need to define $C_{\mu, \delta}f$ for bounded compactly supported functions $f : \mathbb{C} \to \mathbb{C}$, and then the integral above converges for every $z \in \mathbb{C}$. Now, the hypothesis that the Cauchy transform associated to $\mu$ is bounded on $L^2(\mu)$, yields $\mu$-means. Let $\delta > 0$. Since this holds for all $\delta > 0$, it implies that $\mu$ must have linear growth $\mu(B(x, r)) \lesssim r$, see Proposition 6.9 below, and then the definition of $C_{\mu, \delta}f(z)$ makes sense for all $f \in L^2(\mu)$, $z \in \mathbb{C}$, by one application of the Cauchy-Schwarz inequality. This information will not be required in the current section, however.

A main idea behind the proof of Theorem 6.1 is the following striking equation, due to M. Melnikov: if $\mu$ is compactly supported and satisfies $\mu(B(x, r)) \lesssim r$, then

$$\|C_{\mu, \delta}(1)\|_{L^2(\mu)}^2 = \frac{1}{6} c_3^2(\mu) + O(\mu(\mathbb{C})),$$  

where $|O(\mu(\mathbb{C}))| \leq C \mu(\mathbb{C})$, and $C$ only depends on the constants in the linear growth hypothesis $\mu(B(x, r)) \lesssim r$. The function 1 should be interpreted as $\chi_{spt, \mu}$, which is now bounded and compactly supported by hypothesis. I will not prove this inequality in the lecture notes, because it was discussed in the first half of the course by Henri; see Proposition 3.3 in Tolsa’s book [17]. The quantity $c_3^2(\mu)$ is the ($\delta$-truncated) Menger curvature

$$c_3^2(\mu) := \int_{|z-y| > \delta} \int_{|y-z| > \delta} \int_{|x-z| > \delta} c(x, y, z)^2 \, d\mu x \, d\mu y \, d\mu z,$$

where

$$c(x, y, z)^2 := \frac{\text{dist}(z, L_{x,y})^2}{|z - x|^2 |z - y|^2}, \quad z \notin \{x, y\},$$

and $L_{x,y}$ is the line spanned by $x$ and $y$ for $x \neq y$. The numbers $c_3^2(\mu)$ above are positive and increase as $\delta \searrow 0$, so the limit $\delta^2(\mu) \in [1, \infty]$ exists (and has the obvious integral representation). Now, applying equation (6.2) to the restricted measures $\mu|_B$, where $B$ is any bounded Borel set, yields

$$\|C_{\mu, \delta}(\chi_B)\|_{L^2(\mu|_B)}^2 = \|C_{\mu|_B, \delta}(1)\|_{L^2(\mu|_B)}^2 = \frac{1}{6} c_3^2(\mu|_B) + O(\mu(B)).$$

In particular, under the $L^2$-boundedness hypothesis of Theorem 6.1, we get $c_3^2(\mu|_B) \lesssim \|C_{\mu, \delta}(\chi_B)\|_{L^2(\mu)}^2 + \mu(B) \lesssim \|\chi_B\|_{L^2(\mu)}^2 + \mu(B) \sim \mu(B)$.

Since this holds for all $\delta > 0$ uniformly, the conclusion is that

$$\int \int \int_{B \times B \times B} c(x, y, z)^2 \, d\mu x \, d\mu y \, d\mu z \lesssim \mu(B), \quad B \subset \mathbb{C}.\quad (6.3)$$

This is all the information we will need to prove the uniform rectifiability of $\mu$ (Theorem 6.1), so the precise form of the Cauchy transform actually plays a very small role in these
lecture notes. Since 1-AD regular measures are evidently smooth and doubling, the plan is simply to deduce the Carleson condition

$$
\sum_{Q \in \mathcal{D}_{\text{opt},\mu}} \beta_{\mu,2}^2(2Q) \ell(Q) \lesssim \ell(R), \quad R \in \mathcal{D},
$$

(6.4)

from (6.3), and then use Corollary 2.12 to infer that $\mu$ is uniformly rectifiable. In fact, the proof below will show that, for every smooth and doubling measure $\mu$, the condition (6.3) implies

$$
\sum_{Q \in \mathcal{D}_{\text{opt},\mu}} \beta_{\mu,2}^2(2Q) \Theta(2Q)^3 \ell(Q) \lesssim \mu(10R), \quad R \in \mathcal{D},
$$

(6.5)

where $\Theta(2Q)$ is the density ratio $\Theta(Q) = \mu(2Q)/\ell(2Q)$. For 1-AD regular measures $\Theta(2Q) \sim 1$ for all $Q \in \mathcal{D}_{\text{opt},\mu}$ and $\mu(10R) \lesssim \ell(R)$, so (6.5) immediately gives (6.4). I do not know, if (6.5) alone would imply uniform rectifiability for doubling measures.

To prove (6.5), write $\text{spt } \mu =: E$. For $Q \in \mathcal{D}_E$, and any distinct points $x,y \in E \cap 10Q$, note that

$$
\beta_{\mu,2}^2(2Q) \ell(2Q) \leq \frac{\ell(2Q)}{\mu(2Q)} \int_{2Q} \frac{\text{dist}(z, L_{x,y})^2}{\text{diam}(2Q)^2} \, d\mu z
\lesssim \frac{\ell(2Q)^3}{\mu(2Q)} \int_{2Q} \frac{\text{dist}(z, L_{x,y})^2}{|z-x|^2 |z-y|^2} \, d\mu z = \frac{\ell(2Q)^3}{\mu(2Q)} \int_{2Q} c(x,y,z)^2 \, d\mu z.
$$

(6.6)

Since $\mu$ is smooth, there is a constant $c > 0$ such that the annulus

$$
A_{x,Q} := B(x, \ell(Q)) \setminus B(x, c\ell(Q)) \subset 10Q
$$

(6.7)

has measure $\mu(A_{x,Q}) \sim \mu(2Q)$ for all $x \in E \cap 2Q$. Consequently, averaging the inequality (6.6) over all $x \in E \cap 2Q$ and $y \in E \cap A_{x,Q} \subset E \cap 10Q$ gives

$$
\beta_{\mu,2}^2(2Q) \ell(2Q) \lesssim \Theta(2Q)^{-3} \int_{x \in 2Q} \int_{y \in A_{x,Q}} \int_{z \in 2Q} c(x,y,z)^2 \, d\mu x \, d\mu y \, d\mu z.
$$

It remains to sum this inequality over all the cubes $Q \in \mathcal{D}_E$ with $Q \subset R$:

$$
\sum_{Q \in \mathcal{D}_E} \sum_{Q \subset R} \beta_{\mu,2}^2(2Q) \Theta(2Q)^3 \ell(2Q) \lesssim \sum_{Q \in \mathcal{D}_E} \int_{x \in 2Q} \int_{y \in A_{x,Q}} \int_{z \in 2Q} c(x,y,z)^2 \, d\mu x \, d\mu y \, d\mu z
\leq \int \int_{2R \times 2R} \left[ \sum_{Q \subset R} \int_{x \in E \cap 2Q} c(x,y,z)^2 \, d\mu y \right] \, d\mu x \, d\mu z.
$$

Now, it suffices to note that for $x \in 2R$ fixed, the annuli $A_{x,Q}$ have bounded overlap as $Q \subset R$ varies in the collection of squares with $2Q \ni x$:

$$
\sum_{Q \subset R} \chi_{A_{x,Q}} \lesssim \chi_{10R}.
$$

(6.8)

To see this, fix $y \in \mathbb{C}$. Assume that $y \in A_{x,Q}$ for some $Q \subset R$ with $x \in E \cap 2Q$. This forces $y \in 10Q \subset 10R$ by (6.7). Moreover, the requirement $y \in A_{x,Q}$ forces $\ell(Q) \sim |x-y|$, so there are only $\lesssim 1$ side-lengths $2^{-k} = \ell(Q)$ such that this can happen. And for any fixed
side-length $2^{-k}$, there are only $\lesssim 1$ squares $Q$ with $\ell(Q) = 2^{-k}$ and $x \in 2Q$. Combining these facts, there are $\lesssim 1$ squares $Q$ such that $\chi_{A_{x,Q}}(y) = 1$, and this gives (6.8).

All in all, we have now proven that

$$\sum_{Q \in D_k} \beta^2_{\mu}(2Q) \Theta(2Q)^3 \ell(Q) \lesssim \int \int \int_{10R \times 10R \times 10R} c(x, y, z)^2 \, d\mu x \, d\mu y \, d\mu z \quad \text{(6.3)}.$$ 

This is (6.5), so the proof of Theorem 6.1 is complete.

6.2. David’s theorem. This short section contains the following result of G. David [5]:

**Proposition 6.9.** Assume that $\mu$ is a non-atomic Radon measure, and $C_\mu$ is bounded on $L^2(\mu)$. Then $\mu(B(x, r)) \leq C_2 r$ for all balls $B(x, r) \subset \mathbb{R}^2$, where $C_2$ only depend on the $L^2$-boundedness constant for $C_\mu$.

**Remark 6.10.** Note that the non-atomicity is essential. For instance, if $\mu = \delta_0$, then $C_{\mu, \delta}$ is trivially bounded on $L^2(\mu)$ for all $\delta > 0$, because in fact $C_{\mu, \delta}$ equals the zero-operator on $L^2(\mu)$ for any $\delta > 0$. Indeed,

$$C_{\mu, \delta} f(0) = \int_{|w| > \delta} \frac{f(w) \, d\mu(w)}{w} = 0$$

regardless of $f \in L^2(\mu)$.

**Proof of Proposition 6.9.** Assume that $\|C_{\mu, \delta}\|_{L^2(\mu) \rightarrow L^2(\mu)} \leq C_1$, $\delta > 0$, but $\mu$ fails to satisfy the uniform bound $\mu(B(x, r)) \leq 10C_2 r$, for some large $C_2$. This is used to show that $\mu$ has an atom, if $C_2$ is large enough, depending only on $C_1$.

Let $Q_0 \subset \mathbb{R}^2$ be some initial square, not necessarily dyadic, with

$$\Theta_0 := \Theta(Q_0) := \frac{\mu(Q_0)}{\ell(Q_0)} \geq C_2.$$ 

Now, the first step of the plan is to find a significantly smaller sub-square $Q_1 \subset Q_0$ with nearly the same mass as $Q_0$. More precisely, the claim is that for all $N \in \mathbb{N}$, there exists $Q_1 \subset Q_0$ with $\ell(Q_1) \leq \ell(Q_0)/2^N$, satisfying

$$\mu(Q_1) \geq \left(1 - \frac{A}{\Theta_0^2}\right) \mu(Q_0), \quad (6.11)$$

where $A = A(C_1, N) \geq 1$ is a suitable constant (the choice $N = 2$ will work for us in the end). To simplify notation slightly, assume that $Q_0 = [0, 1]^2$, so $\ell(Q_0) = 1$. One may also assume that $N$ is very large to begin with, because the claim is weaker for small $N$.

If (6.11) fails for all squares $Q_1 \subset Q_0$ with $\ell(Q_1) \leq \ell(Q_0)/2^N$, then in particular

$$\mu(2^{8N}Q_1 \cap Q_0) < \left(1 - \frac{A}{\Theta_0^2}\right) \mu(Q_0) \quad (6.12)$$

for all squares $Q_1 \subset Q_0$ with $\ell(Q_1) = 1/2^{10N}$ (since $\ell(2^{8N}Q_1) \leq 1/2^N$ for such squares $Q_1$, noting that in general $\ell(MQ) = (2M - 1)\ell(Q)$ for $M > 1$). Now, pick $Q_1 \in D_{10N} = \{Q \in D : Q \subset Q_0$ and $\ell(Q) = 2^{-10N}\}$ (using the pigeonhole principle) so that

$$\mu(Q_1) \geq \frac{\mu(Q_0)}{2^{20N}}. \quad (6.13)$$
Since (6.12) holds for $Q_1$, 
\[
\mu(Q_0 \setminus 2^{8N}Q_1) \geq \left( \frac{A}{\Theta_0^2} \right) \mu(Q_0),
\]
which implies the existence of another square $Q_2 \in \mathcal{D}_{10N}$ with
\[
Q_2 \subset Q_0 \setminus 2^{8N}Q_1 \quad \text{and} \quad \mu(Q_2) \geq \left( \frac{A}{\Theta_0^2} \right) \frac{\mu(Q_0)}{2^{20N}}. \tag{6.14}
\]
In particular,
\[
\text{dist}(Q_1, Q_2) \geq \ell(2^{8N}Q_1) \sim \frac{1}{2^N}. \tag{6.15}
\]
At this point, consider $C_{\mu, \delta}(xQ_2)(z)$ for $0 < \delta < \text{dist}(Q_1, Q_2)$ and $z \in Q_1$:
\[
C_{\mu, \delta}(xQ_2)(z) = \int_{|w-z| \geq \delta} \frac{\chi_{Q_2}(w)}{z-w} \, d\mu(w) = \int_{Q_2} \frac{d\mu(w)}{z-w}.
\]
For $z \in Q_1$ and $w \in Q_2$, the vectors $z-w$ have essentially constant direction, because the squares $Q_1, Q_2 \in \mathcal{D}_{10N}$ are tiny compared to their separation by (6.15), if $N$ is large. In particular, $N$ can be chosen so large (see computations below) that
\[
|C_{\mu, \delta}(xQ_2)(z)| \sim \int_{Q_2} \frac{d\mu(w)}{|z-w|} \geq \mu(Q_2), \quad z \in Q_1. \tag{6.16}
\]
To prove (6.16) rigorously, let $z_0 \in Q_1$ and $w_0 \in Q_2$ be any points. Then,
\[
\left| \int_{Q_2} \frac{d\mu(w)}{z-w} \right| \geq \left| \int_{Q_2} \frac{d\mu(w)}{z_0-w_0} \right| - \int_{Q_2} \left| \frac{1}{z-w} - \frac{1}{z_0-w_0} \right| \, d\mu(w),
\]
Since $z_0, w_0 \in Q_2 \subset Q_0 = [0, 1]^2$, the first term on the right hand side is $\geq \mu(Q_2)$, as desired in (6.16). For the second term, note that
\[
\left| \frac{1}{z-w} - \frac{1}{z_0-w_0} \right| \leq \frac{|z-z_0|}{|z-w||z_0-w_0|} + \frac{|w-w_0|}{|z-w||z_0-w_0|} \lesssim 2^N \cdot 2^{-10N} = \frac{1}{2^{8N}}
\]
for $z \in Q_1$ and $w \in Q_2$ using (6.15). This gives (6.16) for large enough $N$.

It now follows from (6.16) that, and the $L^2$-boundedness hypothesis on $C_\mu$, that
\[
C_1 \mu(Q_2)^{1/2} \geq \|C_{\mu, \delta}(xQ_2)\|_{L^2(\mu)} \geq \mu(Q_2)\mu(Q_1)^{1/2}.
\]
Combined with the measure estimates (6.13) and (6.14) (and recalling $\ell(Q_0) = 1$, which implies $\Theta_0 = \mu(Q_0)$), this gives
\[
\frac{A^{1/2}}{2^{20N}} = \left( \frac{A}{\Theta_0^2} \right)^{1/2} \frac{\mu(Q_0)}{2^{20N}} \leq \mu(Q_1)^{1/2} \mu(Q_2)^{1/2} \lesssim C_1,
\]
which is impossible for any choice of $A \gg C_1^{2^40N}$. The conclusion is that for some $A \sim C_1^{2^40N}$, there necessarily exists a square $Q_1 \subset Q_0$ with $\ell(Q_1) \leq \ell(Q_0)/2^N$, and satisfying (6.11).

To complete the proof of Proposition 6.9, the observation is iterated to find a sequence of (closed) squares $Q_0 \supset Q_1 \supset \ldots$, where the $\mu$-measure decays so slowly that the intersection $\cap Q_j$ must be an atom for $\mu$. More precisely, start with any (closed) square $Q_0$ with density $\Theta_0 \geq C_2$, as before, and assume that $Q_j$ has been constructed for some $j$,
with density \( \Theta_j = \mu(Q_j)/\ell(Q_j) \geq C_2 \). Apply the construction above with \( N = 2 \), say, to find \( Q_{j+1} \subset Q_j \) with \( \ell(Q_{j+1}) \leq \ell(Q_j)/2^2 \) and

\[
\Theta_{j+1} := \frac{\mu(Q_{j+1})}{\ell(Q_{j+1})} \geq 1 - \frac{A}{\Theta_j^2} \frac{\mu(Q_{j})}{\ell(Q_{j+1})} \geq 2^2 \left( 1 - \frac{A}{\Theta_j^2} \right) \Theta_j.
\]

Now, the key point is that if \( C_2 \) is sufficiently large (depending only on \( A \), which only depends on \( C_1 \)), then the inequality above shows that \( \Theta_{j+1} \geq 2\Theta_j \geq C_2 \). So, the construction can proceed, and one obtains \( \Theta_{j+1} \geq 2\Theta_j \) for all \( j \in \mathbb{N} \). Finally,

\[
\mu(Q_j) \geq \left( 1 - \frac{A}{\Theta_{j-1}^2} \right) \mu(Q_{j-1}) \geq \cdots \geq \left( 1 - \frac{A}{\Theta_0^2} \right)^{j-1} \mu(Q_0) \geq \mu(Q_0)
\]

for uniformly for all \( j \geq 1 \), because the infinite product

\[
\prod_{j=0}^{\infty} \left( 1 - \frac{A}{\Theta_j^2} \right)
\]

converges to a positive number (because \( \sum (A/\Theta_j^2) \leq C_1 \sum 100^{-j} < \infty \)). This proves that

\[
\mu \left( \bigcap_{j \geq 0} Q_j \right) > 0,
\]

as desired. \( \square \)

6.3. The Denjoy conjecture (aka Calderón’s theorem). In this section, which very closely follows (parts of) Section 4 in [17], we use the notation

\[
C_\delta \mu(z) := C_{\mu, \delta} \mu(z).
\]

where \( \mu \) is a finite measure. This allows us to view \( C_\delta \) as an operator acting on the space of complex measures \( M(\mathbb{C}) \).

The historical motivation for studying the connection between geometry, and the boundedness of the Cauchy transform, was to better understand removable sets for bounded analytic functions.

**Definition 6.17.** A compact set \( E \subset \mathbb{C} \) is called removable for bounded analytic functions, or just removable if every bounded analytic function \( f: \mathbb{C} \setminus E \to \mathbb{C} \) is constant.

In particular, every bounded analytic function \( f: \mathbb{C} \setminus E \to \mathbb{C} \) can be extended to an entire function \( f: \mathbb{C} \to \mathbb{C} \); this is why \( E \) is called "removable".

Subsets of lines with positive length are not removable; this is probably a "folklore" result, and I could not find a reference. As early as 1909, A. Denjoy attempted to prove the same for subsets of rectifiable curves (as opposed to subsets of lines). His proof contained a gap, and the statement became known as the Denjoy conjecture. It was resolved by A. Calderón in 1977, who proved that the Cauchy transform is bounded on Lipschitz graphs with sufficiently small constant. We have seen this result – even without the "small constant" restriction – on the course, in Henri’s half. So, the main content of this section is to show, why Calderón’s theorem implies the Denjoy conjecture.

**Theorem 6.18** (Denjoy’s conjecture, or Calderón’s theorem). Let \( \gamma \subset \mathbb{C} \) be a continuum of finite length, and let \( E \subset \gamma \) be a compact subset of positive length. Then \( E \) is not removable.
First of all, the property of being "non-removable" is clearly monotone: if $E_1 \subset E_2 \subset \mathbb{C}$ are compact, and $E_1$ is not removable, then $E_2$ is not removable either (just note that any non-constant bounded analytic function $f : \mathbb{C} \setminus E_1 \to \mathbb{C}$ is also non-constant on $\mathbb{C} \setminus E_2$, by basic properties of analytic functions).

**Lemma 6.19.** Let $\gamma$ be a continuum of finite length, and let $E \subset \gamma$ be a compact subset with $\mathcal{H}^1(E) > 0$. Then, there exists a Lipschitz graph $\Gamma$ such that $\mathcal{H}^1(E \cap \Gamma) > 0$.

**Proof.** Exercise. $\square$

By the monotonicity of non-removability, Denjoy’s conjecture follows, if we can prove the next theorem:

**Theorem 6.20.** Let $\Gamma \subset \mathbb{C}$ be a Lipschitz graph, and let $E \subset \Gamma$ be a compact set with $\mathcal{H}^1(E) > 0$. Then $E$ is not removable.

By the results on the first half of the course (or see [17, Theorem 2.18]), we know the following:

**Proposition 6.21.** Let $E \subset \Gamma$ is as in Theorem 6.20, and let $\mu := \mathcal{H}^1|_E$. Then the Cauchy transform maps $M(\mathbb{C})$ to $L^{1,\infty}(\mu)$ boundedly, which means, by definition, that

$$\mu(\{x \in \mathbb{C} : |\mathcal{C}_\delta \nu(x)| > \lambda\}) \lesssim \frac{\|\nu\|}{\lambda},$$

(6.22)

for all complex Borel measures $\nu$, with implicit constants independent of $\delta > 0$.

The previous proposition is, by far, the hardest part in the proof of Theorem 6.20. Now, Theorem 6.20 will follow, if we just manage to prove the following proposition:

**Proposition 6.23.** Assume that $\mu$ is a Radon measure with $E := \text{spt} \mu$ compact, satisfying $\mu(B(x, r)) \lesssim r$ and the weak-(1,1) bound (6.22). Then $E$ is not removable. More precisely, there exists a function $h : E \to [0, 1]$ with $\int h \, d\mu \gtrsim \mu(E)$, such that $C_{\mu, h}(z) := C(h \, d\mu)(z) := \int \frac{h(w) \, d\mu}{z - w}$, $z \in \mathbb{C} \setminus E$,

defines a non-constant bounded analytic function on $\mathbb{C} \setminus E$.

The proof of the proposition requires two facts from functional analysis. The first is a corollary of the Hahn-Banach theorem:

**Theorem 6.24.** Let $(V, \| \cdot \|)$ be a Banach space, and let $B_1, B_2 \subset V$ be disjoint non-empty convex subsets. Assuming that $B_2$ is open, there exists a number $r \in \mathbb{R}$ and a continuous linear map $\lambda : V \to \mathbb{C}$ such that

$$\Re \lambda(x_1) > r \geq \Re \lambda(x_2) \quad \text{for all } x_1 \in B_1 \text{ and } x_2 \in B_2.$$

**Proof.** See Rudin’s *Functional Analysis* [16], Theorem 3.4(a). $\square$

The second fact identifies the space of complex Radon measures, namely $M(\mathbb{C})$, as the dual of a certain function space:

**Theorem 6.25.** Let $C_0(\mathbb{C})$ be the vector space of all continuous function $\mathbb{C}$, which vanish at infinity: $\varphi \in C_0(\mathbb{C})$, if and only if $\varphi$ is continuous, and for every $\epsilon > 0$, there exists $R_\epsilon > 0$ such that $|\varphi(x)| \leq \epsilon$ for $|x| \geq R_\epsilon$. Equipped with the usual sup-norm, $C_0(\mathbb{C})$ is a Banach space. The dual of $C_0(\mathbb{C})$ is the Banach space $M(\mathbb{C})$, equipped with the total variation norm. More precisely,
if $\Lambda: C_0(\mathbb{C}) \to \mathbb{C}$ is a continuous linear map, then there exists measure $\nu = \nu_\Lambda \in M(\mathbb{C})$ such that

$$\Lambda(\varphi) = \int \varphi \, d\nu, \quad \varphi \in C_0(\mathbb{C}).$$

Proof. This is a version of the Riesz representation theorem, see Rudin’s Real and Complex Analysis, Theorem 6.19. $\square$

Finally, we define a slightly non-standard notion of adjoint. Assume that $T: M(\mathbb{C}) \to C_0(\mathbb{C})$ is a linear map. You should think that $T = C_\delta$, although $C_\delta$ need not quite map $M(\mathbb{C})$ to $C_0(\mathbb{C})$ (we will turn to this issue a bit later). Then, assume that there is another linear map $T^*: M(\mathbb{C}) \to C_0(\mathbb{C})$, satisfying the following relation:

$$\int (T \nu_1) \, d\nu_2 = \int (T^* \nu_2) \, d\nu_1 \tag{6.26}$$

for all $\nu_1, \nu_2 \in M(\mathbb{C})$. Then, we call $T^*$ an adjoint of $T$. We are not claiming uniqueness, continuity, or any other properties usually associated with the notion of “adjoint”. For $T = C_\delta$, an adjoint is simply given by $T^* = -C_\delta$, because

$$\int (C_\delta \nu_1(z)) \, d\nu_2 z = \int \int_{|z-w| \geq \delta} \frac{d\nu_1 w}{z-w} \, d\nu_2 z = \int \left[ -\int_{|z-w| \geq \delta} \frac{d\nu_2 z}{w-z} \right] \, d\nu_1 w = \int (-C_\delta \nu_2(w)) \, d\nu_1 w.$$

Note that if $\mu$ is a measure such that the operators $T = C_\delta$ satisfy (6.22), then clearly $T^* = -C_\delta$ satisfies (6.22) with the same implicit constants.

With that in mind, we prove the following abstract variant of Proposition 6.23:

**Proposition 6.27.** Assume that $T: M(\mathbb{C}) \to C_0(\mathbb{C})$ is a linear operator, and let $T^*: M(\mathbb{C}) \to C_0(\mathbb{C})$ be an adjoint satisfying (6.26). Assume that $\mu$ is a (positive) Radon measure with compact support $E := \text{spt} \, \mu$ such that $T^*$ maps $M(\mathbb{C})$ to $L^{1,\infty}(\mu)$ in the familiar sense that

$$\mu(\{ x \in \mathbb{C} : |T^* \nu(x)| > \lambda \}) \leq C \frac{\|\nu\|}{\lambda}, \quad \nu \in M(\mathbb{C}). \tag{6.28}$$

Then, there exists a Borel function $h: E \to [0,1]$ such that $\|h \, d\mu\| \geq \|\mu\|/2$ and

$$\|T(h \, d\mu)\| < 3C.$$

Proof. Suppose that the claim fails: whenever $h \, d\mu$ lies in

$$G := \{ f \, d\mu : f: E \to [0,1] \text{ is Borel, and } \|f \, d\mu\| \geq \|\mu\|/2 \},$$

one has

$$T(h \, d\mu) \notin B_2 := \{ g \in C_0(\mathbb{C}) : \|g\| \leq 3C \}.$$

Equivalently, $B_1 := T(G)$ is disjoint from $B_2$. Note that both $B_1, B_2$ are convex, and $B_2$ is open. By Theorem 6.24, there exists a continuous linear map $\Lambda: C_0(\mathbb{C}) \to \mathbb{C}$, which we may immediately identify with a measure $\nu = \nu_\Lambda \in M(\mathbb{C})$ by Theorem 6.25, such that

$$\Re \int \varphi_1 \, d\nu > \Re \int \varphi_2 \, d\nu.$$
for all \( \varphi_1 \in B_1 \) and \( \varphi_2 \in B_2 \). Take an inf on the left hand side, and sup on the right hand side: recalling that \( B_1 = T(F) \), this yields

\[
\inf_{f d\mu \in G} \Re \int T(f d\mu) \, d\nu \geq \sup_{\varphi_2 \in B_2} \Re \int \varphi_2 \, d\nu = 3C\|\nu\|.
\]

Now, let \( f := \chi_A \), where

\[
A := \left\{ x \in E : |T^*\nu(x)| \leq \frac{2C\|\nu\|}{\|\mu\|} \right\}.
\]

Then \( f \) is clearly a Borel function taking values in \([0, 1]\). Moreover, by the main assumption (6.28),

\[
\|f \, d\mu\| = \mu(A) = \|\mu\| - \mu(A^c) \geq \|\mu\| - \frac{C\|\nu\|}{2C\|\nu\|/\|\mu\|} = \|\mu\|/2
\]

so \( f \, d\mu \in F \). Hence

\[
3C\|\nu\| \leq \Re \int T(f \, d\mu) \, d\nu \leq \int |T^*\nu| \cdot f \, d\mu \leq \frac{2C}\|\nu\|/\|\mu\| \cdot \mu(A) \leq 2C\|\nu\|.
\]

This is absurd, so the proof is complete. \( \square \)

The only reason, why Proposition 6.27 does not imply Proposition 6.21 directly, is because \( C_\delta \) does not map \( M(\mathbb{C}) \) into \( C_0(\mathbb{C}) \). To fix this little technicality, we need to introduce a "smooth" version of \( C_\delta \). This is fairly standard.

6.3.1. The smooth operators \( \tilde{C}_\delta \). Let \( \varphi : \mathbb{C} \to [0, \infty) \) be some smooth, non-negative radial function with \( \int \varphi = 1 \) and \( \text{spt } \varphi \subset B(0, 1) \). Write \( \varphi_\delta(z) := \delta^{-2}\varphi(x/\delta) \), so that \( \int \varphi_\delta = 1 \), and \( \text{spt } \varphi_\delta \subset B(0, \delta) \). Consider the kernel

\[
\tilde{K}_\delta(z) := \frac{1}{z} * \varphi_\delta.
\]

As the convolution of an \( L^1_{\text{loc}} \)-function, and a compactly supported function, the kernel \( \tilde{K}_\delta \) is continuous, and moreover

\[
|\tilde{K}_\delta(z)| \leq \int \frac{|\varphi_\delta(w)|}{|z-w|} \, dw \leq \|\varphi_\delta\|_\infty \int_{|w| \leq \delta} \frac{dw}{|z-w|} \lesssim \frac{1}{\delta}, \quad z \in \mathbb{C}.
\]

Next, we claim that \( \tilde{K}_\delta(z) = 1/z \) for \( |z| \geq \delta \). To see this, write \( \varphi_\delta(r) \) for the common value of \( \varphi_\delta \) on the circle \( S(0, r) := \{|w - 0| = r\} \) (which exists by radiality). Since \( w \mapsto 1/w \) is harmonic in \( \mathbb{C} \setminus \{0\} \), the average of \( 1/w \) over any circle \( S(z, r) \) not enclosing the origin equals \( 1/z \). This, combined with integration in polar coordinates gives

\[
\tilde{K}_\delta(z) = \int \frac{\varphi_\delta(w)}{z-w} \, dw = c \int_0^\delta r \cdot \varphi_\delta(r) \left[ \int_{S(0,r)} \frac{d\mathcal{H}^1(w)}{z-w} \right] dr
\]

\[
= c \int_0^\delta r \cdot \varphi_\delta(r) \left[ \int_{S(z,r)} \frac{d\mathcal{H}^1(w)}{w} \right] dr = \frac{c}{z} \int_0^\delta r \cdot \varphi_\delta(r) \cdot \mathcal{H}^1(S(0,r)) \, dr
\]

\[
= \frac{1}{z} \int \varphi_\delta(w) \, dw = \frac{1}{z}, \quad \text{for } |z| \geq \delta.
\]
Let $\tilde{C}_\delta$ be the "singular" integral operator associated with $\tilde{K}_\delta$:

$$\tilde{C}_\delta \nu(z) = \int \tilde{K}_\delta(z - w) \, d\nu(w).$$

The fact that $\tilde{K}_\delta$ coincides with $1/z$ outside the ball $B(0, \delta)$ is very convenient for comparing the operators $\tilde{C}_\delta$ and $C_\delta$: if $\nu$ is a complex measure, then

$$|\tilde{C}_\delta \nu(z) - C_\delta \nu(z)| = \left| \int \tilde{K}_\delta(z - w) \, d\nu(w) - \int_{|z - w| \geq \delta} \frac{d\nu(w)}{z - w} \right|$$

$$\leq \int_{|z - w| \leq \delta} |\tilde{K}_\delta(z - w)| \, d|\nu| w \leq \|\tilde{K}_\delta\| : |\nu|(B(z, \delta))$$

$$\lesssim \frac{|\nu|(B(z, \delta))}{\delta} \leq M(|\nu|)(z), \quad z \in \mathbb{C}, \quad \text{(6.29)}$$

where $M(|\nu|)$ is the "radial" maximal function $M(|\nu|)(a) := \sup_{\delta > 0} \delta^{-1} |\nu|(B(z, \delta))$. Recall (or see [17, Theorem 2.5]) that the operator $M$ maps $M(C)$ to $L^{1,\infty}(\mu)$ boundedly, whenever $\mu$ has linear growth (as in Proposition 6.23). So, if $C_\delta$ also maps $M(C)$ to $L^{1,\infty}(\mu)$ boundedly – as assumed in Proposition 6.23 – the conclusion is that the smooth operator $\tilde{C}_\delta$ does the same:

$$\mu(\{x \in \mathbb{C} : |\tilde{C}_\delta \nu(x)| > \lambda\}) \lesssim \frac{||\nu||}{\lambda}. \quad \text{(6.30)}$$

The implicit constants are independent of $\delta > 0$. Note that an adjoint of $\tilde{C}_\delta$ is again given by $(\tilde{C}_\delta)^* = -\tilde{C}_\delta$, repeating the computation from above Proposition 6.27. Now both $\tilde{C}_\delta$ and $(\tilde{C}_\delta)^*$ are linear operators mapping $M(C)$ to $C_0(\mathbb{C})$, and satisfying (6.30).

As a final lemma, we need a comparison between $\tilde{C}_\delta$ and $\tilde{C}_\epsilon$ for $0 < \epsilon \leq \delta$. The proof is so standard that we omit the details (see [17, Lemma 4.4]):

**Lemma 6.31.** Let $\mu$ be a measure satisfying $\mu(B(x, r)) \lesssim r$, and let $\nu$ be any complex measure. Then, for $0 < \epsilon \leq \delta$,

$$||\tilde{C}_\delta \nu|| \leq ||\tilde{C}_\epsilon \nu|| + C||M(|\nu|)||,$$

where $C \geq 1$ is a constant depending only on the function $\varphi$.

We are finally in a position to prove Proposition 6.23.

**Proof of Proposition 6.23.** Let $\mu$ be a measure satisfying the hypotheses of the proposition, with $E = \text{spt} \mu$ compact, so that (6.30) holds for $(\tilde{C}_\delta)^* = -\tilde{C}_\delta$ by the previous discussion. For $\delta > 0$, apply Proposition 6.27 to the operator $\tilde{C}_\delta$: the result is a function $h_\delta : E \to [0, 1]$ such that $||h_\delta \, d\mu|| \geq ||\mu||/2$, and $||\tilde{C}_\delta(h_\delta \, d\mu)|| \lesssim 1$, implicit constants independent of $\delta > 0$.

By Lemma 6.31, we moreover have

$$||\tilde{C}_\delta(h_\epsilon \, d\mu)|| \leq ||\tilde{C}_\epsilon(h_\epsilon \, d\mu)|| + C||M(h_\epsilon \, d\mu)|| \lesssim 1 \quad \text{(6.32)}$$

uniformly for $0 < \epsilon \leq \delta$.

Since $L^{\infty}(\mu)$ is the dual of $L^1(\mu)$, the Banach-Alaoglu theorem states that the sequence $(h_\delta)_{\delta > 0}$ has a weak*-convergent subsequence $(h_j)_{j \in \mathbb{N}} := (h_{\delta_j})_{j \in \mathbb{N}}$ with a limit $h \in \dot{L}^{\infty}(\mu)$. This simply means that

$$\int h_j \cdot g \, d\mu \to \int h \cdot g \, d\mu, \quad g \in L^1(\mu), \quad \text{(6.33)}$$
so in particular (applying the above with $g = 1$), one has $\|h\,d\mu\| \geq \|\mu\|/2$. Now, for $\delta > 0$ and $z \in \mathbb{C}$ fixed, apply (6.33) to $g(z) = K_\delta(z - w)$:

$$|	ilde{C}_\delta(h\,d\mu)(z)| = \left| \int \tilde{K}_\delta(z - w)h(w)\,d\mu w \right| = \lim_{j \to \infty} \left| \int \tilde{K}_\delta(z - w)h_j(w)\,d\mu w \right| \leq \limsup_{j \to \infty} \|\tilde{C}_\delta(h_j\,d\mu)\| \lesssim 1.$$  

The last estimate follows from (6.32). Finally, we infer from (6.29) that

$$\|C_\delta(h\,d\mu)\| \lesssim \|\tilde{C}_\delta(h\,d\mu)\| + \|M(h\,d\mu)\| \lesssim 1$$

uniformly in $\delta > 0$. Consequently, if $z \in \mathbb{C} \setminus E$, we have

$$|C(h\,d\mu)(z)| = \left| \int_E \frac{h(w)\,d\mu w}{z - w} \right| = \lim_{\delta \to 0} \left| \int_{E \cap \{|z - w| > \delta\}} \frac{h(w)\,d\mu w}{z - w} \right| \lesssim 1,$$

which means that $z \mapsto C(h\,d\mu)(z)$ defines a bounded analytic function on $\mathbb{C} \setminus E$. It is an exercise to check that the function is non-constant, hence $E$ is non-removable. The proof of the proposition is complete.  

6.4. Removability of the four corners Cantor set. The purpose of the previous section was to prove that rectifiable sets of positive length are non-removable; now we will see that some purely unrectifiable sets of finite length – including the four corners Cantor set, depicted in Figure 5 – are removable. In fact, all purely unrectifiable sets of finite length are removable; this is a theorem of G. David [6] from 1998. The fact that AD regular purely unrectifiable sets are removable sets was already known a few years earlier. This follows by combining the Mattila-Melnikov-Verdera theorem (Theorem 1.2) with the following result [3, 4] of M. Christ from 1990:

**Theorem 6.34** (Christ). Let $E \subset \mathbb{C}$ be a compact AD regular set. If $E$ is non-removable, then there exists an AD regular set $F \subset \mathbb{C}$ such that $\mathcal{H}^1(E \cap F) > 0$ so that the Cauchy transform associated with $\mu = \mathcal{H}^1|_F$ is bounded on $L^2(\mu)$. In particular (by the later theorem of Mattila-Melnikov-Verdera), $E \cap F$ is uniformly rectifiable.

![Figure 5. The four corners Cantor set.](image)
Since AD regular purely unrectifiable sets certainly cannot contain uniformly rectifiable pieces \( E \cap F \) of positive length, they must be removable.

Unfortunately, the proof of David’s, or even Christ’s, theorem is too long for this course, so we need to take a hands-on approach, proving only the following rather special case.

**Definition 6.35.** Let \( \alpha > 0 \). A Radon measure \( \nu \) on \( \mathbb{C} \) is called \( \alpha \)-non-flat at \( z \), if \( \text{spt} \nu \cap B(z, r) \) is not contained in any cone centred at \( z \), with opening angle \( \alpha \), for any \( r > 0 \). A measure is simply called non-flat at \( z \), if it is \( \alpha \)-non-flat at \( z \) for some \( \alpha > 0 \).

**Theorem 6.36.** Let \( E \subset \mathbb{C} \) be compact and AD regular, and write \( \mu := \mathcal{H}^1|_E \). Assume that for \( \mu \) almost every \( a \in \mathbb{C} \), every tangent measure \( \nu \in \Tan(\mu, a) \) is non-flat at 0. Then E is removable.

This theorem, which is a slightly easier variant of Theorem 19.17 in Mattila’s book [12] (our exposition follows his), quite clearly implies that the four corners Cantor set is removable. It would be easy to relax the hypotheses somewhat: the AD regularity \( \mu \) (our exposition follows his), quite clearly implies that the four corners Cantor set is removable. Then there is a non-constant bounded analytic function \( f \) analytic with \( \|f\|_\infty \leq 1 \) and

\[
f(\infty) := \lim_{z \to \infty} f(z) = 0.
\]

Then, there exists a measure \( \sigma \in M(\mathbb{C}) \) with \( \text{spt} \sigma \subset E \) such that \( |\sigma(B(x, r))| \leq r \) for all discs \( B(x, r) \), and

\[
f(z) = \mathcal{C}(\sigma)(z) = \int \frac{d\sigma w}{z - w}, \quad z \in \mathbb{C} \setminus E.
\]

Moreover, \( \sigma = \varphi \cdot \mathcal{H}^1|_E \) for some function \( \varphi : E \to \mathbb{C} \) with \( \|\varphi\|_{L^\infty(\mathcal{H}^1)} \leq 1 \).

**Proof.** Contained in Janne’s presentation. \( \square \)

Now, in the proof of Theorem 6.36, we start with a counter assumption: \( E \) is not removable. Then there is a non-constant bounded analytic function \( f : \mathbb{C} \setminus E \to \infty \). The limit \( f(\infty) \) exists by elementary theory of analytic functions (note that \( z \mapsto f(1/z) \) is defined in a neighbourhood of the origin, and has a removable singularity at the origin), and \( g = (f - f(\infty))/\|f\|_\infty \) is a a non-constant analytic function satisfying the assumptions of the previous lemma. Let

\[
\sigma = \varphi \, d\mu = \varphi \cdot \mathcal{H}^1|_E
\]

be the measure given by the lemma, associated with \( g \); note that \( \varphi \not\equiv 0 \), because \( g \not\equiv 0 \). Then also \( |\sigma| = |\varphi| \, d\mu \), from which it follows that

\[
|\sigma|(B(x, r)) \lesssim r, \quad x \in \mathbb{C}, \ r > 0.
\]

(6.38)

Since \( \mu \) is AD regular, every tangent measure \( \nu \in \Tan(\mu, a) \), \( a \in E \), has the form

\[
\nu = c \cdot \lim_{i \to \infty} r_i^{-1} T_{a,r_i} \mu,
\]

(6.39)

where \( c > 0 \), and \((r_i)_{i \in \mathbb{N}}\) is a sequence of positive radii tending to zero as \( i \to \infty \). To see this, fix \( a \in E \), and let \( \nu = \lim_{i \to \infty} c_i T_{a,r_i} \mu \in \Tan(\mu, a) \). Then, because \( \nu \not\equiv 0 \) by
assumption (trivial measures are excluded from the definition of tangent measures), we find some some \( R > 0 \) such that

\[
0 < \nu(U(0, R)) \leq \liminf_{i \to \infty} c_i \mu(U(a, Rr_i)) \lesssim \limsup_{i \to \infty} c_i R r_i
\]

\[
\lesssim \limsup_{i \to \infty} c_i \mu(B(a, Rr_i)) \leq \nu(B(0, R)) < \infty.
\]

This implies that the the numbers \( c_i r_i \) lie, for large enough \( i \in \mathbb{N} \), on some compact interval \([a, b] \subset (0, \infty)\), and hence there is a subsequence \( c_i r_i \to c \in [a, b] \). Then, it is easy to check that \( \nu = c \cdot \lim_{i \to \infty} r_i^{-1} T_{a, r_i, \varphi} \sigma \), as claimed in (6.39).

Now, if \( \nu \in \Tan(\mu, a) \) as in (6.39), and \( \varphi \) (as in \( \sigma = \varphi d\mu \)) is non-vanishing and continuous in a neighbourhood of \( a \), it is easy to see (exercise) that

\[
\nu = \tilde{c} \cdot \lim_{i \to \infty} r_i^{-1} T_{a, r_i, \varphi} \sigma
\]

with \( \tilde{c} = c/\varphi(a) \). In general, applying the Lebesgue differentiation theorem to \( \varphi \), one can prove that at \( |\sigma| \) almost every point \( a \in \mathbb{C} \), every tangent measure \( \nu \in \Tan(\mu, a) \) has the form (6.40); since \( \sigma \) is a non-trivial measure, "\( |\sigma| \)-almost every point" implies "\( \mu \)-positively many points".

Further, it follows from (6.39) every \( \nu \in \Tan(\mu, a) \), \( a \in \text{spt} \mu \), is AD regular: if \( U(x, r) \) any open ball (centred on \( \text{spt} \mu \) or not), then

\[
\nu(U(x, r)) \leq c \liminf_{i \to \infty} r_i^{-1} \mu(U(a + x, r r_i)) \lesssim c \liminf_{i \to \infty} r r_i = cr,
\]

For the converse inequality, note that if \( x \in \text{spt} \nu \) and \( r > 0 \), then \( \nu(U(x, r/2)) > 0 \) for all \( r > 0 \), and it follows from the estimate above that \( U(a + x, r r_i/2) \cap E \neq \emptyset \) for all large enough indices \( i \). By the AD regularity of \( \mu \), this implies \( \mu(B(a + x, r r_i)) \gtrsim r r_i \), and so

\[
\nu(B(x, r)) \gtrsim c \limsup_{i \to \infty} r_i^{-1} \mu(B(a + x, r r_i)) \gtrsim cr.
\]

It will also be useful to note that \( 0 \in \text{spt} \nu \) for all \( \nu \in \Tan(\mu, a), a \in E \). To this end, fix \( \nu \) as above, and \( r > 0 \):

\[
\nu(B(0, r)) \geq c \limsup_{i \to \infty} r_i^{-1} \mu(B(0, r r_i)) \gtrsim cr > 0.
\]

In summary, if \( \mu \) is AD regular, then for following hold for all \( a \in \text{spt} \mu \):

- Every \( \nu \in \Tan(\mu, a) \) has the form \( \nu = c \lim_{i \to \infty} r_i^{-1} T_{a, r_i, \varphi} \sigma \), with \( c > 0 \) and \( r_i \to 0 \).
- Every \( \nu \in \Tan(\mu, a) \) is AD regular. In particular \( \text{spt} \nu \neq \mathbb{C} \).
- \( 0 \in \text{spt} \nu \) for all \( \nu \in \Tan(\mu, a) \).

Now, we will show that tangent measures interact well with the Cauchy transform: the maximal Cauchy transform of any tangent of \( \mu \) is bounded at 0:

**Lemma 6.41.** If \( a \in E \) is such that \( \nu \in \Tan(\mu, a) \) is a tangent measure as in (6.40), then

\[
\sup_{0 < r < R < \infty} \left| \int_{B(0, R) \setminus B(0, r)} \frac{d\nu w}{w} \right| < \infty.
\]

**Proof.** The claim will follow, if we find a dense set of radii \( 0 < r < R < \infty \), and a constant \( C \geq 1 \), such that

\[
\left| \int_{B(0, R) \setminus B(0, r)} \frac{d\nu w}{w} \right| \leq C,
\]
After this, for arbitrary radii $0 < r < R < \infty$, we can find sequences $(r_i)$ and $(R_i)$ from the dense collection such that $0 < r < r_i < R_i < R < \infty$, and $r_i \searrow r$ and $R_i \nearrow R$. Then the sets $B(0, R_i) \setminus B(0, r_i)$ converge in a monotone way to $U(0, R) \setminus B(0, r)$, and consequently
\[
\left| \int_{B(0, R) \setminus B(0, r)} \frac{d\nu w}{w} \right| \leq \int_{\partial B(0, R)} \frac{d\nu w}{w} + \lim_{i \to \infty} \int_{B(0, R_i) \setminus B(0, r_i)} \frac{d\nu w}{w} \leq \nu(B(0, R)) + C.
\]
Since $\nu$ is (upper) AD regular, the right hand side has a uniform bound.

Now, fix $a \in E$ as in the hypotheses, and let
\[
\nu = c \cdot \lim_{j \to \infty} r_j^{-1} T_{a, r_j} \sigma \in \Tan(\mu, a),
\]
where $c \in \mathbb{C} \setminus \{0\}$ (note that $c$ has the form "$c = c/\varphi(a)$" from (6.40), so $c$ can have an imaginary part, if $\varphi(a)$ does; this will not affect anything). Next, let $0 < r < R < \infty$ be radii such that $\nu(S(0, r)) = 0 = \nu(S(0, R))$; this holds but all but countably many pairs $0 < r < R < \infty$ and is used to infer that if
\[\int_{B(0, R) \setminus B(0, r)} \psi \, d\nu = \lim_{j \to \infty} \int_{B(0, R) \setminus B(0, r)} \psi \, d\nu_j, \tag{6.42}\]
whenever $\psi$ is a continuous function on $B(0, R) \setminus B(0, r)$ and $\nu_j \to \nu$ weakly.$^6$

After this observation, we just compute as follows:
\[
\left| \int_{B(0, R) \setminus B(0, r)} \frac{d\nu w}{w} \right| \overset{(6.42)}{=} \lim_{j \to \infty} \left| \frac{1}{r_j} \int_{B(0, R) \setminus B(0, r)} \frac{d(T_{a, r_j} \sigma)w}{w} \right| \leq \lim_{j \to \infty} \left| \frac{1}{r_j} \int_{B(a, R_j) \setminus B(a, r_j)} \frac{d\sigma w}{(a - w)/r_j} \right| \leq 2 \cdot C^*(\sigma)(a).
\]

Here $C^*(\sigma)$ is the maximal Cauchy transform
\[C^*(\sigma)(a) = \sup_{\delta > 0} |C_\delta(\sigma)(a)|.\]
(The numbers $C_\delta(\sigma)(a)$ are well-defined, since $\sigma$ is compactly supported; the tangent measures are typically not compactly supported, which explains the need for the double truncation in the formulation of the lemma.) The proof of the lemma is now complete, as soon as we verify that
\[C^*(\sigma)(a) \lesssim \|C(\sigma)\|_{L^\infty(\mathbb{C} \setminus E)} + M(|\sigma|)(a), \tag{6.43}\]
where $M(|\sigma|)$ is the "radial" maximal function $M(|\sigma|)(a) = \sup_{r > 0} |\sigma|(B(a, r))/r$, which is now uniformly bounded by (6.38). Also, recall that $\|C_\sigma\|_{L^\infty(\mathbb{C} \setminus E)} = \|g\|_{L^\infty(\mathbb{C} \setminus E)} < \infty$ by assumption.

$^6$The general principle here is the following: if $\nu_i \to \nu$ weakly, and $B$ is any Borel set with $\nu(\partial B) = 0$, then $\nu_i(B) \to \nu(B)$. This is an exercise.
To prove (6.43), fix $\delta > 0$. We first claim that there exists $b \in B(a, \delta/2) \setminus E$ such that
\begin{equation}
\int_{B(a,\delta)} \frac{d|\sigma|w}{|b-w|} \lesssim M(|\sigma|)(a).
\end{equation}

Assuming this for a moment, we can estimate as follows:
\begin{align*}
|C_\delta(\sigma)(a) - C(\sigma)(b)| &= \left| \int_{C \setminus B(a,\delta)} \frac{d\sigma w}{a-w} - \int_{B(a,\delta)} \frac{d\sigma w}{b-w} \right| \\
&\leq \int_{C \setminus B(a,\delta)} \frac{|a-b|}{|a-w||b-w|} d|\sigma|w + \int_{B(a,\delta)} \frac{d|\sigma|w}{|b-w|}
\end{align*}

The second term is bounded by $M(|\sigma|)(a)$ by (6.44), while in the first term the crucial points to note are the following: $|a-b| \leq \delta$, and $|a-w| \sim |b-w|$ for all $w \in C \setminus B(a,\delta)$, because $b \in B(a,\delta/2)$. So, the first term is comparable to
\[ \delta \cdot \int_{C \setminus B(a,\delta)} \frac{d|\sigma|w}{|a-w|^2} \lesssim \delta \cdot \sum_{j=0}^\infty \frac{1}{2j\delta} \cdot |\sigma|(B(a, 2^{j+1}\delta)) \lesssim M(|\sigma|)(a). \]

Since $\delta > 0$ was arbitrary, and $|C(\sigma)(b)| \leq \|C(\sigma)\|_{L^\infty(C \setminus E)}$, this completes the proof of (6.43), modulo finding the point $b \in B(a,\delta/2) \setminus E$ satisfying (6.44). This is done by a simple averaging trick:
\[ \frac{1}{\delta^2} \int_{B(a,\delta/2)} \left[ \int_{B(a,\delta)} \frac{d|\sigma|w}{|b-w|} \right] d\mathcal{L}^2(b) = \frac{1}{\delta^2} \int_{B(a,\delta)} \left[ \int_{B(a,\delta/2)} \frac{d\mathcal{L}^2(b)}{|b-w|} \right] d|\sigma|w \lesssim \frac{|\sigma|(B(a,\delta))}{\delta}, \]
noting (in the inner integral) that $B(a,\delta/2) \subset B(w,2\delta)$, whenever $w \in B(a,\delta)$. The proof of the lemma is complete.

A very useful property of tangent measures is that "taking tangents twice" does not add much information:

**Lemma 6.45.** Let $\mu$ be a Radon measure on $\mathbb{R}^n$. Then, at $\mu$ almost every $a \in \mathbb{R}^n$, every tangent measure $\nu \in \Tan(\mu, a)$ has the following property: $\Tan(\nu, x) \subset \Tan(\mu, a)$ for all $x \in \text{spt} \nu$.

Taking this result for granted (it is Theorem 14.16 in [12]), we are prepared to prove Theorem 6.36:

**Proof of Theorem 6.36.** The idea is to find a special tangent measure $\lambda \in \Tan(\mu, a)$ with $\text{spt} \lambda$ contained in a half-plane, and then apply Lemma 6.41 to produce a contradiction (against the counter assumption that $E$ is not removable).

Start by choosing any tangent measure $\nu \in \Tan(\mu, a)$, with $a \in E = \text{spt} \mu$ such that the conclusion of Lemma 6.45 is valid, and the hypotheses of Lemma 6.41 are valid. Since $\mu$ is AD regular, $\nu$ is also AD regular by the discussion before Lemma 6.41. Typically $\text{spt} \nu$ is not compact, but anyway the support of a 1-AD regular measure cannot be $C$. So, there is a point $z \in \mathbb{C} \setminus \text{spt} \nu$, which then of course satisfies $\rho := \text{dist}(z, \text{spt} \nu) > 0$. Note that
\begin{equation}
U(z, \rho) \cap \text{spt} \nu = \emptyset \quad \text{and} \quad \partial B(z, \rho) \cap \text{spt} \nu \neq \emptyset,
\end{equation}
see Figure 6. Then, pick a point $y \in \partial B(z, \rho) \cap \text{spt} \nu$. Since $\nu$ is AD regular, there exists a tangent measure
\[ \lambda \in \Tan(\nu, y) \subset \Tan(\mu, a), \]
using Lemma 6.45 in the second inclusion. In particular, $0 \in \text{spt } \lambda$. Moreover, it follows from the first equation in (6.46) that the support of $\lambda$ is entirely contained in some closed half-space $H$, with the $0 \in \partial H$ and $\partial H \perp (z - y)$. For simplicity of notation, assume that $H$ is the lower half-plane $\{(x, y) : y \leq 0\} = \{w : \text{Im } w \leq 0\}$. Now, recall also the main assumption of the theorem: $\lambda$ is $\alpha$-non-flat at $0$ for some $\alpha > 0$. In particular, this implies that $B(0, r) \cap \text{spt } \lambda$ is not contained in $C_\alpha$ for any $r > 0$, where $C_\alpha$ is the cone $C_\alpha := \{w : |\text{Im } w| \leq \alpha |w|\}$, see Figure 7. Since $\text{spt } \lambda \subset H$, the conclusion is that for any $r > 0$, there exists a point $w \in B(0, r) \cap \text{spt } \lambda$ with

$$w \in H \setminus C_\alpha = \{w : \text{Im } w < -\alpha |w|\} =: G_\alpha.$$ 

Note that if $w \in G_\alpha$, then $B(w, \alpha |w|/2) \subset G_{\alpha/2} = \{w : \text{Im } w < -\alpha |w|/2\}$.

With the observations above in mind, it is possible to find a sequence of points $\{w_j\}_{j \in \mathbb{N}}$ in $\text{spt } \lambda$, converging to $0$, such that the balls $B_j := B(w_j, \alpha |w_j|/2)$ are disjoint, and satisfy

$$B_j \subset G_{\alpha/2} \cap \mathbb{C} \setminus B(0, |w_{k+1}|], \quad 1 \leq j \leq k. \quad (6.47)$$

Since $\lambda \in \text{Tan}(\mu, a)$, and $a$ was assumed to satisfy the hypotheses of Lemma 6.41, we have

$$\sup_{\epsilon > 0} \left| \int_{B(0,1) \setminus B(0,\epsilon)} \frac{d\lambda w}{w} \right| < \infty. \quad (6.48)$$
On the other hand, noting that $1/w = \bar{w}/|w|^2$, and using the fact that $\text{Im } w \leq 0$ for all $w \in \text{spt } \lambda$, plus the AD regularity of $\lambda$, we infer that

$$\left| \int_{B(1) \setminus B(0,|u_{k+1}|)} \frac{d\lambda w}{w} \right| \geq \int_{B(1) \setminus B(0,|u_{k+1}|)} \frac{\text{Im } \bar{w}}{|w|^2} d\lambda w \geq \sum_{j=1}^{k} \int_{B_j} -\frac{\text{Im } w}{|w|^2} d\lambda w \geq \frac{\alpha}{2} \sum_{j=1}^{k} \int_{B_j} d\lambda w \geq \frac{\alpha}{2} \sum_{j=1}^{k} -\frac{\lambda(B_j)}{|w_j|} \sim \alpha k.$$  

This contradicts (6.48) for large enough $k$, and the proof of the theorem is complete. \hfill \Box

### 7. The Geometric Construction of M. Badger and R. Schul

This section contains the proof of the geometric construction, Theorem 5.20, which is repeated below for convenience:

**Theorem 7.1.** Let $n \geq 2$, $A > 1$, $x_0 \in \mathbb{R}^n$ and $r_0 > 0$. Let $(V_k)_{k \in \mathbb{N}}$ be a sequence of non-empty finite subsets of $B(x_0, Ar_0)$ such that the following conditions are satisfied:

(V\text{sep}) The distance between distinct points in $V_k$ is at least $2^{-k}r_0$.

(V\text{\dag}) For all $v \in V_k$ there exists $v^{\dag} \in V_{k+1}$ with $|v - v^{\dag}| < 2A^{-k+1}r_0$.

(V\uparrow) For all $v \in V_{k+1}$, there exists $v^{\uparrow} \in V_k$ with $|v - v^{\uparrow}| < 2A^{-k}r_0$.

Further, assume that for all $k \geq 1$ and for all $v \in V_k$ there is a line $\ell_v = \ell_{k,v} \subset \mathbb{R}^n$ and a number $\alpha_v = \alpha_{k,v} \geq 0$ such that

$$\sup_{x \in (V_{k+1} \cup V_k) \cap B(v, 65A2^{-k}r_0)} \text{dist}(x, \ell_v) \leq \alpha_v 2^{-k}r_0. \quad (7.2)$$

Then the sets $V_k$ converge in the Hausdorff metric to a compact set $V \subset \overline{B(x_0, Ar_0)}$, and there exists a compact, connected set $\Gamma \subset \overline{B(x_0, Ar_0)}$ such that $\Gamma \supset V$, and

$$\mathcal{H}^1(\Gamma) \lesssim_{A,n} r_0 + \sum_{k \in \mathbb{N}} \sum_{v \in V_k} \alpha_v^2 2^{-k}r_0. \quad (7.3)$$

To achieve a slight simplification in the proof, I record the following:

**Proposition 7.4.** It suffices to prove Theorem 7.1 under the additional hypothesis that if either

- $k \in \mathbb{N}$ and $v, v' \in V_k$ and $w, w' \in V_k$ are distinct pairs of points, or
- $j < k$ and $v, v' \in V_j$ and $w, w' \in V_k$,

then the intersection $[v, v'] \cap [w, w']$ has zero length.

**Proof.** It is clear that the extra hypothesis can be achieved by perturbing the points in the various sets $V_k$ by arbitrarily small amounts. Note that these perturbations must be made so that no pair in $V_j$ is also contained in $V_k$ for any $k > j$. If every point of $V_k$ is moved by less than $2^{-k}r_0$, then $(V\text{sep}) - (V\uparrow)$ continue to hold. In case some of the numbers $\alpha_v = 0$, then any perturbation may cause (7.2) to fail, but the assumption $\alpha_v > 0$ can be made without loss of generality (check!). Finally, if the perturbations are small enough, the sets $\tilde{V}_k$ have the same limit set as the $V_k$’s. Thus, it suffices to cover the limit set of the $\tilde{V}_k$’s by a continuum $\Gamma$ satisfying (7.3).

The convergence of the sequence $V_k$ is, in fact, rather simple and based on $(V_{III})$ alone:
Proposition 7.5. Let $V_0, V_1, \ldots$ be subsets of some fixed ball $B(x_0, A r_0)$. If the sets $V_k$ satisfy $(\forall k)$, then they converge to a compact set $V \subset \overline{B(x_0, A r_0)}$ in the Hausdorff metric.

Proof. Exercise (or read the paper of Badger and Schul).

The statement of Theorem 7.1 is clearly "scaling invariant", i.e. one may assume $r_0 = 1$.

Also, to avoid trivialities, I will assume that

$$\text{card } V_k \geq 2, \quad k \in \mathbb{N}.$$  

Among other things, the next lemma defines a useful ‘ordering” for the points in $(V_k \cup V_{k-1}) \cap B(v, 65 A 2^{-k})$, assuming that the number $\alpha_{v,k}$ is sufficiently small. The lemma also gives some explanation for the "square" in $\alpha^2_v$.

Lemma 7.6. Let $0 \leq \alpha \leq 1/16$. Assume that $V \subset \mathbb{R}^n$ is a $1$-separated set with $\text{card } V \geq 2$, and there exist lines $\ell_1, \ell_2$ such that

$$\text{dist}(v, \ell_i) \leq \alpha, \quad v \in V, \ i \in \{1, 2\}.$$  

Let $\pi_i$ be the orthogonal projection to $\ell_i$. Then, one may identify both $\ell_1 \ell_2$ with $\mathbb{R}$ in such a way that

$$\pi_1(v) \leq \pi_1(v') \iff \pi_2(v) \leq \pi_2(v'), \quad v, v' \in V.$$  

Moreover, if $v_1, v_2$ are consecutive points relative to the order given by $\pi_1$ (equivalently $\pi_2$), then

$$\mathcal{H}^1([u_1, u_2]) < (1 + 3 \alpha^2) \cdot \mathcal{H}^1([\pi_1(u_1), \pi_1(u_2)]), \quad [u_1, u_2] \subset [v_1, v_2].$$  

Also,

$$\mathcal{H}^1([u_1, u_2]) < (1 + 12 \alpha^2) \cdot \mathcal{H}^1([\pi_1(u_1), \pi_1(u_2)]), \quad [u_1, u_2] \subset \ell_2.$$  

7.1. Construction of the continuums. In this section, the points of $V_k$ will be covered by a (nearly) piecewise linear set $\Gamma_k$, whose connectedness and length will be discussed in later sections. Each set $\Gamma_k$ will consist of a finite number of edges $[v, v']$ between vertices $v, v' \in V_k$, plus a finite number of more complicated connected sets called bridges, to be defined presently.

For a vertex $v \in V_k$, define the extension $E(v) = E(v, k)$ inductively as follows. Let $v_0 = v \in E(v)$, and assume that $v_j$ has been defined for some $j \geq 0$. Set $v_{j+1} := v_j^+$ (this was the closest “next generation” vertex to $v_j$). Then, define

$$E(v) := \bigcup_{j=0}^{\infty} [v_j, v_{j+1}].$$  

Now, the bridge between two generation $k$ vertices $v, v'$ is

$$B(v, v') = B(v, v', k) := [v, v'] \cup E(v) \cup E(v').$$  

Remark 7.7. In the special (but already interesting) case $V_k \subset V_{k+1} \subset \ldots$, the extension $E(v)$ simplifies to $\{v\}$ and $B(v, v') = [v, v'].$
7.1.1. The induction begins. In this tiny subsection, the initial curve $\Gamma_0$ is defined (which covers $V_0$). Consider a pair $v, v' \in V_0$. If $|v - v'| < 30A$, then $[v, v'] \subset \Gamma_0$. Otherwise, set $B(v, v') \subset \Gamma_0$. In other words,

$$\Gamma_0 := \bigcup_{|v-v'|<30A} [v, v'] \cup \bigcup_{|v-v'|\geq 30A} B(v, v').$$

7.1.2. The construction of $\Gamma_k$ based on $\Gamma_{k-1}$. Here comes a key point: bridges stay, edges don’t. Thus, if a bridge is contained in $\Gamma_{k-1}$, then it will also be contained in $\Gamma_k$. The edges of $\Gamma_{k-1}$, however, will be thrown away and replaced by new material (edges and bridges) in $\Gamma_k$. In symbols, $\Gamma_k$ will look like this:

$$\Gamma_k := \bigcup_{v \in V_k} \bigcup_{j=0}^{k-1} \bigcup_{v' \in \Gamma_{j-1} \cap N(v)} B(v', v'').$$

(7.8)

Here $\Gamma_{k,v}$ is a "local part" of $\Gamma_k$ constructed inside the "neighbourhood"

$$\mathcal{N}(v) := B(v, 65A2^{-k}).$$

Note that $\mathcal{N}$ is the ball appearing in (7.2).

The local parts will look very different depending on whether $\alpha_v$ is "large" or "small". Here "small" simply means

$$0 < \alpha_v < \epsilon := 1/32,$$

and "large" means $\alpha_v \geq \epsilon$. The threshold $\epsilon = 1/32$ has been chosen so that Lemma 7.6 can be applied to any number smaller than $2\epsilon$.

Case (L). Assume that $v \in V_k$ and $\alpha_v \geq \epsilon$. Let $v', v'' \in V_k \cap N(v)$. If $|v' - v''| < 30A2^{-k}$, add the edge $[v', v'']$ to $\Gamma_{v,k}$. In the opposite case $|v' - v''| \geq 30A2^{-k}$, add the bridge $B(v', v'')$ to $\Gamma_{v,k}$. The case (L) is complete.

Here (L) obviously stands for "large". The "small" case (S) is more complicated and divides further into various sub-cases. However, before starting, a crucial point is worth emphasising:

Principle. At any stage of the construction, two points $v, v' \in V_k$ will be joined by and a finite sequence of edges in $\Gamma_k$ if and only if $|v - v'| < 30A2^{-k}$. Check that this is the true for the cases above, and keep this in mind in the future!

Case (S). Assume that $\alpha_v < \epsilon$ and note that the "ordering" Lemma 7.6 now applies to all points in $(V_{k-1} \cup V_k) \cap \mathcal{N}(v)$ and the line $\ell_v$ (once the picture is scaled by $2^k$). In particular, once an orientation for $\ell_v$ has been fixed, it makes sense to write things like "$v'$ is to the left from $v''$". I will also write $v' < v''$, if $v'$ is to the left from $v''$.

With this ordering in mind, enumerate the points in $V_k \cap \mathcal{N}(v)$ from left to right as $v_0 < v_1 < v_2 < \ldots < v_m$, where $v_0 = v$. It may happen that $v_0$ is the only element on this list! I will begin by describing how the "right half" $\Gamma^R_v$ of $\Gamma_v$ looks like. This is the part of $\Gamma_v$, whose construction involves the points $v_i$ with $i > 0$ (should any exist, which is neither clear nor assumed at this point). The "left half" will eventually be treated symmetrically.

Start from $v$ and start moving right along the sequence $v_1, v_2, \ldots$. Include the edge $[v_i, v_{i+1}]$ to $\Gamma^R_v$ as long as

$$|v_{i+1} - v_i| < 30A2^{-k} \quad \text{and} \quad v_{i+1} \in B(v, 30A2^{-k}).$$

(7.9)
These conditions could easily hold for all the points \( v_1, \ldots, v_m \). If one of the conditions eventually fails, stop right there! The construction will now divide into sub-cases depending on what happened.

**Subcase (S-NT).** Here "NT" stands for "non-terminal", because this is the sub-case, where the algorithm above produced at least one edge. In other words, \( |v_1 - v| < 30A2^{-k} \), and the edge \([v,v]_1\) (and possibly much more) was added to \( \Gamma_v^R \). The construction of \( \Gamma_v^R \) is complete in this simple sub-case.

**Subcase (S-T).** Here "T" stands for "terminal", because this is the sub-case, where the algorithm above left us empty-handed: either \( v_1 \) does not exist at all, or \( |v_1 - v| \geq 30A2^{-k} \), and no edges were added. In this case the vertex \( v \) will be called terminal to the right (terminal to the left will be defined similarly while constructing \( \Gamma_v^L \)). Now, the construction of \( \Gamma_v^R \) will depend on how the previous generation points \( V_{k-1} \cap \mathcal{N}(v) \) are positioned. Again, since \( \alpha_v \) is very small, and the points in \( V_{k-1} \cap \mathcal{N}(v) \) lie at distance \( \alpha_v 2^{-k} \) from \( \ell_v \), they can be arranged from "left to right" as

\[
\ldots < w_{r-1} < w_r < w_1 < \ldots < w_s,
\]

where \( w_0 = v^* \in B(v, A2^{-k}) \) is the closest point of \( V_{k-1} \) to \( v \). It goes without saying that "left to right" means the same order as with the points \( v_i \) above. Let \( w_r \) be the right-most vertex on the list above, which still lies in \( B(v, 2A2^{-k}) \).

Consider the following two sub-sub-cases:

**(S-TT)** This stands for "terminal terminal", because the definition of \( \Gamma_v^R \) will simply be \( \{v\} \). And what is this case? It occurs, if either \( r = s \) (so everything in \( V_{k-1} \cap \mathcal{N}(v) \) to the right from \( u_0 \) is contained in \( B(v, 2A2^{-k}) \)), or then \( |w_r - w_{r+1}| \geq 30A2^{-(k-1)} \), so there is no \( \Gamma_{k-1} \) edge joining \( w_r \) and \( w_{r+1} \) (recall the Principle!).

**(S-TB)** This stands for "terminal bridge", because now – and only now – a bridge will be added. Since (S-TT) does not occur, the point \( w_{r+1} \) exists and satisfies

\[
|w_r - w_{r+1}| \leq 30A2^{-(k-1)} = 60A2^{-k}.
\]

By the assumption \((V^+)\), one moreover has \( |w_{r+1} - w_{r+1}^+| < A2^{-k} \), which implies that

\[
|w_{r+1} - v| \leq |v - v^*| + |w_r - w_{r+1}| + |w_{r+1} - w_{r+1}^+| < 2A2^{-k} + 30A2^{-k} + A2^{-k} = 63A2^{-k}.
\]

The upshot is that \( w_{r+1}^+ \in V_k \cap \mathcal{N}(v) \setminus \{v\} \) (\( w_{r+1}^+ \neq v \) because it’s quite far away; check!), and hence \( v_1 \) exists. But since we ended up in the "T-cases", we know that \( |v - v_1| \geq 30A2^{-k} \). Now, a bridge \( B(v, v_1) \) is added to \( \Gamma_v^R \), and the construction of \( \Gamma_v^R \) is complete.

The construction of the "left part" \( \Gamma_v^L \) is symmetric. I make one last remark about Case (S-TB). Let \( v_1 \) and \( w_{r+1} \) be as above. It is useful to observe that while nothing necessitates that \( v_1 = w_{r+1} \), we still have

\[
|v_1 - w_{r+1}| < 2A2^{-k}. \tag{7.10}
\]

Indeed, note that \( v_1^* \in V_{k-1} \cap \mathcal{N}(v) \), hence \( v_1 \) must lie "to the right" from \( w_{r+1} \), since \( w_{r+1} \) is the first vertex "to the right" from \( w_r \). Now, if (7.10) failed, the situation would be as in Figure 8 below.

7This is no typo: the "2" should not be "30".
Figure 8. The position of the points $v_1, v_1', w_{r+1}, w_{r+1}'$ in the absurd situation that (7.10) failed. Note how $w_{r+1}'$ must lie to the left of $v_1$, which contradicts the definition of $v_1$ as the first vertex in $V_k$ to the right from $v$.

But since $w_{r+1}' \in V_k \cap N(v)$, it should lie "to the right" from $v_1$, which is now evidently impossible, using $|w_{r+1}' - w_r| < A2^{-k}$, the counterassumption $|v_1 - w_{r+1}| \geq 2A2^{-k}$, and the fact that $\alpha_v < \epsilon$ (think of how the proof would go precisely, if all the points actually lay on $\ell_v$). This proves (7.10).

The construction of $\Gamma_v$ is complete, and the set $\Gamma_k$ is obtained by performing the same algorithm for every point $v \in V_k$ (and keeping the old bridges, as indicated in (7.8)).

8. The Connectedness of $\Gamma_k$

So far, all we (are supposed to) know is that $V_k \subset \Gamma_k$, and each pair $v, v' \in V_k$ with $|v - v'| \leq 30A2^{-k}$ is connected by a sequence of edges in $\Gamma_k$. In this section, we will verify that $\Gamma_k$ is, indeed, a connected set containing $V_k$. Equivalently, an arbitrary pair of points $v, v' \in V_k$ can be connected by a tour inside $\Gamma_k$.

The proof runs by induction on $k$. First, it is clear that $\Gamma_0$ is connected, and indeed every pair $v, v' \in V_0$ is connected in $\Gamma_0$ by either an edge or a bridge. Now, suppose (inductively) that every pair of points in $V_{k-1}$ can be connected by a tour in $\Gamma_{k-1}$. Then, fix $v_1, v_p \in \Gamma_k$ (the index parameter $p \in \mathbb{N}$ will be explained soon), and let $w_1 := v_1'$. By the inductive hypothesis, the points $w_1$ and $w_p$ can be joined by a tour in $\Gamma_{k-1}$, namely

$$w_1, w_2, \ldots, w_p$$

where $w_1$ is connected to $w_{i+1}$ by either an edge or a bridge contained in $\Gamma_{k-1}$. Note that

$$|v_1 - w_1| < A2^{-k} \quad \text{and} \quad |v_p - w_p| < A2^{-k}$$

by condition $(V^\dagger)$. Now, suppose inductively that $1 \leq t \leq p - 1$, and there exists a vertex $v_t \in V_k$ such that $|v_t - w_t| < A2^{-k}$, and $v_1$ is connected to $v_t$ by a tour in $\Gamma_k$. This is clearly true for $t = 1$, and the whole proof is about showing the same for $t = p$. If $t = p - 1$, then $v_{t+1} = v_p$, and it satisfies $|v_{t+1} - w_{t+1}| < A2^{-k}$; if $t \leq p - 2$, simply let $v_{t+1} \in V_k$ be any point satisfying $|v_{t+1} - w_{t+1}| < A2^{-(k-1)}$ (which exists by $(V^\dagger)$). In both cases, it remains to demonstrate that $v_t$ and $v_{t+1}$ can be connected by a tour contained in $\Gamma_k$.

The proof divides into two cases, depending on whether $w_1$ is connected to $w_{t+1}$ by a bridge or an edge contained in $\Gamma_k$.

---

8By definition, a tour is a finite sequence of vertices $x_1, \ldots, x_p \in V_k$, where $x_i$ is connected to $x_{i+1}$ by either an edge or a bridge in $\Gamma_k$. Note that the bridge need not necessarily be of the form $B(x, x_{i+1})$; the requirement is simply that $B(x, y) \subset \Gamma_k$ and $x_i, x_{i+1} \in B(x, y)$. 

Case (Bridge). Assume that \( w_t \) is connected to \( w_{t+1} \) by a bridge contained in \( \Gamma_{k-1} \). This means that there are vertices \( x, y \in V_j \) for some \( 0 \leq j \leq k-1 \) such that \( w_t, w_{t+1} \in B(x, y) \). Let \( v_t' := w_t^j \in B(x, y) \) and \( v_{t+1}' := w_{t+1}^k \in B(x, y) \). It is not automatically clear that \( v_t' = v_t \) and \( v_{t+1}' = v_{t+1} \), but it turns out that \( |v_t - v_t'| \) and \( |v_{t+1} - v_{t+1}'| \) are so small that there are connecting edges. For example,

\[
|v_t - v_t'| \leq |v_t - w_t| + |w_t + v_t'| \leq 2A2^{-(k-1)} < 30A2^{-k},
\]

which by the Principle implies that \( v_t \) is connected to \( v_t' \) by an edge in \( \Gamma_k \). The same is true for the pair \( v_{t+1}, v_{t+1}' \). Since \( v_t' \) and \( v_{t+1}' \) are connected by the bridge \( B(x, y) \subset \Gamma_k \), it follows that \( v_t \) and \( v_{t+1} \) are connected by a tour in \( \Gamma_k \).

Case (Edge). Assume that \( w_t \) is connected to \( w_{t+1} \) by an edge in \( \Gamma_{k-1} \). In particular,

\[
|w_t - w_{t+1}| < 30C2^{-(k-1)} = 60C2^{-k}
\]

by the Principle, and consequently

\[
|v_t - v_{t+1}| \leq |v_t - w_t| + |w_t - w_{t+1}| + |w_{t+1} - v_{t+1}| < 2C2^{-k} + 60C2^{-k} + 2C^{-k} = 64C2^{-k}.
\]

This means that \( v_{t+1} \in \Gamma_k \cap \mathcal{N}(v_t) \) and vice versa. Consequently, if \( \alpha_{v_t} \geq \epsilon \), then either \( [v_t, v_{t+1}] \subset \Gamma_{v_t} \subset \Gamma_k \) or \( B(v_t, v_{t+1}) \subset \Gamma_{v_t} \subset \Gamma_k \) by Case (L), and we are done.

Now, assume that \( \alpha_{v_t} < \epsilon \), so \( \Gamma_{v_t} \subset \Gamma_k \) is constructed by one of sub-cases in Case (S). As before, the assumption \( \alpha_{v_t} < \epsilon \) means that the points in \( V_k \cap \mathcal{N}(v_t) \) can be ordered "along" the line \( \ell_{v_t} \), and the notions of "left" and "right" make sense. Assume that \( v_{t+1} \) is "to the right" of \( v_t \), and the points of \( V_k \cap \mathcal{N}(v_t) \) lying between \( v_t \) and \( v_{t+1} \) are indexed as

\[
v_t = z_1 < z_2 < \ldots < z_q = v_{t+1}.
\]

Using the fact that \( |v_t - v_{t+1}| < 64C2^{-k} \), and that \( \alpha_{v_t} \) is very small, it is easy to convince oneself that

\[
v_t, v_{t+1} \in B(z_i, 65A2^{-k}) = \mathcal{N}(z_i), \quad 1 \leq i \leq q.
\]

(I omit the proof, because it is so clear that this holds for some suitable choice of \( \epsilon \), even if it this were not exactly \( \epsilon = 1/32 \).) Consequently, if it happens that \( \alpha_{z_i} = \epsilon \) for even one index \( 1 \leq i \leq q \), then either \( [v_t, v_{t+1}] \subset \Gamma_{z_i} \subset \Gamma_k \) or \( B(v_t, v_{t+1}) \subset \Gamma_{z_i} \subset \Gamma_k \) by Case (L), and we are happy.

So, the remaining case is where \( \alpha_{z_i} < \epsilon \) for all \( 1 \leq i \leq q \). The plan is to show that \( z_i \) is connected to \( z_{i+1} \), for \( 1 \leq i \leq q-1 \), by either an edge or a bridge contained in \( \Gamma_{z_i} \subset \Gamma_k \); it will then follow that \( v_t \) is connected to \( v_{t+1} \) by a tour in \( \Gamma_k \), as claimed. If \( |z_i - z_{i+1}| < 30A2^{-k} \), then we are in Case (S-NT) and \( [z_t, z_{t+1}] \subset \Gamma_{z_t} \subset \Gamma_k \).

![Figure 9. The final case in the proof of connectedness.](image-url)
Otherwise, $|z_i - z_{i+1}| \geq 30A2^{-k}$, and we are in either Case (S-TT) or (S-TB). In Case (S-TB), the bridge $B(z_i, z_{i+1})$ is contained in $\Gamma_i$, and we are happy. So, it suffices to rule out Case (S-TT), where potentially $\Gamma_{i+1}^k = \{z_i\}$. This uses the fact that $[w_t, w_{t+1}]$ is an edge in $\Gamma_{k-1}$, so in particular $|w_t - w_{t+1}| < 30A2^{-(k-1)}$. Recall the point $w_{t+1}^\pi \in B(z_i, 2A2^{-k})$ from the Case (S-TT) associated with $z_i$; now this case could only occur, if the "next point of $\Gamma_{k-1}$ to the right" from $w_{t+1}^\pi$ was very far away (at distance $\geq 30A2^{-(k-1)}$) or did not exist at all. But since $z_i$ satisfies $v_t < z_i < v_{t+1}$, and

$$|z_i - v_{t+1}| \geq |z_i - z_{i+1}| \geq 30A2^{-k},$$

we can infer that $w_{t+1} \in \Gamma_{k-1} \cap \mathcal{N}(z_i)$ lies (strictly) to the right from $w_{t+1}^\pi$, and satisfies $|w_{t+1}^\pi - w_{t+1}| \leq |w_t - w_{t+1}| < 30A2^{-(k-1)}$. Hence also the "next point of $\Gamma_{k-1}$ to the right from $w_{t+1}^\pi$" is at distance $< 30A2^{-(k-1)}$, and Case (S-TT) cannot occur at $z_i$. The proof of connectedness is complete.

9. LENGTH ESTIMATES

Now we really arrive at the core of the proof. The aim will be to prove that

$$\mathcal{H}^1(\Gamma_k) \lesssim 1 + \sum_{j \leq k} \sum_{v \in V_j} \alpha_v^2 2^{-j}. \quad (9.1)$$

(All we normalised $r_0 = 1$, then $r_0$ would appear above in place of 1). The obvious first attempt would be to estimate

$$\mathcal{H}^1(\Gamma_k) \leq \mathcal{H}^1(\Gamma_{k-1}) + C \sum_{v \in V_k} \alpha_v^2 2^{-k},$$

since this estimate could be iterated $k$ times to produce (9.1). Before getting down on the actual details, I briefly discuss why this should work, and why it actually does not quite work. Let $v \in V_k$. The basic (and slightly naive) idea is estimate

$$\mathcal{H}^1(\Gamma_v) \leq \mathcal{H}^1(\Gamma_{k-1} \cap \mathcal{N}(v)) + C\alpha_v^2 2^{-k}. \quad (9.2)$$

If $\alpha_v \geq \epsilon$, this is trivially true (with implicit constants depending on $\epsilon$ of course), since it is easy to check that $\mathcal{H}^1(\Gamma_v) \lesssim 2^{-k}$. If $\alpha_v < \epsilon$, the situation might look like the one in Figure 10. In particular, the vertices of $\Gamma_k$ are ordered linearly relative to $\ell_v$. Now, if cheating and over-optimism are allowed for a moment, the length of all $\Gamma_k$ inside $\mathcal{N}(v)$ can be estimated as follows. First, for every edge $[v', v''] \subset \Gamma_k$ with $v', v'' \in V_k \cap \mathcal{N}(v)$, use Lemma 7.6 to infer that

$$\mathcal{H}^1([v', v'']) \leq \mathcal{H}^1([\pi_{\ell_v} (v'), \pi_{\ell_v} (v'')]) + C\alpha_v^2 2^{-k}.$$
simply because $\pi_{\ell_v}$ is a 1-Lipschitz mapping and does not increase length. This is precisely the estimate we were after. Of course, in real life there may be bridges contained in $\Gamma_v$, and there will be problems near the boundary of $\mathcal{N}(v)$ (where edges are no longer spanned by two vertices in $V_k \cap \mathcal{N}(v)$). These problems are reason for serious headache, but they are little compared to the following big issue: even if $\Gamma_v$ has non-trivial length, the set $\Gamma_{k-1} \cap \mathcal{N}(v)$ can be absolutely tiny, in fact a point! This can happen, for instance, if $v$ is a vertex on a bridge and has many neighbours in $V_k \cap \mathcal{N}(v)$, but $w = v^\uparrow$ has no neighbours in $V_{k-1} \cap \mathcal{N}(v)$, see Figure 11.

---

**Figure 10.** The set $\Gamma_k$ is drawn in black, and the part of $\Gamma_{k-1}$ inside $\mathcal{N}(v)$ is drawn in green. The line $\ell_v$, which now coincidentally happens to pass through $v$, is drawn in blue.

**Figure 11.** The vertex $v$ belongs to the same bridge as $w = v^\uparrow \in V_{k-1}$ and $w^\downarrow \in V_{k-2}$. The vertex $w$ is the only vertex of $V_{k-1}$ inside $\mathcal{N}(v)$; in particular, $w = (v')^\uparrow$ for all the vertices $v' \in V_k \cap \mathcal{N}(v)$, which forces these vertices to lie quite close to $w$ – and each other. Thus, these vertices are inter-connected by edges in $\Gamma_k$. 
Clearly, there is no chance for anything like (9.2) to work in this scenario. This is the main problem in the proof, and it is resolved with a very clever "pre-payment" scheme. This is formalised through the notion of virtual credit, which we now start to discuss.

9.1. Virtual credit. The virtual credit associated with \( v \in V_k \) is the number

\[
\$v = \$v,k = \$k := 3A2^{-k}.
\]

The virtual credit of a bridge \( B(v, v') \), \( v, v' \in V_k \) is

\[
\$v,v' := \sum_{w \in B(v, v')} \$w = 2 \cdot 3A \sum_{j \geq k} 2^{-k} = 12^{-k}.
\]

Virtual credit is an allegory of life itself: it comes and goes, not everyone has it all the time, and everyone loses it in the end. For every \( k \geq 0 \), we will inductively define a set

\[
R_k \subset \bigcup_{\ell \geq k} V_\ell,
\]

whose elements are "rich", and have credit at time \( k \). Other elements are poor and have nothing. The total virtual credit at time \( k \) equals

\[
\$(R_k) := \sum_{v \in R_k} \$v.
\]

The initial set \( R_0 \) of rich vertices consists of all the elements of \( \bigcup_{k \geq 0} V_k \), which make an appearance in the definition of \( \Gamma_0 \). In other words, every \( v \in V_0 \) lies in \( R_0 \), and also \( B(v, v') \subset R_0 \) for all \( v, v' \in V_0 \). To get the induction rolling, I state two key properties, which will always be required from \( R_k \):

(BP) The "bridge property" states that whenever \( B(v, v') \subset \Gamma_k \) with \( v, v' \in V_j \) and \( 0 \leq j \leq k \), then \( R_k \) contains all vertices in \( B(v, v') \cap \bigcup_{l \geq k+1} V_l \).

(TVP) The "terminal vertex property" is a bit more complicated, but it essentially states that "those without neighbours are rich". To be precise, fix \( v \in V_k \), and let \( \ell \) be any line such that

\[
\text{dist}(y, \ell) < \epsilon \cdot 2^{-k}, \quad y \in V_k \cap B(v, 30.42^{-k}).
\]

Then the points in \( V_k \cap BN(v, 30.42^{-k}) \) are arranged so that "left" and "right" make sense. In case there is no vertex of \( V_k \) either to the left or to the right of \( v \) inside \( B(v, 30.42^{-k}) \), then \( v \in R_k \).

These properties are trivially satisfies by \( R_0 \). So, next we assume that \( R_{k-1} \) has already been defined for some \( k \geq 1 \), satisfying the properties (BP) and (TVP). Let us see, how to define \( R_k \). First, initialise \( R_k \) by setting

\[
R_k := R_{k-1} \setminus [V_{k-1} \cup V_k].
\]

Next, we consider each element \( v \in V_k \) and add vertices to \( R_k \) according to the familiar cases (L), (S) etc.

(L) If \( \alpha_v \geq \epsilon \), then all vertices \( v' \in V_k \cap N(v) \) are added to \( R_k \). Also, if \( B(v', v'') \subset \Gamma_{k,v} \), then \( B(v', v'') \subset R_k \).

(S-NT) If \( \alpha_v < \epsilon \), and both \( \Gamma^R_v \) and \( \Gamma^L_v \) were defined via Case (S-NT), then no vertices are added to \( R_k \). Thus, \( v \) is terminal to neither left nor right.

(S-TT) If \( \alpha_v < \epsilon \) and either \( \Gamma^R_v \) or \( \Gamma^L_v \) was defined via Case (S-TT), then \( v \) is added to \( R_k \).
(S-TB) If $\alpha_v < \epsilon$ and $\Gamma^R_v$ was defined via Case (S-TB), add $B(v, v_1)$ to $R_k$. Similarly, if $\Gamma^L_v$ was defined via Case (S-TB), then add $B(v_{-1}, v)$ to $R_k$.

This completes the definition of $R_k$, and it is clear that $R_k$ satisfies the bridge property (BP) (note that if $v \in R_{k-1} \cap V_{k+1}$, then also $v \in R_k \cap V_{k+1}$). It is also clear (by induction) that $R_k \subset \bigcup_{i\geq k} V_i$. It remains to verify that $R_k$ satisfies the terminal vertex property (TVP). So, fix $v \in V_k$ and let $\ell$ be any such line that

$$\text{dist}(y, \ell) < \epsilon \cdot 2^{-k}, \quad y \in V_k \cap B(v, 30A2^{-k}). \quad (9.3)$$

Assume there is no vertex either to the "left" or "right" of $v$ in the ordering of $V_k \cap B(v, 30A2^{-k})$ with respect to $\ell$. Since $\ell$ is a completely arbitrary line, the estimate (9.3) tells us nothing about $\ell_v$ or $\alpha_v$: in particular, it could happen that $\alpha_v \geq \epsilon$. But in this case $v \in R_k$ by item (L) above, so we are happy. So, assume $\alpha_v < \epsilon$. Then the set $V_k \cap B(v, 30A2^{-k})$ is also ordered relative to $\ell_v$, and, by choosing orientations correctly, these orderings agree by Lemma 7.6. Thus, the fact that there is no vertex to the "left" or "right" from $v$ means that $v$ is either terminal to the left or right, and hence one of the items (S-TT) or (S-TB) occur. In both cases $v \in R_k$, and the induction is complete.

9.2. Proof of the length estimate (9.1). Now we are finally set to prove the estimate (9.1), which is repeated below:

$$\mathcal{H}^1(\Gamma_k) \leq 1 + \sum_{j \leq k} \sum_{v \in V_j} \alpha_v^2 2^{-j}. \quad (9.4)$$

The key auxiliary estimate is the following. Let $\text{Edges}(k)$ be the edges $[v, v'] \subset \Gamma_k$, and let $\text{Bridges}(k)$ be the bridges $B(v, v') \subset \Gamma_k$ with $v, v' \in V_k$ (note that $\Gamma_k$ may contain other bridges than those included in the "generation $k$ bridges" $\text{Bridges}(k)$). Then

$$\sum_{[v, v'] \in \text{Edges}(k)} \mathcal{H}^1([v, v']) + \sum_{B(v, v') \in \text{Bridges}(k)} \mathcal{H}^1(B(v, v')) + $(R_k)$ \leq \sum_{[w, w'] \in \text{Edges}(k-1)} \mathcal{H}^1([w, w']) + \frac{13}{15} \sum_{B(v, v') \in \text{Bridges}(k)} \mathcal{H}^1([v, v']) + $(R_{k-1}) + C \sum_{v \in V_k} \alpha_v^2 2^{-k}. \quad (9.5)$$

The reader might first think that there are typos on line (9.5), but there are none: the sum should not run over the bridges of generation $k - 1$, and we really want to sum over $\mathcal{H}^1([v, v'])$ instead of $\mathcal{H}^1(B(v, v'))$.

9.2.1. Proof of (9.4) based on (9.5). By definition of $\Gamma_k$,

$$\mathcal{H}^1(\Gamma_k) \leq \sum_{[v, v'] \in \text{Edges}(k)} + \sum_{j \leq k} \sum_{B(w, w') \in \text{Bridges}(j)} \mathcal{H}^1(B(w, w')),$$
and hence (9.5) leads to

\[
\mathcal{H}^1(\Gamma_k) + \$ (R_k) \\
\leq \sum_{[w,w'] \in \text{Edges}(k-1)} \mathcal{H}^1([w,w']) + \sum_{j \leq k-1} \sum_{B(w,w') \in \text{Bridges}(j)} \mathcal{H}^1(B(w,w')) + \$ (R_{k-1}) \\
+ \frac{13}{15} \sum_{B(v,v') \in \text{Bridges}(k)} \mathcal{H}^1([v,v']) + C \sum_{v \in V_k} \alpha^2 2^{-k} \\
= \mathcal{H}^1(\Gamma_{k-1}) + \$ (R_{k-1}) + \frac{13}{15} \sum_{B(v,v') \in \text{Bridges}(k)} \mathcal{H}^1([v,v']) + C \sum_{v \in V_k} \alpha^2 2^{-k}.
\]

Now, performing the same estimate on \( \mathcal{H}^1(\Gamma_{k-1}) + \$ (R_{k-1}) \), and continuing in the same manner \( k \) times, leads to

\[
\mathcal{H}^1(\Gamma_k) \leq \mathcal{H}^1(\Gamma_0) + \$ (R_0) + \frac{13}{15} \sum_{j \leq k} \sum_{B(v,v') \in \text{Bridges}(j)} \mathcal{H}^1([v,v']) + C \sum_{j \leq k} \sum_{v \in V_j} \alpha^2 2^{-j}.
\] (9.6)

To conclude (9.4) from here, note that \([v,v'] \subset B(v,v') \subset \Gamma_k\) for all \( B(v,v') \in \text{Bridges}(j) \) and for all \( j \leq k \). Moreover, the sets \([v,v']\) arising this way are essentially disjoint, meaning that

\[
\mathcal{H}^1([v,v'] \cap [w',w'']) = 0.
\]

for distinct pairs \( v, v' \) and \( w, w' \). This follows immediately from the initial reduction we made in Proposition 7.4. Consequently,

\[
\frac{13}{15} \sum_{j \leq k} \sum_{B(v,v') \in \text{Bridges}(j)} \mathcal{H}^1([v,v']) \leq \frac{13}{15} \mathcal{H}^1(\Gamma_k).
\]

Now (9.4) follows from (9.6), combined with the nearly trivial estimate

\[
\mathcal{H}^1(\Gamma_0) + \$ (R_0) \lesssim 1.
\]

9.3. Proof of the estimate (9.5). I repeat the estimate below:

\[
\sum_{[v,v'] \in \text{Edges}(k)} \mathcal{H}^1([v,v']) + \sum_{B(v,v') \in \text{Bridges}(k)} \mathcal{H}^1(B(v,v')) + \$ (R_k) \\
\leq \sum_{[w,w'] \in \text{Edges}(k-1)} \mathcal{H}^1([w,w']) + \frac{13}{15} \sum_{B(v,v') \in \text{Bridges}(k)} \mathcal{H}^1([v,v']) \\
+ \$ (R_{k-1}) + C \sum_{v \in V_k} \alpha^2 2^{-k}.
\] (9.7)

Staring at the left hand side for a moment reveals that it can be split into "local" terms of the form

\[
\sigma(v) := \sum_{[v',v''] \in \text{Edges}(k,v)} \mathcal{H}^1([v',v'']) + \sum_{B(v',v'') \in \text{Bridges}(k,v)} \mathcal{H}^1(B(v',v'')) + \$ (R_k(v)), \quad v \in V_k,
\]

where \text{Edges}(k,v) and \text{Bridges}(k,v) stand for the bridges and edges added to \( \Gamma_v \), and \( R_k(v) \) is the part of \( R_k \) constructed with \( v \in V_k \) fixed (recall the definition of \( R_k \)). However, simply estimating the left hand side of (9.7) by a sum of the local terms \( \sigma(v) \) over \( v \in V_k \) is wasteful: for instance, each edge \([v',v''] \subset \Gamma_k \) only needs to be counted once,
even if it may (and most often will) be included in $\Gamma_v$, for several distinct vertices $v$. Of course, even a wasteful estimate can sometimes work, but here it is too rough. Namely, every term on the right hand side of (9.7) can also be used only once to "pay" for something on the left hand side, and this is essentially why terms on the left hand side should also be accounted for precisely once.

At a high level, proving the estimate (9.7) thus has two challenges: first, to estimate each term $\sigma(v)$ separately by something appearing on the right hand side of (9.7), and, second, to make sure that nothing on the right hand side gets used twice in such estimates, when $v \in V_k$ varies. Thus, the proof will (formally speaking) contain the construction of an injective mapping $\Psi$ from all the terms on the left hand side to those on the right hand side. I will never attempt to write the complete expression of $\Psi$ down, but this philosophy is good to keep in mind.

The proof now begins. The local terms $\sigma(v)$ with $\alpha_v \geq \epsilon$ allow for a very care-free estimate:

9.3.1. *Edges, bridges and virtual credit nearby a case (L) vertex.* Assume that $\alpha_v \geq \epsilon$. Then

$$\sigma(v) \lesssim 2^{-k} \lesssim \alpha_v 2^{-k},$$

which is certainly good enough for us. The other cases will be more involved, but we can already benefit from the fact that Case (L) has been settled: in case an edge, or bridge, or virtual credit appearing below also happens to appear in some local term $\sigma(v)$ with $\alpha_v \geq \epsilon$, then we know that this edge/bridge/virtual credit has already been accounted for, and can be ignored.

9.3.2. *Virtual credit and parts of edges very close to Case (S-TT) vertices.* Assume that $\alpha_v < \epsilon$, so that both $V_k \cap \mathcal{N}(v)$ and $V_{k-1} \cap \mathcal{N}(v)$ are ordered along the line $\ell = \ell_v$, and "left" and "right" make sense. In this subsection, we will not handle the full sum $\sigma(v)$ for a Case (S-TT) vertex $v \in V_k$, but only a part of it. The rest will come later. For the moment, we are indeed just interested in bounding the quantity

$$S_v + \sum_{[v',v''] \in \text{Edges}(k,v)} \mathcal{H}^1([v',v''] \cap B(v, 2A2^{-k}))$$

(9.8)

for a fixed vertex $v \in V_k$, for which either $\Gamma_v^L$ or $\Gamma_v^R$ was defined through Case (S-TT). Note that in this case $v \in R_k$, so $S_v$ is indeed a part of $\$(R_k)$ and appears on the left hand side of (9.7). The quantity in (9.8) will be bounded by either $s_w$ or $s_w + s_{w'}$ for certain $w, w' \in R_{k-1}$ (which are, in turn, terms appearing on the right hand side of (9.7)).

There are a few cases to consider. You probably need to recall what Case (S-TT) means: in particular, recall the definition of the left-most and right-most vertices

$$w_l \in V_{k-1} \cap B(v, 2A2^{-k}) \quad \text{and} \quad w_r \in V_{k-1} \cap B(v, 2A2^{-k}).$$

Observe that certainly

$$\text{dist}(y, \ell) < \epsilon \cdot 2^{-(k-1)}, \quad y \in V_{k-1} \cap B(w, 30A2^{-(k-1)}), \quad w \in \{w_l, w_r\},$$

(9.9)

because $B(w, 30A2^{-(k-1)}) \subset B(v, 65A2^{-k})$, and using the definition of $\alpha_v < \epsilon$. Now, assume that $\Gamma_v^R$, for instance, was defined through Case (S-TT). By definition, this means that there is no vertex of $V_{k-1}$ "to the right" from $w_r$ within $B(w, 30A2^{-(k-1)})$, and by the terminal vertex property (TVP), we conclude that $w_r \in R_{k-1}$. Now, once more using the fact that $\alpha_v < \epsilon$, recalling the length estimate in Lemma 7.6, and observing that there are
no edges passing in the "right half" of $B(v, 2A^2 - k)$ (this is precisely because $\Gamma^R_v$ is defined via Case (S-TT), see Figure 12), one obtains

$$w + \sum_{[v',v''] \in \text{Edges}(k,v)} H^1([v',v''] \cap B(v, 2A^2 - k)) = 3A2^{-k} + (1+3\epsilon^2)2A2^{-k} < 3A2^{-(k-1)} = w_r,$$

Of course, the estimate above would hold even with $w_r \notin R_{k-1}$, but the point is that now $w_r$ is something appearing on the right hand side of (9.7). In case $\Gamma^L_v$ was defined via Case (S-TT), the estimate is the same, with $w_r$ replaced by $w_l$, and "left" and "right" interchanged.

Are we done? No! It might occur that a single vertex $w \in R_{k-1}$ is needed in the treatment of many distinct Case (S-TT) vertices $v$ as above (so the injectivity of the mapping $\Psi$ is in jeopardy). More precisely, there could be several vertices $v$ such that either

- $w = w^r_v$, where $\Gamma^R_v$ was defined via Case (S-TT), or
- $w = w^L_v$, where $\Gamma^L_v$ was defined via Case (S-TT).

I emphasise again that this is a real problem, because $w_r$ may only be counted once in $\$(R_{k-1})$ on the right hand side of (9.7). It turns out that several can be exactly twice (as we will shortly see), but even so $w_r$ is not large enough alone. To start tackling this problem, first note that whenever $w$ arises as $w^r_v$ or $w^L_v$ for some such vertex $v$, then $w \in B(v, 2A2^{-k})$, so $v \in B(w, 2A2^{-k})$.

Now, assume that there are at least two Case (S-TT) vertices $v_1, v_2 \in V_k \cap B(w, 2A2^{-k})$, which give rise to the same $w$ in the argument above. Then, inside the ball $B(w, 40A2^{-k})$ (this is just some ball large enough to contain all the interesting action) the vertices of $V_k$ are ordered with respect to $\ell = \ell_{v_1}$, say, and "left" and "right" make sense. Assume that $v_1$ is "left" from $v_2$: then $v_1$ is terminal to the left, and $v_2$ is terminal to the right. Since both are Case (S-TT) vertices, the conclusion is that $\Gamma^L_{v_1}$ must have been defined via Case (S-TT), and $\Gamma^R_{v_2}$ must have been defined via Case (S-TT). Now, we are in trouble, if $$w = w^L_{v_1} \quad \text{and} \quad w = w^R_{v_2}.$$ If this is really the case, then the fact that $\Gamma^L_{v_1}$ was defined via Case (S-TT) implies that there are no vertices of $V_{k-1}$ within distance $30A2^{-(k-1)}$ to the left from $w^L_{v_1} = w$. Similarly, there are no vertices of $V_{k-1}$ within distance $30A2^{-(k-1)}$ to the right from $w$. So, $w$

\[\text{Figure 12. The task is to estimate the length of the red curve.}\]
is in fact the only vertex in $V_{k-1} \cap B(w, 30A2^{-(k-1)})$, which forces
$$v_1^\uparrow = w = v_2^\uparrow.$$  
This gives the slightly improved estimate
$$|v_1 - v_2| \leq |v_1 - w| + |w - v_2| \leq 2A2^{-k}.$$  
In particular,
$$v \in \Gamma$$
and "right" make sense. In this subsection, we assume that at least one of
$$B$$
is only needed once
$$v \in V$$
and hence
$$w \in V$$
lie between
$$v$$
and
$$w.$$  
The final observation is that since $w$ is not connected by an edge to any other vertices in $V_{k-1}$ (all such vertices are too far away, at distance $30A2^{-(k-1)}$ at least), but since $w$ is still in some manner connected to other vertices in $\Gamma_{k-1}$, it must be the case that $w$ belongs to a bridge $B(x, y)$ for some $x, y \in V_j$, $j < k - 1$. Thus also $w^k \in B(x, y) \cap V_k$, and hence $w^k \in R_{k-1}$ by the bridge property (BP) of virtual credit. Since
$$S_{v_1} + S_{v_2} = S_{w^k} + \frac{S_w}{2},$$
we now arrive at the estimate
$$L + S_{v_1} + S_{v_2} \leq S_w + S_{w^k}.$$  
This means that even in the worst case, when $w$ is needed twice, every element of $R_{k-1}$ is only needed once in (this part of) the estimate (9.7)!

9.3.3. Virtual credit, edges and bridges near Case (S-TB) vertices. Suppose that $v \in V_k$ has $\alpha_v < \epsilon$, so that both $V_k \cap N(v)$ and $V_{k-1} \cap N(v)$ are ordered along the line $\ell = \ell_v$, and "left" and "right" make sense. In this subsection, we assume that at least one of $\Gamma^R_v$ and $\Gamma^L_v$ was defined via Case (S-TB). For instance, assume that $\Gamma^R_v$ was defined via Case (S-TB), so that $v_1, v_2$ the "next vertex to the right" from $v$ inside $N(v)$ exists, and lies at distance $\geq 30A2^{-k}$ from $v$, see Figure 13. This time we will handle the following part of $\sigma(v) + \sigma(v_1)$:
$$S_{v,v_1} + \sum_{[v', v'']} H^1([v', v''] \cap B\{v, v_1\}, 2A2^{-k}) + H^1(B(v, v_1)).$$  
Here $B\{v, v_1\}, 2A2^{-k}) := B(v, 2A2^{-k}) \cup B(v_1, 2A2^{-k})$. We will bound (9.10) by
$$H^1([w, w_{r+1}]) + \frac{13}{15}H^1([v, v_1]).$$
Note that both these terms above appear on the right hand side of (9.7), and we have not used them in the previous cases, so we are free to waste them here. Start by observing
Figure 13. The possible location of the edges in $\Gamma_k \cap B(\{v, v_1\}, 2A2^{-k})$ are marked in red, a part of the bridge in $\Gamma_k$ is marked in blue, and the edge $[w_r, w_{r+1}] \subset \Gamma_{k-1}$ is marked in green. The fact that $w_{r+1} \in B(v_1, 2A2^{-k})$ uses (7.10).

That

$$\mathcal{H}^1(B(v, v_1)) \leq \mathcal{H}^1([v, v_1]) + 2 \sum_{j=0}^{\infty} A2^{-j} = \mathcal{H}^1([v, v_1]) + 4A2^{-k},$$

by repeated application of the hypothesis $(V^1)$. Also, recall that $A_{v, v_1} = 12A2^{-k}$. The lengths of the edges inside $B(v, 2A2^{-k})$ can be estimated by

$$\sum_{[v', v''] \subseteq \text{Edges}(k)} \mathcal{H}^1([v', v''] \cap B(v, 2A2^{-k})) \leq (1 + 3\epsilon^2)2A2^{-k} < 3A2^{-k}$$

by Lemma 7.6, and the same holds for the edges inside $B(v_1, 2A2^{-k})$. All in all, (9.10) turns out to be at most

$$(4 + 12 + 6)A2^{-k} + \mathcal{H}^1([v, v_1]) = 22A2^{-k} + \mathcal{H}^1([v, v_1]).$$

Next, note that

$$\mathcal{H}^1([w_r, w_{r+1}]) \geq \mathcal{H}^1([v, v_1]) - 4A2^{-k},$$

because $w_r \in B(v, 2A2^{-k})$ and $v_1 \in B(w_{r+1}, 2A2^{-k})$ by (7.10). This proves that

$$(9.10) \leq \mathcal{H}^1([w_r, w_{r+1}]) + 26A2^{-k} \leq \mathcal{H}^1([w_r, w_{r+1}]) + \frac{13}{15} \mathcal{H}^1([v, v_1]),$$

using the assumption $|v - v_1| \geq 30A2^{-k}$.

Again, it is a legitimate concern, whether this case is now really complete: could it, again, happen that a term of the form $\mathcal{H}^1([w_r, w_{r+1}])$ or $\frac{13}{15} \mathcal{H}^1([v, v_1])$ comes up several times, as one varies the point $v$? This does happen, indeed, but only twice: for $v$ itself, and then $v_1$, and this is precisely why these two terms received a symmetrical treatment above. You should now think, how the sum (9.10) had looked like, had we started off with $v_1$ instead of $v$. If $\alpha_{v_1} < \epsilon$, then you will find that $\Gamma_{v_1}^k$ is defined via Case (S-TB), and $v$ is the first vertex to the left from $v_1$. Consequently, (9.10) looks the same for $v$ and $v_1$. But since these terms of (9.10) only need to be counted once in the sum (9.7), this is ok.

Now, we prove that the terms $\mathcal{H}^1([w_r, w_{r+1}])$ or $\frac{13}{15} \mathcal{H}^1([v, v_1])$ can only arise from this case for $v$ or $v_1$. So, assume that $v' \neq v$ is another vertex with this property, and let us prove that $v' = v_1$. Since $v'$ is relevant for this case, $v'$ must have either a left or a right
neighbour \( v'_{-1} \) or \( v'_1 \) such that \( B(v'_{-1}, v') \subset \Gamma_k \) or \( B(v'_1, v'_1) \subset \Gamma_k \). Assume, say, that \( v' \) has such a left neighbour \( v'_{-1} \).

Now, the proof above applied to the pair \( v', v'_{-1} \) gives rise to the terms

\[
\frac{13}{15} \mathcal{H}^1([v'_{-1}, v']) \quad \text{and} \quad \mathcal{H}^1([w'_{-1}, w'_1]),
\]

which are needed in estimating the analogue of (9.10). We potentially run into trouble, if either \([v'_{-1}, v'] = [v, v_1]\) or \([w'_{-1}, w'_1] = [w_r, w_{r+1}]\), because the terms \( \mathcal{H}^1([v, v_1]) \) and \( \mathcal{H}^1([w_r, w_{r+1}]) \) are needed (also) in connection with \( v \). In case \([v'_{-1}, v'] = [v, v_1]\), then \( v' = v_1 \), as claimed.

What if \([w'_{-1}, w'_1] = [w_r, w_{r+1}]\), so that \( w'_{-1} = w_r =: w_1 \) and \( w'_1 = w_{r+1} =: w_2 \)? Then

\[
v, v'_{-1} \in B(w_1, 2A2^{-k}) \quad \text{and} \quad v_1, v' \in B(w_2, 2A2^{-k}),
\]

But now all the points \( v, v', v'_{-1}, v_1 \) are linearly ordered with respect to \( \ell_v \), say, and this easily implies \( v_1 = v' \). Otherwise \( v_1 \) would either be strictly to the left or right from \( v' \). If left, then \( v'_{-1} \) certainly would not be the nearest point left from \( v' \). If right, then \( v_1 \) certainly would not be the nearest point right from \( v \). This proves that \( v_1 = v' \).

We soon move to the last case: it will be crucial to keep in mind that certain edges \( [w, w'] \in \text{Edges}(k-1) \) have already been used in the current case. So, the reader needs to make sure that those edges do not get used again!

### 9.3.4. Whatever remains

What actually remains? It is probably a good idea to have a look at (9.7) once more:

\[
\sum_{[v,v'] \in \text{Edges}(k)} \mathcal{H}^1([v, v']) + \sum_{B(v, v') \in \text{Bridges}(k)} \mathcal{H}^1(B(v, v')) + \$ (R_k)
\leq \sum_{[w,w'] \in \text{Edges}(k-1)} \mathcal{H}^1([w, w']) + \frac{13}{15} \sum_{B(v, v') \in \text{Bridges}(k)} \mathcal{H}^1([v, v']) + \$ (R_{k-1}) + C \sum_{v \in V_k} \alpha_v 2^{-k}.
\]

It is clear that \( \mathcal{H}^1(B(v, v')) \) has been dealt with for every bridge \( B(v, v') \in \text{Bridges}(k) \) by the previous case. It is not true that every element of \( R_k \) has been taken care of: we have only accounted for the new additions to \( R_k \) at step \( k \) (let us denote them by \( N_k \)), and a quick look at Section 9.3.2 reveals that we have done so by using exclusively the virtual credit in \( R_{k-1} \cap [V_{k-1} \cup V_k] \). But

\[
R_k \setminus N_k \subset R_{k-1} \setminus [V_{k-1} \cup V_k],
\]

because \( R_k \) was initialised by deleting everything from \( R_{k-1} \cap [V_{k-1} \cup V_k] \), so

\[
\$ (R_k \setminus N_k) \leq \$ (R_{k-1} \setminus [V_{k-1} \cup V_k]).
\]

This implies that \( \$ (R_k) \) is now completely accounted for. Consequently, all that remains "to be paid for" on the left hand side of (9.11) are certain edges, and parts thereof. More precisely, in Sections 9.3.1–9.3.3, we have already taken care of edges, and parts thereof, which are contained in either \( \mathcal{N}(v) \) for some \( v \in V_k \) with \( \alpha_v \geq \epsilon \), or alternately \( B(v, 2A2^{-k}) \) for some vertex \( v \in V_k \), which is terminal to either left or right.
Now, fix an edge \([v, v'] \subset \Gamma_k\) with \(v, v' \in V_k\) and such that \(\max\{\alpha_v, \alpha_{v'}\} < \epsilon\). Let \([u, u'] \subset [v, v']\) be (any) maximal sub-segment, which stays at distance \(\geq 2A2^{-k}\) from all terminal vertices. Let \(\pi\) be the orthogonal projection to the line \(\ell_v\). By Lemma 7.6,

\[
\mathcal{H}^1([u, u']) \leq (1 + 3\alpha_v^2)\mathcal{H}^1([\pi(u), \pi(u')]) \leq \mathcal{H}^1([\pi(u), \pi(u')]) + 90\alpha_v^2 2^{-k},
\]

since \(\pi\) is 1-Lipschitz and \(|u - u'| \leq 30A2^{-k}\) by the Principle. Assume, say, that \(u\) lies to the left form \(u'\) in the order relative to \(\ell_v\) (this makes sense, as \(\alpha_v < \epsilon\)). Then, let \(u_l \in V_k\) (resp. \(u'_l \in V_k\)) be the closest vertex to the left from \(u\) (resp. right from \(u'\)) with

\[
\pi(u_l) < \pi(u) - A2^{-k} \quad \text{and} \quad \pi(u'_l) > \pi(u) + A2^{-k}.
\]

Why do they exist? If, for instance, all vertices to the left from \(u\) within \(\mathcal{N}(v)\) satisfied the opposite inequality, then all of them would certainly lie in \(B(u, 2A2^{-k})\), and the leftmost of them would be terminal to the left, contrary to the definition of \(u\). The situation is depicted in Figure 14. Finally, use hypothesis \((V^\uparrow)\) to find vertices \(w_l = u_l^\uparrow\) and \(w'_l = (u'_l)^\uparrow\) with \(|w_l - u_l| < A2^{-k}\) and \(|w'_l - u'_l| < A2^{-k}\). It follows that

\[
\pi(w_l) < \pi(u) < \pi(u') < \pi(w'_l).
\]

(9.12)

Moreover, the vertices \(w_l\) and \(w'_r\) can be connected by a finite sequence of edges inside \(\Gamma_{k-1} \cap \mathcal{N}(v)\), which essentially follows from the Principle: \(w_l\) is fairly close to \(v\) (at distance \(\leq 32.42^{-k} \leq 30A2^{-k-1}\)), so \(w_l\) and \(v^\uparrow\) can first be connected by edges in \(\Gamma_{k-1}\) inside \(\mathcal{N}(v)\). Then, \(v^\uparrow\) can be connected to \(w'_r\), since \(w'_r\) is not much further from \(v^\uparrow\) than \(|v - v'| \leq 30A2^{-k} \leq 30A2^{-k-1}\).

Now, it follows from (9.12) and the discussion above that for every point \(t \in [\pi(u), \pi(u')]\), there is a point \(x(t) \in \Gamma_{k-1}\), belonging to one of the edges connecting \(w_l\) to \(w'_r\), such that \(\pi(x(t)) = t\) (see Figure 14). Moreover, this point \(x\) can be chosen so that

\[
|x(t) - t| \leq \alpha_v 2^{-k}.
\]

(9.13)

(Indeed, since the end-points of all edges \([w, w']\) fully contained \(\Gamma_{k-1} \cap \mathcal{N}(v)\) satisfy \(\text{dist}(w, \ell_v) < \alpha_v 2^{-k}\), the same distance bound remains true for any points on \([w, w']\).)

Consequently, using again the fact that \(\pi\) is 1-Lipschitz,

\[
\mathcal{H}^1([\pi(u), \pi(u')]) \leq \mathcal{H}^1(\mathcal{E}(u, u')),
\]

where \(\mathcal{E}(u, u') := \{(x(t) \in \Gamma_{k-1} : t \in [u, u']\}\). All in all,

\[
\mathcal{H}^1([u, u']) \leq \mathcal{H}^1(\mathcal{E}(u, u')) + 90\alpha_v^2 2^{-k},
\]

(9.14)
where $\mathcal{H}^1(\mathcal{E}(u, u'))$ is certainly a part of the sum

$$
\sum_{[w, w'] \in \text{Edges}(k-1)} \mathcal{H}^1([w, w']).
$$

It is also easy to see that those edges $[w, w']$ do appear in this manner, which arose (and whose length was already used) in the previous section. Any such edge $[w, w']$ had the property that both $w$ and $w'$ lay at distance $\leq 2A2^{-k}$ from Case (S-TB) vertices in $V_k$, and there are none of those close enough.

So, the only remaining problem is that the sets $\mathcal{E}(u, u')$ can have some overlap as $u, u'$ vary. Assume that two sets of the form $\mathcal{E}_1 := \mathcal{E}(u_1, u'_1)$ and $\mathcal{E}_2 := \mathcal{E}(u_2, u'_2)$ meet at a point $\xi \in \Gamma_{k-1}$, where $[u_1, u'_1]$ and $[u_2, u'_2]$ are distinct segments. The first task is to show that "all the action happens in a single local picture $\mathcal{N}(\nu)$". Here are some basic facts: $[u_1, u'_1] \subset [v_1, v'_1]$ and $[u_2, u'_2] \subset [v_2, v'_2]$ for certain vertices $v_1, v_2, v'_1, v'_2$ with $\alpha$-numbers at most $\epsilon$, satisfying

$$
|v_1 - v'_1| < 30A2^{-k} \quad \text{and} \quad |v_2 - v'_2| < 30A2^{-k}.
$$

Moreover, it is easy to check that $\mathcal{E}(u_1, u'_1) \subset B(v_1, 32A2^{-k}) \cap B(v'_1, 32A2^{-k})$ and similarly $\mathcal{E}(u_2, u'_2) \subset B(v_2, 32A2^{-k}) \cap B(v'_2, 32A2^{-k})$. In particular, the point $\xi$ lies in all of the balls above. Now, it does not make a big difference, which vertex $v_i$ or $v'_i$ we declare as our "centre" $v$; say $v := v_1$. Then $\alpha_v < \epsilon$, and all the sets above lie well inside $\mathcal{N}(v)$ (check this; or if you are lazy, just assume that the constant "65" in the definition of $\mathcal{N}(v)$ is replaced by $10^{10}$ – it's precise value is totally irrelevant in future applications).

The points $v_i$ and $v'_i$, $i \in \{1, 2\}$ are now linearly ordered relative to $\ell = \ell_v$. It cannot happen that $v_1 = v'_1$ and $v_2 = v'_2$, because the sets $\mathcal{E}(u, u')$ arising from a single edge $[v_1, v'_1]$ are clearly disjoint (all those sets are defined the fixed projection $\pi = \pi_v$, so the points $x(t) \in \Gamma_{k-1}$ are distinct for disjoint intervals $[u_1, u'_1], [u_2, u'_2] \subset [v_1, v'_1]$).

Now there are essentially two different possibilities: either all the points $v_i, v'_i$ are distinct, or then, say $v'_1 = v_2$ and the three points $v_1, v_2, v'_2$ are consecutive in the linear order relative to $\ell$. The first situation actually cannot occur, if $\xi$ exists: this follows from the separation $|v'_1 - v_2| \geq 2^{-k}$ and the fact that $\epsilon$ is so small, which implies that the projections $\pi_{v_1}$ and $\pi_{v_2}$ are nearly the same (as will be discussed carefully below). Checking the details is a bit tedious, but the situation is shown in Figure 15.

![Figure 15](image-url)

**Figure 15.** This is the case, where $v_1 < v'_1 < v_2 < v'_2$. The edges in $\Gamma_k$ are drawn in red, and the edges in $\Gamma_{k-1}$ are shown in green. The parts of $\Gamma_{k-1}$ required to pay for $[v_1, v'_1]$ (or anything in between) are separated by the dotted lines. As you can see, these parts are distinct, because the separation between $|v'_1 - v_2| \geq 2^{-k}$ is large compared to $\alpha_v 2^{-k}$. 

Now, we are left with the case $v_1' = v_2$. I rename the three vertices $v_1, v_1', v_2$ as $v_1, v_2, v_3$, and the assumption is that $v_1 < v_2 < v_3$ are consecutive vertices in the linear order relative to $\ell_{v_1}$ (or any other $\ell_{v_i}$, because these orders are compatible due to small $\alpha$-numbers). This situation is depicted in Figure 16. This might be clear to the reader, but let us briefly repeat: what causes the overlap? The segments $[u_1, u_1']$ and $[u_2, u_2']$ are contained in the two distinct, consecutive segments $[v_1, v_2]$ and $[v_2, v_3]$. The edges in $\Gamma_k$ are drawn in deep red, and the edges in $\Gamma_{k-1}$ are shown in green. The overlap $\mathcal{E}(v_1, v_2) \cap \mathcal{E}(v_2, v_3)$ is show in bright red.

![Figure 16](image)

**Figure 16.** This is the case, where $v_1 < v_2 < v_3$, and moreover $[u_1, u_1'] = [v_1, v_2]$ and $[u_2, u_2'] = [v_2, v_3]$. The edges in $\Gamma_k$ are drawn in deep red, and the edges in $\Gamma_{k-1}$ are shown in green. The overlap $\mathcal{E}(v_1, v_2) \cap \mathcal{E}(v_2, v_3)$ is show in bright red.

It now suffices to show that (9.14) (see Figure 17). formed by the three points $\Delta_1 := \ell_{v_1} \cap \ell_{v_2}, \Delta_2 = \pi_{v_1}(v_3)$ and $\Delta_3 = \pi_{v_2}(v_3)$). Then the angle at $\Delta_1$ is obviously $\theta$, and so the sine of $\theta$ is

$$\sin \theta = \frac{|\Delta_2 - \Delta_3|}{|\Delta_2 - \Delta_1|} \lesssim \frac{\alpha \cdot 2^{-k}}{2^{-k}} = \alpha.$$ 

Hence $\theta \lesssim \alpha$, and the proof of Theorem 7.1 is complete.
Figure 17. The overlap $E(v_1, v_2) \cap E(v_2, v_3)$ is shown in bright red. It’s total length is clearly bounded by $\lesssim \theta \cdot \alpha 2^{-k}$. The triangle, from which $\theta$ can be solved, is marked by the three blue discs: the intersection of $\ell_{v_1}$ and $\ell_{v_2}$, the projection $\pi_{v_1}(v_3)$, and the projection $\pi_{v_2}(\pi_{v_1}(v_3))$.

References

University of Helsinki, Dept. of Mathematics and Statistics
E-mail address: tuomas.orponen@helsinki.fi