The Tb-theorem on non-homogeneous spaces that proves a conjecture of Vitushkin

F. Nazarov, S. Treil and A. Volberg*

Contents

F. Nazarov, S. Treil, A. Volberg

Abstract

This article was written in 1999, and was posted as a preprint in CRM (Barcelona) preprint series n^0 519 in 2000. However, recently CRM (Barcelona) erased all preprints dated before 2006 from its site, and this paper became inacessible. It has certain importance though, as the reader shall see. Meanwhile this paper in bits and pieces appeared in several book formats, namely in Volberg's lecture notes [Vo], in Doudziak's book [Du], and in Tolsa's book [To].

Formally this paper is a proof of the (qualitative version of the) Vitushkin conjecture. The last section is concerned with the quantitative version. This quantitative version turns out to be very important. It allowed Xavier Tolsa to close the subject concerning Vtushkin's conjectures: namely, using the quantitative nonhomogeneous Tb theorem proved in the present paper, he proved the semiadditivity of analytic capacity. Another "theorem", which is implicitly contained in this paper, is the statement that any non-vanishing L^2-function is accretive in the sense that if one has a finite measure μ on the complex plane $\mathbb{C}$ that is Ahlfors at almost every point (i.e. for μ-almost every $x \in \mathbb{C}$ there exists a constant $M > 0$ such that $\mu(B(x,r)) \leqslant Mr$ for every $r > 0$) then any one-dimensional antisymmetric Calderón-Zygmund operator K (i.e. a Cauchy integral type operator) satisfies the following "all-or-nothing" princple: if there exists at least one function $\varphi \in L^2(\mu)$ such that $\varphi(x) \neq 0$ for μ-almost every $x \in \mathbb{C}$ and such that *the maximal singular operator* $K^*\varphi \in L^2(\mu)$, then there exists an everywhere positive weight $w(x)$, such that K acts from $L^2(\mu)$ to $L^2(wd\mu)$. In particular, there exists a a set E of positive μ-measure, $\mu(E) > 0$, such that operator K is a bounded operator from $L^2(E,\mu)$ to itself. Moereover, a concrete estimate can be given for the bound of its norm and the portion $\mu(E)/\|\mu\|$ if we have quantitative information on how non-zero is φ and haow small is $\|K^*\varphi\|$.

*All authors are partially supported by the NSF grant DMS 9970395

Table of contents

0. What this is all about

Let us be a little bit more specific. The analytic capacity of a compact set on the plane was defined by Ahlfors in 1947 as

$$\gamma(E) = \sup_f \lim_{z\to\infty} |z\, f(z)|,$$

where the supremum is taken over all analytic functions in the complement of E such that $|f(z)| \le 1$ and $f(\infty) = 0$. Ahlfors showed that $\gamma(E) = 0$ if and only if E is removable for bounded analytic functions. It was very interesting to find a geometric characterization. This is often called the Painlevé problem since Painlevé started to study it more than 100 years ago.

Vituskin's conjecture (1967): for sets E such that $\mathcal{H}^1(E) < \infty$, $\gamma(E) = 0$ if and only if $\mathcal{H}^1(E \cap \Gamma) = 0$ for every rectifiable curve Γ.

Alberto Calderón and Guy David found the geometric characterization of sets of positive analytic capacity and finite length (= finite $\mathcal{H}^1$-measure), thus proving one half of Vitushkin's conjecture each.

Theorem: Let E be a compact on the plane with $\mathcal{H}^1(E) < \infty$. Then $\gamma(E) = 0$ if and only if $\mathcal{H}^1(E \cap \Gamma) = 0$ for every rectifiable curve Γ.

Here $\mathcal{H}^1$ is 1-dimensional Hausdorff measure. The sets of finite 1-dimensional Hausdorff measure with the latter condition satisfied are called purely unrectifiable according to Federer. Besicovitch studied them and multidimensional analogs in the 1920's and 1930's and proved many very difficult and beautiful results about such sets. He called them irregular.

The "only if" part of the theorem has been proved by Calderón in 1977. It amounts to establishing that the Cauchy integral operator on Lipschitz curves is bounded on L^2 (Calderón's problem, which he solved in 1977 for small Lipschitz constants: this turned out to be sufficient for the "only if" part). The "if" part was considered to be super difficult. Finally it was proved by Guy David in 1997 [D1] using also [DM]. But actually this was only the "analytic part" of the proof. The "geometric part" was fortunately known because of the fantastic idea of Melnikov and Verdera [MV] and a geometric theorem due to David and Léger [L].

Here we give another (probably simpler and more streamlined, more conceptual) proof of the "if" part in the theorem, actually of the "analytic" part.

To explain the approach we need the notion of the Cauchy integral operator. So let E in the plane have finite $\mathcal{H}^1(E)$. Call $\mu = \mathcal{H}^1|E$. The Cauchy singular integral operator C_μ is

$$C_\mu g(z) = \lim_{\delta\to 0} \int_{E\setminus B(z,\delta)} \frac{g(\zeta)}{\zeta - z}\, d\mu(\zeta)\,.$$

Actually, if $z \in E$, it is not clear when the limit exists (while outsude of E the definition is always fine). So we introduce the maximal Cauchy singular integral operator C^*_μ:

$$C^*_\mu g(z) = \sup_{\delta>0} \Big| \int_{E\setminus B(z,\delta)} \frac{g(\zeta)}{\zeta - z}\, d\mu(\zeta) \Big|$$

and the "cut-off" Cauchy singular integral operator C^δ_μ:

$$C^\delta_\mu g(z) = \int_{E\setminus B(z,\delta)} \frac{g(\zeta)}{\zeta - z}\, d\mu(\zeta)\,.$$

Suppose $\gamma(E) > 0$. One should find a rectifiable Γ such that $H^1(E \cap \Gamma) > 0$.

The analytic part here will end by constructing a positive (this is very important, let us say this again, positive) ϕ such that

$$|C^*_\mu \phi(z)| \le 1 \quad \forall z \in \mathbb{C}\,.$$

Setting $\nu = \phi\, d\mu$ and applying to this positive measure the permutation idea from [MV] one gets

$$c^2(\nu) := \int\int\int c(x,y,z)^2 d\nu(x)d\nu(y)d\nu(z) < \infty$$

where $c(x,y,z)$ is the reciprocal of the radius of the circle passing through x, y, z. The quantity $c(\nu)$ is called the Menger curvature of the measure ν.

The following theorem is from the abovementioned "geometric part" of the proof. It is due to David and Léger [L].

Theorem: If $\nu = \phi\, dH^1|E, \phi \geqslant 0, \phi \in L^\infty(E)$, $\mathcal{H}^1(E) < \infty$ and $c^2(\nu) < \infty$, then there are rectifiable curves Γ_i such that $\nu(\mathbb{C} \setminus \cup_{i=1}^{\infty}\Gamma_i) = 0$.

Now we see that after constructing a **positive** ϕ such that $|C^*_\mu \phi(z)| \leq 1, \quad \forall z \in \mathbb{C}$, one refers to the geometric papers [MV] and [L] to finish the proof of Vitushkin's conjecture.

How to find such a positive ϕ? We have only the information that $\gamma(E) > 0$ and $\mathcal{H}^1(E) < \infty$. The first condition means that there is a nonconstant bounded analytic function f in $\mathbb{C} \setminus E$ vanishing at infinity. The second condition quite easily shows that this f is represented as a Cauchy integral of $\phi\, d\, H^1|E = \phi\, d\, \mu$: $f(z) = C_\mu \phi(z)$, $\forall z \in \mathbb{C} \setminus E$. We do not explain this now, but it is easy to assume that our $\mu := \mathcal{H}^1|E$ satisfies $\mu(B(z,r)) \leq C\, r$ for all $z \in \mathbb{C}$ and all $r > 0$. Then not only $C_\mu \phi(z)$ is bounded on $\mathbb{C} \setminus E$, but one can prove that there exists a finite constant C such that

$$|C^*_\mu \phi(z)| \leq C < \infty \qquad \forall z \in \mathbb{C}\,.$$

But this is not at all what we need—even though it seems precisely what we wanted. The main problem is that ϕ is complex valued function! It is impossible to prove that it is positive. (Actually positivity will generically never happen.)

Here is the main result to the proof of which the rest of the paper is devoted:

Main Theorem: Let μ denote $\mathcal{H}^1|E$ for a set E of finite 1-dimensional Hausdorff measure. If there is a nonzero $\phi \in L^\infty(E)$ (this L^∞ part can be weakened) such that $\sup_{z\in\mathbb{C}} |C^*_\mu \phi(z)| \leq Const < \infty$, then there exists a nonnegative bounded function ψ, which is strictly positive on the set of positive measure μ, such that $\sup_{z\in\mathbb{C}} |C^*_\mu \psi(z)| \leq Const < \infty$.

Actually the fact that we work with $\mathcal{H}^1$ is not important. Another way of expressing the essence of the previous theorem is to formulate its analog, which is as follows:

Theorem (on bounded Cauchy transforms of measures): Let ν denote a nonzero complex measure with compact support on the plane. Let its Cauchy transform C_ν be uniformly bounded: $\sup_{z\in\mathbb{C}\setminus\mathrm{supp}(\nu)} |C_\nu(z)| \leq Const < \infty$. Suppose that the area of $\mathrm{supp}\,\nu$ is zero. Then there exists a positive measure μ, absolutely continuous with respect to ν, such that its Cauchy transform is uniformly bounded too: $\sup_{z\in\mathbb{C}\setminus\mathrm{supp}(\mu)} |C_\mu(z)| \leq Const < \infty$.

We are grateful to V. Lomonosov and N.K. Nikolski who pointed out to us that this result has the following interpretation as a result about normal operators.

Theorem (on resolvents of normal operators): Let N be a normal operator whose spectrum $\sigma(N)$ has zero area. Let $R_\lambda, \lambda \in \mathbb{C} \setminus \sigma(N)$, denote its resolvent. If there are two vectors f, g such that g belongs to the closed linear span of $\{N^k f\}_{k\geqslant 0}$, $g \neq 0$, and such that $(R_\lambda f, g)$ is a bounded function on $\mathbb{C}\setminus\sigma(N)$, then there exists a nonzero vector h in the closed linear span of $\{N^k f\}_{k\geqslant 0}$, such that $(R_\lambda h, h)$ is a bounded function on $\mathbb{C} \setminus \sigma(N)$.

In other words, if a compact set supports a complex measure with bounded nonzero Cauchy transform, then this compact set supports a positive measure with bounded (and also automatically nonzero) Cauchy transform. Also if the resolvent of a normal operator is uniformly bounded on a pair of vectors f, g, ($g \neq 0$ being in the invariant subspace generated by f) then it is uniformly bounded on certain h, h, $h \neq 0$.

So this is what we will be proving using the "perfect hair" approach in what now follows.

Few words about methods used in the proof.

The probabilistic argument is a very important thing here. It is used to compensate for the roughness of our underlying measure. The other people have used before the arguments involving many dyadic lattices at once. We mean a paper by Garnett and Jones called "BMO from diadic BMO" [GJ].

We use dyadic martingale decomposition in our proof. We want to mention that looking at dyadic martingale decomposition is also a variation of an old theme, initiated, at least in the context of the Cauchy integral, by Coifman, Jones and Semmes in their paper [CJS]. There they proved a $T(b)$ theorem for the Cauchy integral using a Haar basis adapted to b. The main strategy of our proof is looking at dyadic martingale decomposition, but a random one!

Going further.

Let us recall the definitions of the Cauchy capacities. The first is *the complex Cauchy capacity* (not a very good name because it is a non-negative set function). We define it for $\nu \in M_c(K)$:= complex measures supported on K.

$$\gamma_c(K) := \{\sup|\nu(K)| : |C^\nu(z)| \le 1 \forall z \in \mathbb{C} \setminus K,\, \nu \in M_c(K)\}\,.$$

The second is *the positive Cauchy capacity* or just *the Cauchy capacity*:

$$\gamma_c(K) := \{\sup\mu(K) : |C^\mu(z)| \le 1 \forall z \in \mathbb{C} \setminus K,\, \nu \in M_+(K)\}\,.$$

Here $M_+(K)$ is a set of all positive measures supported on K. Obviously,

$$\gamma_+(K) \le \gamma_c(K) \le \gamma(K)\,.$$

We actually prove in this paper the following theorem (a sort of inverse to the previous left inequality).

Theorem. Let K be a compact set of zero area. Then

$$\gamma_+(K) \geq A\Big(1 + \Big(\frac{\operatorname{diam} K}{\gamma_c(K)}\Big)^2\Big(\frac{\|\nu\|}{\gamma_c(K)}\Big)^{42}\Big)^{-1/2}\gamma_c(K)\,, \tag{INV}$$

where ν is a measure that (almost) gives the supremum in the definition of γ_c. Its total variation in (INV) hinders us from proving that

$$\gamma_c \geq A\gamma_+\,.$$

Recently Xavier Tolsa [XT3] used (INV) and a very clever "induction on scales" that appeared in the preprint by J. Mateu, X. Tolsa and J. Verdera [MTV], in which it is shown that the condition conjectured by Mattila characterizes the Cantor sets of vanishing analytic capacity, to prove: .

$$\gamma_c \geq A\gamma_+\,.$$

This solves an old open problem. Actually, this implies the positive answer to Vitushkin's question whether the analytic capacity is semi-additive (with absolute constant). In fact, it is relatively easy to prove that γ_+ is semi-additive (see [NTV2], [NTV3]). The uniform comparability of γ_c and γ_+ implies uniform comparability of γ and γ_+ (indeed, this is just an easy approximation argument using the fact that for any compact set which is a finite union of rectifiable curves, γ_c coincides with γ_+).

Acknowledgements. We are grateful to Michael Frazier and Joan Verdera for many helpful remarks.

I. Suppressed operators K_Φ

Let Φ be a nonnegative Lipschitz function, i.e., $\Phi(x) \geqslant 0$ for every $x \in \mathbb{C}$ and

$$|\Phi(x) - \Phi(y)| \leqslant |x-y| \qquad \text{for every } x, y \in \mathbb{C}.$$

Define

$$k_\Phi(x,y) = \frac{\overline{x-y}}{|x-y|^2 + \Phi(x)\Phi(y)}\,.$$

Lemma: The kernel k_Φ is an antisymmetric Calderon-Zygmund kernel. It is also really well suppressed at the points where $\Phi(x) > 0$ or $\Phi(y) > 0$. Namely,

$$|k_\Phi(x,y)| \leqslant \frac{1}{\max\{\Phi(x), \Phi(y)\}} \qquad \text{for all } x, y \in \mathbb{C}.$$

Proof: Clearly,

$$|k_\Phi(x,y)| \leqslant \frac{1}{|x-y|} \qquad \text{and} \qquad k_\Phi(x,y) = -k_\Phi(y,x).$$

Since k_Φ is antisymmetric, to prove the second claim of the lemma, it is enough to show that $|k_\Phi(x,y)| < \frac{1}{\Phi(x)}$ for all $x, y \in \mathbb{C}$. We have $\Phi(y) \geqslant \Phi(x) - |x-y|$. Therefore

$$|k_\Phi(x,y)| \leqslant \frac{|x-y|}{|x-y|^2 + \Phi(x)(\Phi(x) - |x-y|)} = \frac{|x-y|}{|x-y|^2 + \Phi(x)^2 - \Phi(x)|x-y|}$$
$$= \frac{|x-y|}{\Phi(x)|x-y| + (\Phi(x) - |x-y|)^2} \leqslant \frac{1}{\Phi(x)},$$

and we are done.

To prove the first claim of the lemma, let us show that

$$|\nabla_x k_\Phi(x,y)| \leqslant \frac{4}{|x-y|^2}.$$

Indeed,

$$|\nabla_x k_\Phi(x,y)| \leqslant \frac{1}{|x-y|^2 + \Phi(x)\Phi(y)} + \frac{2|x-y|^2 + |x-y|\Phi(y)}{[|x-y|^2 + \Phi(x)\Phi(y)]^2}$$
$$\leqslant \frac{3}{|x-y|^2} + \frac{|x-y|\Phi(y)}{[|x-y|^2 + \Phi(x)\Phi(y)]^2} = \frac{3}{|x-y|^2} + \frac{\Phi(y)}{[|x-y|^2 + \Phi(x)\Phi(y)]}|k_\Phi(x,y)|$$
$$\leqslant \frac{3}{|x-y|^2} + \frac{\Phi(y)}{|x-y|^2}\frac{1}{\Phi(y)} = \frac{4}{|x-y|^2},$$

finishing the proof of the lemma.

From now on, we will denote by K_Φ the operator with kernel k_Φ.

Pick some very small number $\delta > 0$. It will stay fixed throughout the rest of the paper and will be used in many formulae without any special comment. The reader may think that δ is just an abbreviation for 45^{-239}.

II. Perfect random dyadic lattices and good functions

Let μ be a measure on the complex plane $\mathbb{C}$ satisfying $0 < \mu(\mathbb{C}) < +\infty$.

Assume that $\mathcal{D}$ is a random dyadic lattice (this phrase means that we have a *family* of dyadic lattices endowed with some probability P, and we use the letter $\mathcal{D}$ to denote *an element* in the family), and let Λ, $\{\Delta_Q\}_{Q\in\mathcal{D}}$ be the (random) family of projections associated with $\mathcal{D}$. As usual, this means that

$$\Lambda, \Delta_Q : L^2(\mu) \to L^2(\mu), \quad \Delta_Q\Lambda = \Lambda\Delta_Q = 0 \text{ for all } Q \in \mathcal{D}, \qquad \Delta_Q\Delta_R = 0 \text{ when } Q \neq R,$$

and for every function $\varphi \in L^2(\mu)$, one has

$$\varphi = \Lambda\varphi + \sum_{Q\in\mathcal{D}} \Delta_Q\varphi,$$

where the series converges at least in $L^2(\mu)$. Assume also that for every $\varphi \in L^2(\mu)$,

$$2^{-1}||\varphi||^2_{L^2(\mu)} \leqslant ||\Lambda\varphi||^2_{L^2(\mu)} + \sum_{Q\in\mathcal{D}} ||\Delta_Q\varphi||^2_{L^2(\mu)} \leqslant 2||\varphi||^2_{L^2(\mu)}.$$

Remark:

Let us make a couple of useful observations about such families of projections.

First of all, note that for every sequence of complex numbers $\{c_Q\}_{Q\in\mathcal{D}}$ that is finite in the sense that only finitely many c_Q do not vanish, we have

$$2^{-1}\sum_{Q\in\mathcal{D}}|c_Q|^2||\Delta_Q\varphi||^2_{L^2(\mu)}\leqslant\Big\|\sum_{Q\in\mathcal{D}}c_Q\Delta_Q\varphi\Big\|^2_{L^2(\mu)}\leqslant 2\sum_{Q\in\mathcal{D}}|c_Q|^2||\Delta_Q\varphi||^2_{L^2(\mu)}.$$

Indeed, consider the function $\widetilde{\varphi}:=\sum_{Q\in\mathcal{D}}c_Q\Delta_Q\varphi$ and note that $\Lambda\widetilde{\varphi}=0$, $\Delta_Q\widetilde{\varphi}=c_Q\Delta_Q\varphi$. Now it remains only to apply our assumption to the function $\widetilde{\varphi}$ instead of φ itself.

Now take any function $\psi\in L^2(\mu)$. We have

$$\Big|\sum_{Q\in\mathcal{D}}c_Q\langle\Delta_Q\varphi,\psi\rangle\Big|=\Big|\Big\langle\sum_{Q\in\mathcal{D}}c_Q\Delta_Q\varphi,\psi\Big\rangle\Big|$$

$$\leqslant\Big\|\sum_{Q\in\mathcal{D}}c_Q\Delta_Q\varphi\Big\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)}\leqslant\sqrt{2}\|\psi\|_{L^2(\mu)}\Big[\sum_{Q\in\mathcal{D}}|c_Q|^2||\Delta_Q\varphi||^2_{L^2(\mu)}\Big]^{\frac12}.$$

In particular, this means that if $\mathcal{F}\subset\mathcal{D}$ is some family of dyadic squares, then

$$\sum_{Q\in\mathcal{F}}|\langle\Delta_Q\varphi,\psi\rangle|\leqslant\sqrt{2}\|\psi\|_{L^2(\mu)}\Big[\sum_{Q\in\mathcal{F}}||\Delta_Q\varphi||^2_{L^2(\mu)}\Big]^{\frac12}$$

(just take $c_Q=0$ for $Q\notin\mathcal{F}$ and choose c_Q for $Q\in\mathcal{F}$ in such a way that $|c_Q|=1$ and $c_Q\langle\Delta_Q\varphi,\psi\rangle=|\langle\Delta_Q\varphi,\psi\rangle|$; if the family $\mathcal{F}$ is infinite, do it for all its finite subfamilies and then pass to the supremum).

Also, let us take any finite family $\mathcal{F}\subset\mathcal{D}$ such that $\|\Delta_Q\varphi\|_{L^2(\mu)}>0$ for every $Q\in\mathcal{F}$. Take $c_Q=0$ for $Q\notin\mathcal{F}$ and choose c_Q for $Q\in\mathcal{F}$ in such a way that $|c_Q|=\dfrac{|\langle\Delta_Q\varphi,\psi\rangle|}{||\Delta_Q\varphi||^2_{L^2(\mu)}}$ and $c_Q\langle\Delta_Q\varphi,\psi\rangle=\dfrac{|\langle\Delta_Q\varphi,\psi\rangle|^2}{||\Delta_Q\varphi||^2_{L^2(\mu)}}$.

Then we get

$$\sum_{Q\in\mathcal{F}}\frac{|\langle\Delta_Q\varphi,\psi\rangle|^2}{||\Delta_Q\varphi||^2_{L^2(\mu)}}\leqslant\sqrt{2}\|\psi\|_{L^2(\mu)}\Big[\sum_{Q\in\mathcal{F}}\frac{|\langle\Delta_Q\varphi,\psi\rangle|^2}{||\Delta_Q\varphi||^2_{L^2(\mu)}}\Big]^{\frac12},$$

or, which is the same,

$$\sum_{Q\in\mathcal{F}}\frac{|\langle\Delta_Q\varphi,\psi\rangle|^2}{||\Delta_Q\varphi||^2_{L^2(\mu)}}\leqslant 2\|\psi\|^2_{L^2(\mu)}.$$

Now, of course, the summation on the left can be extended to all squares Q for which $\|\Delta_Q\varphi\|_{L^2(\mu)}>0$.

We will not need anything beyond this, so we are not going to say the magic words that the projections Λ and $\{\Delta_Q\}_{Q\in\mathcal{D}}$ generate a Riesz basis of subspaces in $L^2(\mu)$ to a reader who does not want to hear them.

Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two independent copies of the random dyadic lattice $\mathcal{D}$. Suppose that there is some rule which allows one to tell, for every square $Q_1 \in \mathcal{D}_1$, whether it is "bad" or "good" with respect to the lattice $\mathcal{D}_2$. Of course, since $\mathcal{D}_1$ and $\mathcal{D}_2$ are copies of *the same* random dyadic lattice, we can use the same rule to define bad squares in $\mathcal{D}_2$ with respect to $\mathcal{D}_1$.

Our next assumption is that bad squares are very rare. Namely, we suppose that for every fixed $\mathcal{D}_1$ and for every $Q_1 \in \mathcal{D}_1$, the probability

$$P_{\mathcal{D}_2}\{Q_1 \text{ is bad}\} \leqslant \delta$$

(and vice versa, of course).

If all the above assumptions are satisfied, we will say that $\mathcal{D}$ is a *perfect random dyadic lattice.*

Let again $\mathcal{D}_1$ and $\mathcal{D}_2$ be two independent copies of a random dyadic lattice $\mathcal{D}$.

A function $\varphi_1 \in L^2(\mu)$ is called good (the full name should be $\mathcal{D}_1$-good with respect to the lattice $\mathcal{D}_2$, or something like that) if for every bad square $Q_1 \in \mathcal{D}_1$, we have

$$\Delta_{Q_1}\varphi_1 = 0.$$

Even if a function $\varphi_1 \in L^2(\mu)$ is not good, we still can write the decomposition

$$\varphi_1 = \Big[\Lambda_1\varphi_1 + \sum_{\substack{Q_1\in\mathcal{D}_1,\\ Q_1 \text{ is good}}} \Delta_{Q_1}\varphi_1\Big] + \sum_{\substack{Q_1\in\mathcal{D}_1,\\ Q_1 \text{ is bad}}} \Delta_{Q_1}\varphi_1 =: (\varphi_1)_{\text{good}} + (\varphi_1)_{\text{bad}}.$$

Note that this decomposition depends on both $\mathcal{D}_1$ and $\mathcal{D}_2$, and therefore $(\varphi_1)_{\text{good}}$ and $(\varphi_1)_{\text{bad}}$ are random functions even if φ_1 is a sure function. If the dyadic lattice $\mathcal{D}$ is perfect, it is easy to show that always

$$\|(\varphi_1)_{\text{good}}\|_{L^2(\mu)}, \|(\varphi_1)_{\text{bad}}\|_{L^2(\mu)} \leqslant 2\|\varphi_1\|_{L^2(\mu)}.$$

What is more, if φ_1 does not depend on $\mathcal{D}_2$, then for every fixed $\mathcal{D}_1$,

$$\mathbb{E}_{\mathcal{D}_2}\|(\varphi_1)_{\text{bad}}\|^2_{L^2(\mu)} \leqslant 4\delta\|\varphi_1\|^2_{L^2(\mu)}.$$

Indeed, we have

$$\mathbb{E}_{\mathcal{D}_2}\|(\varphi_1)_{\text{bad}}\|^2_{L^2(\mu)} = \mathbb{E}_{\mathcal{D}_2}\Big\|\sum_{\substack{Q_1\in\mathcal{D}_1,\\ Q_1 \text{ is bad}}} \Delta_{Q_1}\varphi_1\Big\|^2_{L^2(\mu)} \leqslant 2\mathbb{E}_{\mathcal{D}_2}\sum_{\substack{Q_1\in\mathcal{D}_1,\\ Q_1 \text{ is bad}}} \|\Delta_{Q_1}\varphi_1\|^2_{L^2(\mu)}$$

$$= 2\sum_{Q_1\in\mathcal{D}_1} P_{\mathcal{D}_2}\{Q_1 \text{ is bad}\}\|\Delta_{Q_1}\varphi_1\|^2_{L^2(\mu)} \leqslant 2\delta\sum_{Q_1\in\mathcal{D}_1}\|\Delta_{Q_1}\varphi_1\|^2_{L^2(\mu)} \leqslant 4\delta\|\varphi_1\|^2_{L^2(\mu)}.$$

Hence for all sure functions φ_1, we have

$$\mathbb{E}\|(\varphi_1)_{\text{bad}}\|^2_{L^2(\mu)} \leqslant 4\delta\|\varphi_1\|^2_{L^2(\mu)}.$$

III. Perfect hair

Let again μ be a measure on the complex plane $\mathbb{C}$ satisfying $0 < \mu(\mathbb{C}) < +\infty$.

Assume that we have a perfect random dyadic lattice $\mathcal{D}$ (i.e., a family of dyadic lattices endowed with some probability so that the assumptions of the previous section are satisfied) and suppose that with every dyadic lattice $\mathcal{D}$ in that family a nonnegative Lipschitz function $\Phi_{\mathcal{D}}$ is associated in such a way that the following properties hold: *1)* $\mu\{x \in \mathbb{C} : \Phi_{\mathcal{D}}(x) > 0\} \leqslant \delta\mu(\mathbb{C})$ *for every* $\mathcal{D}$*; 2) For every two dyadic lattices* $\mathcal{D}_1, \mathcal{D}_2$*, for every Lipschitz function* Θ *satisfying* $\inf_{\mathbb{C}} \Theta > 0$*,* $\Theta \geqslant \delta \max(\Phi_{\mathcal{D}_1}, \Phi_{\mathcal{D}_2})$*, and for any two good functions* φ_1 *and* φ_2 *(with respect to the lattices* $\mathcal{D}_1$ *and* $\mathcal{D}_2$*, correspondingly), we have*

$$\big|\langle\varphi_1, K_\Theta\varphi_2\rangle\big| \leqslant N\|\varphi_1\|_{L^2(\mu)}\|\varphi_2\|_{L^2(\mu)},$$

where N *is some (large) positive constant, not depending on* φ_1*,* φ_2 *or* Θ*.*

(The assumption $\inf_{\mathbb{C}} \Theta > 0$ is purely technical, of course: it just allows us to avoid tiresome discussions concerning the definition of $K_\Theta\varphi_2$: the kernel is uniformly bounded, the measure is finite, so everything makes sense.) Then we will say that we have *"perfect hair"*.

Our first aim is to show that every perfect hair generates a bounded (in $L^2(\mu)$) operator, which coincides with the Cauchy integral operator everywhere outside an exceptional set of small μ-measure.

IV. Truncated mathematical expectation

Let ξ be a nonnegative random variable and let $0 < \beta < 1$. Define

$$\mathbb{E}_\beta\xi := \inf\Big\{\int_{\Omega\setminus\Omega_1} \xi\, dP \,:\, P\{\Omega_1\} \leqslant \beta\Big\}$$

Note that *A) If* $P\{\xi > 0\} \leqslant \beta$*, then* $\mathbb{E}_\beta\xi = 0$*; B)* $P\{\xi \geqslant \beta^{-1}\mathbb{E}_\beta\xi\} \leqslant 2\beta$*; C) If* $\Phi_\omega(x)$ *(*$x \in \mathbb{C}$*) is a random nonnegative Lipschitz function, then* $\mathbb{E}_\beta\Phi_\omega(x)$ *is a certain nonnegative Lipschitz function.*

V. How to use perfect hair

Theorem:
Assume that we have a perfect hair. Let $\beta = \sqrt{\delta}$. Let $\Phi := \mathbb{E}_\beta\Phi_{\mathcal{D}}$.
Then
1) $\mu\{x \in \mathbb{C} : \Phi(x) > 0\} \leqslant \sqrt{\delta}\mu(\mathbb{C})$; 2) The operator K_Φ acts in $L^2(\mu)$ in the sense that

$\sup_{\lambda>0} ||K_{\Phi+\lambda}||_{L^2(\mu)\to L^2(\mu)} < +\infty$.

Proof:

The first claim is easy: note that

$$\mathbb{E}\mu\{x \in \mathbb{C} : \Phi_{\mathcal{D}}(x) > 0\} \leqslant \delta\mu(\mathbb{C}),$$

and thereby for the set

$$E := \{x \in \mathbb{C} : P\{\Phi_{\mathcal{D}}(x) > 0\} \geqslant \beta = \sqrt{\delta}\},$$

we have $\mu(E) \leqslant \sqrt{\delta}\mu(\mathbb{C})$. It remains only to recall that, according to property (A) of the truncated mathematical expectation, $\Phi = \mathbb{E}_\beta \Phi_{\mathcal{D}} \equiv 0$ outside E.

Now we will prove even a little bit more than the second claim. Namely, we will show that

$$\sup\{||K_\Theta|| : \Theta \text{ is Lipschitz }, \Theta \geqslant \Phi\} < +\infty$$

(in the same sense as above; see the exact formulation below).

Fix $\lambda > 0$ and let

$$N_\lambda = \sup\{||K_\Theta|| : \Theta \text{ is Lipschitz }, \Theta \geqslant \Phi + \lambda\}.$$

Clearly, for every $\lambda > 0$, we have $N_\lambda \leqslant \frac{\mu(\mathbb{C})}{\lambda} < +\infty$. We are going to prove that N_λ is bounded by some constant *independent* of λ.

"Space" and "frequency" reductions

Choose $\Theta \geqslant \Phi + \lambda$ and functions $\varphi_1, \varphi_2 \in L^2(\mu)$ with $||\varphi_1||_{L^2(\mu)} = ||\varphi_2||_{L^2(\mu)} = 1$ such that

$$\left|\langle \varphi_1, K_\Theta \varphi_2 \rangle\right| \geqslant \tfrac{9}{10} N_\lambda.$$

Consider two independent copies $\mathcal{D}_1$ and $\mathcal{D}_2$ of a perfect random dyadic lattice $\mathcal{D}$. Let

$$S := \{x \in \mathbb{C} : \max_{j=1,2} \Phi_{\mathcal{D}_j}(x) \geqslant \beta^{-1}\Phi(x)\}.$$

Put

$$\varphi_j' := \varphi_j \chi_S, \qquad \widetilde{\varphi}_j := \varphi_j \chi_{\mathbb{C}\setminus S} = \varphi_j - \varphi_j' \qquad j = 1, 2$$

("space" reduction) and, at last,

$$\psi_j := (\widetilde{\varphi}_j)_{\text{good}} = \widetilde{\varphi}_j - (\widetilde{\varphi}_j)_{\text{bad}}, \qquad j = 1, 2$$

("frequency" reduction).

We expect both reductions to be just "minor corrections". Soon we will show that this really is the case, namely, that

$$||\varphi_j'||_{L^2(\mu)}, ||(\widetilde{\varphi}_j)_{\text{bad}}||_{L^2(\mu)} \leqslant \frac{1}{10} \qquad (*)$$

with probability close to 1. Now let us demonstrate that these reductions really make sense.

Pick a pair of dyadic lattices $\mathcal{D}_1$ and $\mathcal{D}_2$, for which $(*)$ holds. We have

$$\langle\widetilde{\varphi}_1, K_\Theta\widetilde{\varphi}_2\rangle = \langle\varphi_1, K_\Theta\varphi_2\rangle - \langle\varphi_1', K_\Theta\varphi_2\rangle - \langle\widetilde{\varphi}_1, K_\Theta\varphi_2'\rangle$$

and thereby

$$\left|\langle\widetilde{\varphi}_1, K_\Theta\widetilde{\varphi}_2\rangle\right| \geqslant \tfrac{9}{10}N_\lambda - \tfrac{2}{10}\|K_\Theta\| \geqslant \tfrac{7}{10}N_\lambda$$

(here we used the obvious estimate $\|\widetilde{\varphi}_1\|_{L^2(\mu)} \leqslant \|\varphi_1\|_{L^2(\mu)} = 1$ together with $(*)$ to get the first inequality).

The key observation about the space reduction is that

$$\langle\widetilde{\varphi}_1, K_\Theta\widetilde{\varphi}_2\rangle = \iint k_\Theta(x_1,x_2)\widetilde{\varphi}_1(x_1)\widetilde{\varphi}_2(x_2)\,d\mu(x_1)d\mu(x_2) =$$

$$\iint k_{\Theta'}(x_1,x_2)\widetilde{\varphi}_1(x_1)\widetilde{\varphi}_2(x_2)\,d\mu(x_1)d\mu(x_2) = \langle\widetilde{\varphi}_1, K_{\Theta'}\widetilde{\varphi}_2\rangle,$$

where

$$\Theta' := \max\{\Theta, \beta\Phi_{\mathcal{D}_1}, \beta\Phi_{\mathcal{D}_2}\}.$$

We still have Θ' Lipschitz and satisfying $\Theta' \geqslant \Phi + \lambda$, but now also $\Theta' \geqslant \delta\max\{\Phi_{\mathcal{D}_1}, \Phi_{\mathcal{D}_2}\}$, and therefore we only need to make the functions $\widetilde{\varphi}_j$ good to apply property (2) of perfect hair and to finish the story. This is exactly what the frequency reduction does. Like above, we can write

$$\langle\psi_1, K_{\Theta'}\psi_2\rangle = \langle\widetilde{\varphi}_1, K_{\Theta'}\widetilde{\varphi}_2\rangle - \langle(\widetilde{\varphi}_1)_{\text{bad}}, K_{\Theta'}\widetilde{\varphi}_2\rangle - \langle\psi_1, K_{\Theta'}(\widetilde{\varphi}_2)_{\text{bad}}\rangle$$

and thereby

$$\left|\langle\psi_1, K_{\Theta'}\psi_2\rangle\right| \geqslant \tfrac{7}{10}N_\lambda - \tfrac{3}{10}\|K_{\Theta'}\| \geqslant \tfrac{4}{10}N_\lambda$$

(here we used the estimate $\|\psi_j\|_{L^2(\mu)} \leqslant \|\widetilde{\varphi}_j\|_{L^2(\mu)} + \|(\widetilde{\varphi}_j)_{\text{bad}}\|_{L^2(\mu)} < 2$ together with $(*)$ to get the first inequality).

Now, according to property (2) of perfect hair, the left hand part does not exceed $4N$ and we get $N_\lambda \leqslant 10N$. It remains only to prove that $(*)$ holds with probability close to 1.

Note that for any given point $x \in \mathbb{C}$, we have $P\{x \in S\} \leqslant 4\beta$, and therefore,

$$\mathbb{E}\|\varphi_j'\|^2_{L^2(\mu)} \leqslant 4\beta, \qquad j = 1, 2.$$

Hence,

$$P\big\{\|\varphi_j'\|_{L^2(\mu)} \geqslant \beta^{\frac{1}{3}}\big\} \leqslant 4\beta^{\frac{1}{3}} \qquad j = 1, 2.$$

Now we would like to say that the norms of the functions $(\widetilde{\varphi}_j)_{\text{bad}}$ are small as well. Unfortunately, as constructed, each of them depends on both $\mathcal{D}_1$ and $\mathcal{D}_2$. So it seems that we can only apply the obvious estimate $\|(\widetilde{\varphi}_j)_{\text{bad}}\|_{L^2(\mu)} \leqslant 2\|(\widetilde{\varphi}_j)\|_{L^2(\mu)} \leqslant 2$, which is clearly useless.

Note, nevertheless, that

$$(\widetilde{\varphi}_j)_{\text{bad}} = (\varphi_j)_{\text{bad}} - (\varphi_j')_{\text{bad}}.$$

The norm of $(\varphi'_j)_{\text{bad}}$ does not exceed $2\|\varphi'_j\|_{L^2(\mu)}$. As to $(\varphi_j)_{\text{bad}}$, we can apply the estimate for sure functions to φ_j, which yields

$$\mathbb{E}||(\varphi_j)_{\text{bad}}||^2_{L^2(\mu)} \leqslant 4\delta.$$

So finally we conclude that with probability at least $1-8\beta^{\frac{1}{3}}-8\delta^{\frac{1}{3}} > \frac{9}{10}$ all the norms in the left hand part of $(*)$ are bounded by $2\beta^{\frac{1}{3}}+\delta^{\frac{1}{3}} < \frac{1}{10}$.

VI. Lyric deviation: Hausdorff measure and analytic capacity

We will start with a couple of definitions.

The 1-dimensional Hausdorff measure

Let $\varepsilon > 0$. For every set $E \subset \mathbb{C}$ define

$$\mathcal{H}_\varepsilon(E) := \inf\Big\{\sum_j r_j \,:\, E \subset \bigcup_j B(x_j, r_j), \quad x_j \in \mathbb{C},\, r_j \leqslant \varepsilon\Big\}$$

(the infimum is taken over all (countable) coverings of E by open disks $B(x_j, r_j)$ with radii $r_j \leqslant \varepsilon$).

It is clear that $\mathcal{H}_\varepsilon$ is an outer measure and that if $\varepsilon' \leqslant \varepsilon''$, then $\mathcal{H}_{\varepsilon'}(E) \geqslant \mathcal{H}_{\varepsilon''}(E)$ for every $E \subset \mathbb{C}$. Since every monotone function has a limit (maybe, infinite), we can define

$$\mathcal{H}(E) := \lim_{\varepsilon\to 0} \mathcal{H}_\varepsilon(E) = \sup_{\varepsilon>0} \mathcal{H}_\varepsilon(E).$$

It is a trivial exercise to show that $\mathcal{H}$ is an outer measure. However, it is much better than just that, namely, $\mathcal{H}$ is *a Borel measure*. The proof of this remarkable theorem can be found in any (good) textbook on measure theory. We can only regret that it is not included in the Leningrad (or Michigan) State University analysis course.

Analytic capacity

Let $F \subset \mathbb{C}$ be a compact set. We will say that F has positive analytic capacity if there exists a bounded analytic function $f : \mathbb{C} \setminus F \to \mathbb{C}$, which is not identically 0 and such that $f(x) \to 0$ as $x \to \infty$. Assume now that we have a compact set F of positive analytic capacity and such that $\mathcal{H}(F) < +\infty$. Let f be the corresponding bounded analytic function.

VII. Cauchy integral representation

We devote this section to a well-known representation of bounded analytic function outside of a compact of finite length. Since F is compact, we can consider only finite coverings in the definition of $\mathcal{H}_\varepsilon(F)$. Now for every positive integer n, construct a covering

$$\bigcup_{j=1}^{N(n)} B(x_j^{(n)}, r_j^{(n)}) \supset F$$

such that all $r_j^{(n)} \leqslant \frac{1}{n}$, $\sum_j r_j^{(n)} \leqslant \mathcal{H}(F) + \frac{1}{n}$ and $B(x_j^{(n)}, r_j^{(n)}) \cap F \neq \emptyset$ for every j. Let $\Omega_n = \mathbb{C} \setminus \operatorname{clos} \bigcup_j B(x_j^{(n)}, r_j^{(n)})$ and let $\Gamma_n := \partial\Omega_n$. Note that Γ_n is a good contour (consisting of finitely many arcs) and that $\Gamma_n \subset \mathbb{C} \setminus F$. Therefore for every point $x \in \Omega_n$, we can write the standard Cauchy formula

$$f(x) = -\frac{1}{2\pi i} \oint_{\Gamma_n} \frac{f(y)\,dy}{x-y} = \int_{\mathbb{C}} \frac{d\nu_n(y)}{x-y},$$

where ν_n is a complex-valued measure defined (on Borel sets, say) by

$$\nu_n(E) = -\frac{1}{2\pi i} \oint_{\Gamma_n \cap E} f(y)\,dy.$$

Note that the variations of the complex-valued measures ν_n are uniformly bounded (by $||f||_{L^\infty}(\mathcal{H}(F)+1)$, say), and therefore (passing to a subsequence, if needed) we may assume that ν_n weakly converge to a complex-valued Borel measure ν (over the space $C_0(\mathbb{C})$ of all compactly supported complex-valued continuous functions on $\mathbb{C}$). Note now that Ω_n contains all points $x \in \mathbb{C}$ for which $\operatorname{dist}(x, F) > \frac{1}{n}$. Therefore for any $\eta \in C_0(\mathbb{C})$ satisfying $\operatorname{supp} \varphi \subset \mathbb{C} \setminus F$, we have

$$\int_{\mathbb{C}} \eta\, d\nu = \lim_{n\to\infty} \int_{\mathbb{C}} \eta\, d\nu_n = 0,$$

i.e., $\operatorname{supp} \nu \subset F$. Now passing to the limit in the Cauchy formula above, we see that for every $x \in \mathbb{C} \setminus F$,

$$f(x) = \int_{\mathbb{C}} \frac{d\nu(y)}{x-y}.$$

Our next step will be to show that for every Borel set $E \subset \mathbb{C}$ we have

$$|\nu|(E) \leqslant \|f\|_{L^\infty} \mathcal{H}F(E),$$

where $\mathcal{H}F(E) := \mathcal{H}(E \cap F)$.

Recall that every finite Borel measure μ is regular in the sense that for every Borel set E one can find an open set $G \supset E$ such that $\mu(G \setminus E)$ is as small as one wants. Therefore it is enough to prove this inequality for open sets only.

Recall also that for an open set G,

$$|\nu|(G) = \sup\Big\{\Big|\int_{\mathbb{C}} \eta\, d\nu\Big| \,:\, \eta \in C_0(\mathbb{C}), \operatorname{supp} \eta \subset G, ||\eta||_{L^\infty} \leqslant 1\Big\}.$$

Therefore we need only to show that for every such η,

$$\Big|\int_{\mathbb{C}} \eta\, d\nu\Big| \leqslant \|f\|_{L^\infty} \mathcal{H}F(G).$$

But we have

$$\int_{\mathbb{C}} \eta\, d\nu = \lim_{n\to\infty} \int_{\mathbb{C}} \eta\, d\nu_n,$$

and for every n,

$$\Big|\int_{\mathbb{C}} \eta\, d\nu\Big| \leqslant \|f\|_{L^\infty} \sum_{j : B(x_j^{(n)}, r_j^{(n)}) \cap \operatorname{supp} \eta \neq \emptyset} r_j^{(n)}.$$

Now notice that if $\frac{1}{n} < \operatorname{dist}(\operatorname{supp}\eta, \partial G)$, then the disks $B(x_j^{(n)}, r_j^{(n)})$ that intersect $\operatorname{supp}\eta$ cannot participate in the covering of $F \setminus G$. Therefore,

$$\sum_{j: B(x_j^{(n)}, r_j^{(n)}) \cap \operatorname{supp}\eta \neq \emptyset} r_j^{(n)} \leqslant \sum_{j=1}^{N(n)} r_j^{(n)} - \sum_{j: B(x_j^{(n)}, r_j^{(n)}) \cap F \setminus G \neq \emptyset} r_j^{(n)}$$

$$\leqslant \mathcal{H}(F) + \frac{1}{n} - \mathcal{H}_{\frac{1}{n}}(F \setminus G) \to \mathcal{H}(F) - \mathcal{H}(F \setminus G) = \mathcal{H}(F \cap G) = \mathcal{H}F(G)$$

as $n \to \infty$, proving the claim.

Applying the Radon-Nykodim theorem, we conclude that there exists a Borel measurable function h satisfying $\|h\|_{L^\infty} \leqslant \|f\|_{L^\infty}$ and such that

$$f(x) = \int_{\mathbb{C}} \frac{h(y)}{x-y} d\mathcal{H}F(y)$$

for every $\mathbb{C} \setminus F$ (there is no problem with convergence here, because, as we remember, the integral is actually taken over F). Note also that since $f(x) \neq 0$ for at least one $x \in \mathbb{C}$, we should have $\mathcal{H}F\{y \in \mathbb{C} : h(y) \neq 0\} > 0$. As a trivial consequence, we observe that $\mathcal{H}(F) > 0$.

VIII. The Ahlfors radius $\mathcal{R}(x)$

Now take some large $M > 1$. We will call a disk $B(x, r)$ $(x \in \mathbb{C}, r > 0)$ *non-Ahlfors*, if

$$\mathcal{H}F(B(x, r)) > Mr.$$

For every point $x \in \mathbb{C}$ define its Ahlfors radius $\mathcal{R}(x)$ by

$$\mathcal{R}(x) := \sup\{r > 0 : B(x, r) \text{ is non-Ahlfors}\}.$$

Since f is bounded on $\mathbb{C} \setminus F$, so is the Cauchy integral $\int_{\mathbb{C}} \frac{h(y)}{x-y} d\mathcal{H}F(y)$. We are going to show that, in a sense, this integral stays bounded on F as well (where f, generally speaking, does not exist). Namely, for every $x \in \mathbb{C}$

$$\sup_{\varepsilon > \mathcal{R}(x)} \left| \int_{\mathbb{C} \setminus B(x, \varepsilon)} \frac{h(y)}{x-y} d\mathcal{H}F(y) \right| \leqslant 7M\|f\|_{L^\infty}.$$

Proof:

Note first of all, that the condition $\mathcal{H}(F) < +\infty$ implies that the 2-dimensional Lebesgue measure $m(F) = 0$. Indeed, for every covering $\bigcup_j B(x_j, r_j) \supset F$, we have

$$m(F) \leqslant \pi \sum_j r_j^2 \leqslant \pi(\max_j r_j) \sum_j r_j.$$

Therefore

$$m(F) \leqslant \pi\varepsilon\mathcal{H}_\varepsilon(F) \leqslant \pi\varepsilon\mathcal{H}(F)$$

for every $\varepsilon > 0$, and we are done.

Now compare $\int_{\mathbb{C}\setminus B(x,\varepsilon)} \frac{h(y)}{x-y} d\mathcal{H}F(y)$ to

$$\frac{4}{\pi\varepsilon^2}\int_{B(x,\frac{\varepsilon}{2})} f(z)\,dm(z) = \frac{4}{\pi\varepsilon^2}\int_{B(x,\frac{\varepsilon}{2})\setminus F} f(z)\,dm(z),$$

which is clearly bounded by $\|f\|_{L^\infty} \leqslant M\|f\|_{L^\infty}$. Using the Cauchy integral representation for f, we see that the difference equals

$$-\frac{4}{\pi\varepsilon^2}\int_{B(x,\frac{\varepsilon}{2})}\Big(\int_{B(x,\varepsilon)} \frac{h(y)}{z-y} d\mathcal{H}F(y)\Big) dm(z) +$$

$$\int_{\mathbb{C}\setminus B(x,\varepsilon)} h(y)\Big(\frac{1}{x-y} - \frac{4}{\pi\varepsilon^2}\int_{B(x,\frac{\varepsilon}{2})} \frac{dm(z)}{z-y}\Big) d\mathcal{H}F(y) =: I' + I''.$$

The integral I' allows the rough estimate

$$|I'| \leqslant \frac{4}{\pi\varepsilon^2}\|f\|_{L^\infty}\int_{B(x,\varepsilon)}\Big(\int_{B(x,\frac{\varepsilon}{2})} \frac{dm(z)}{|z-y|}\Big) d\mathcal{H}F(y).$$

Since the inner integral does not exceed $\pi\varepsilon$ for every $y \in \mathbb{C}$, we get

$$|I'| \leqslant 4\varepsilon^{-1}\|f\|_{L^\infty}\mathcal{H}F(B(x,\varepsilon)) \leqslant 4M\|f\|_{L^\infty},$$

provided that $\varepsilon \geqslant \mathcal{R}(x)$.

To estimate I'', note that

$$\Big|\frac{1}{x-y} - \frac{4}{\pi\varepsilon^2}\int_{B(x,\frac{\varepsilon}{2})} \frac{dm(z)}{z-y}\Big| = \frac{4}{\pi\varepsilon^2}\Big|\int_{B(x,\frac{\varepsilon}{2})} \frac{z-x}{(x-y)(z-y)} dm(z)\Big| \leqslant \frac{\varepsilon}{|x-y|^2}$$

because $|z-x| \leqslant \frac{\varepsilon}{2}$ and $2|z-y| \geqslant |x-y|$ for every $y \in \mathbb{C}\setminus B(x,\varepsilon)$, $z \in B(x,\frac{\varepsilon}{2})$.

Thus

$$|I''| \leqslant \|f\|_{L^\infty}\int_{\mathbb{C}\setminus B(x,\varepsilon)} \frac{\varepsilon}{|x-y|^2} d\mathcal{H}F(y).$$

To estimate the last integral, we need the following obvious lemma, which we will frequently use in the future.

Comparison Lemma: Let $S \subset \mathbb{C}$. Assume that we have a measure μ satisfying

$$\mu\{x \in \mathbb{C} : \operatorname{dist}(x,S) < r\} \leqslant Mr \qquad \text{for every } r \geqslant R_0$$

and a nonnegative continuous decreasing function $U(t)$ $(t > 0)$.

Then for every $R \geqslant R_0$

$$\int_{\{x:\operatorname{dist}(x,s)\geqslant R\}} U(\operatorname{dist}(y,S))d\mu(y) \leqslant M\Big(RU(R) + \int_R^{+\infty} U(t)dt\Big).$$

Note also that the quantity in parentheses can be viewed as the integral over the whole ray $[0,\infty)$ of $\min\{U(t),U(R)\}$ and therefore is a decreasing function in R. So, we can replace R on the right hand side by any lesser number if we want to.

The Comparison Lemma (with $S=\{x\}$, $R_0=\mathcal{R}(x)$, $R=\varepsilon$ and $U(t)=\frac{\varepsilon}{t^2}$) yields

$$\int_{\mathbb{C}\setminus B(x,\varepsilon)}\frac{\varepsilon}{|x-y|^2}d\mathcal{H}F(y)\leqslant M\Big(1+\int_{\varepsilon}^{+\infty}\frac{\varepsilon}{t^2}dt\Big)=2M,$$

and thereby $|I''|\leqslant 2M\|f\|_{L^\infty}$. It remains only to add the estimates to get the desired inequality.

The additional assumption $\varepsilon\geqslant\mathcal{R}(x)$ in the formulation of the last statement seems quite unpleasant. We would prefer to have a result that is valid for every $\varepsilon>0$. This can be achieved if we replace the kernel $\frac{1}{x-y}$ by the suppressed kernel k_Φ with a Lipschitz function Φ satisfying $\Phi(x)\geqslant\delta\mathcal{R}(x)$ for every $x\in\mathbb{C}$.

Lemma:

For every $x\in\mathbb{C}$

$$\sup_{\varepsilon>0}\Big|\int_{\mathbb{C}\setminus B(x,\varepsilon)}k_\Phi(x,y)h(y)d\mathcal{H}F(y)\Big|\leqslant(11+\delta^{-1})M\|f\|_{L^\infty}.$$

Proof: Recall that the kernel k_Φ is given by

$$k_\Phi(x,y)=\frac{\overline{x-y}}{|x-y|^2+\Phi(x)\Phi(y)}.$$

Put $r:=\Phi(x)$, $R:=\delta^{-1}\Phi(x)\,(\geqslant\mathcal{R}(x))$ and, at last, $R':=\max\{\varepsilon,R\}$. Write

$$\int_{\mathbb{C}\setminus B(x,\varepsilon)}k_\Phi(x,y)h(y)d\mathcal{H}F(y)=\int_{\mathbb{C}\setminus B(x,R')}\dots+\int_{B(x,R)\setminus B(x,\varepsilon)}\dots=:I'+I''$$

(R in the second integral is not a misprint: we need this second term only for $R>\varepsilon$ when $R'=R$).

Recall that $|k_\Phi(x,y)|\leqslant\frac{1}{\Phi(x)}=r^{-1}$ for all $y\in\mathbb{C}$.

Thus

$$|I''|\leqslant\|f\|_{L^\infty}r^{-1}\mathcal{H}F(B(x,R))\leqslant\|f\|_{L^\infty}r^{-1}MR=\delta^{-1}M\|f\|_{L^\infty}.$$

As to I', let us compare it to $\widetilde{I}:=\int_{\mathbb{C}\setminus B(x,R')}\frac{h(y)}{x-y}d\mathcal{H}F(y)$, which is bounded by $7M\|f\|_{L^\infty}$, because $R'\geqslant\mathcal{R}(x)$. The difference does not exceed

$$\|f\|_{L^\infty}\int_{\mathbb{C}\setminus B(x,R')}\Big|\frac{1}{x-y}-\frac{\overline{x-y}}{|x-y|^2+\Phi(x)\Phi(y)}\Big|d\mathcal{H}F(y).$$

Representing $\frac{1}{x-y}$ as $\frac{\overline{x-y}}{|x-y|^2}$ and observing that for every two numbers $t,s>0$, one has

$$\frac{1}{t}-\frac{1}{t+s}=\frac{s}{t(t+s)}\leqslant\frac{s}{t^2},$$

we get

$$\Big|\frac{1}{x-y}-\frac{\overline{x-y}}{|x-y|^2+\Phi(x)\Phi(y)}\Big|\leqslant\frac{\Phi(x)\Phi(y)}{|x-y|^3}\leqslant\frac{r(r+|x-y|)}{|x-y|^3}.$$

Applying the Comparison Lemma again, we see that

$$|I' - \widetilde{I}| \leqslant M\|f\|_{L^\infty}\Big[r\frac{r(r+r)}{r^3} + \int_r^{+\infty}\frac{r(r+t)}{t^3}dt\Big] = \tfrac{7}{2}M\|f\|_{L^\infty} \leqslant 4M\|f\|_{L^\infty}$$

(here, in order to simplify calculations, we used the possibility to replace R' by the lesser number r).

Now it remains only to bring all the estimates together to get the conclusion of the lemma.

IX. The exceptional set H

The demand $\Phi(x) \geqslant \delta\mathcal{R}(x)$ is much less restrictive than it seems at first glance. Let us show that if M is sufficiently large, then $\mathcal{R}(x) = 0$ for most x. Indeed, for every non-Ahlfors point $x \in \mathbb{C}$, one can find a disk $B(x, r)$ such that $\mathcal{H}F(B(x, r)) > Mr$. Using the Vitali covering theorem, we can construct a countable family of *pairwise disjoint* non-Ahlfors disks $B(x_j, r_j)$ such that every non-Ahlfors disk $B(x, r)$ is contained in the union

$$H := \bigcup_j B(x_j, 5r_j).$$

Note that $r_j < \frac{\mathcal{H}F(B(x_j,r_j))}{M}$ and therefore

$$\sum_j r_j < \frac{\mathcal{H}F(\mathbb{C})}{M}.$$

Observing that every term in the sum is not greater than the whole sum, we conclude that

$$\mathcal{H}_{\frac{5\mathcal{H}F(C)}{M}}(H) \leqslant \frac{5\mathcal{H}F(\mathbb{C})}{M},$$

and thereby,

$$\mathcal{H}F(\mathbb{C}\setminus H) = \mathcal{H}(F\setminus H) \geqslant \mathcal{H}_{\frac{5\mathcal{H}F(\mathbb{C})}{M}}(F\setminus H) \geqslant \mathcal{H}_{\frac{5\mathcal{H}F(\mathbb{C})}{M}}(F) - \mathcal{H}_{\frac{5\mathcal{H}F(\mathbb{C})}{M}}(H) \geq$$

$$\mathcal{H}_{\frac{5\mathcal{H}F(\mathbb{C})}{M}}(F) - \frac{5\mathcal{H}F(\mathbb{C})}{M} \to \mathcal{H}(F) = \mathcal{H}F(\mathbb{C})$$

as $M \to +\infty$. Thus $\mathcal{H}F(H) = \mathcal{H}F(\mathbb{C}) - \mathcal{H}F(\mathbb{C}\setminus H) \to 0$ as $M \to +\infty$, proving the claim.

Now define

$$\widetilde{\mathcal{R}}(x) := \operatorname{dist}(x, \mathbb{C}\setminus H).$$

Clearly $\widetilde{\mathcal{R}}$ is a Lipschitz function. Since every non-Ahlfors disk is contained in H, we have $\widetilde{\mathcal{R}}(x) \geqslant \mathcal{R}(x)$. At last $\mathcal{H}F\{x \in \mathbb{C} : \widetilde{\mathcal{R}}(x) > 0\} = \mathcal{H}F(H)$ can be made as small as one wants by choosing the constant M large enough.

X. Localization

Let x_0 be any L^2-Lebesgue point of h with respect to the measure $\mathcal{H}F$ satisfying $h(x_0) \neq 0$. Recall that it means

$$\mathcal{H}F(B(x_0, r)) > 0 \qquad \text{for every } r > 0;$$

$$\frac{1}{\mathcal{H}F(B(x_0, r))} \int_{B(x_0,r)} |h(x) - h(x_0)|^2 \, d\mathcal{H}F(x) \to 0 \qquad \text{as } r \to 0.$$

Since the measure $\mathcal{H}F$ is finite, $\mathcal{H}F$-almost every point $x \in \mathbb{C}$ is a Lebesgue point of a bounded function h (actually this statement is true for *any* $L^2(\mathcal{H}F)$-function h). On the other hand, as we have seen above, $\mathcal{H}F\{x \in \mathbb{C} : h(x) \neq 0\} > 0$. So, the needed point x_0 really exists.

Now choose $0 < \rho < \frac{1}{8}$ so small that

$$\frac{1}{\mathcal{H}F(B(x_0, \rho))} \int_{B(x_0,\rho)} |h(x) - h(x_0)|^2 \, d\mathcal{H}F(x) < \delta^4 |h(x_0)|^2.$$

Choose $M > 1$ so large that for the corresponding exceptional set H, we have

$$\mathcal{H}F(H) \leqslant \frac{\delta}{3} \mathcal{H}F(B(x_0, \rho)).$$

Now let $\rho' < \rho$ be so close to ρ that

$$\mathcal{H}F(B(x_0, \rho) \setminus B(x_0, \rho')) < \frac{\delta}{3} \mathcal{H}F(B(x_0, \rho)).$$

Let

$$\widetilde{\Phi}(x) := \max\{\widetilde{\mathcal{R}}(x), |x - x_0| - \rho'\}.$$

Note that $\widetilde{\Phi}(x)$ is a nonnegative Lipschitz function majorizing the Ahlfors radius $\mathcal{R}(x)$ and that

$$\mathcal{H}F\{x \in B(x_0, \rho) : \widetilde{\Phi}(x) > 0\} \leqslant \frac{2\delta}{3} \mathcal{H}F(B(x_0, \rho)).$$

Define the Borel measure μ by

$$\mu(E) := \mathcal{H}F(E \cap B(x_0, \rho)).$$

Note that for every Lipschitz function $\Theta \geqslant \delta\widetilde{\Phi}$ and for every $x \in \mathbb{C}$ we have

$$K^\sharp_\Theta h(x) := \sup_{\varepsilon > 0} \left| \int_{\mathbb{C} \setminus B(x,\varepsilon)} k_\Theta(x, y) h(y) \, d\mu(y) \right| \leqslant$$

$$\left[(11 + \delta^{-1}) M + \delta^{-1} (\rho - \rho')^{-1} \mathcal{H}F(\mathbb{C}) \right] \|f\|_{L^\infty} =: B |h(x_0)|.$$

Indeed, if we replace $d\mu(y)$ by $d\mathcal{H}F(y)$, we will have the bound $(11 + \delta^{-1}) M \|f\|_{L^\infty}$ for the supremum. The difference between the corresponding integrals does not exceed

$$\int_{\mathbb{C} \setminus B(x_0,\rho)} |k_\Theta(x, y)| \cdot |h(y)| \, d\mathcal{H}F(y) \leqslant \|f\|_{L^\infty} \int_{\mathbb{C} \setminus B(x_0,\rho)} \frac{d\mathcal{H}F(y)}{\Theta(y)} \leqslant \|f\|_{L^\infty} \frac{\mathcal{H}F(\mathbb{C})}{\delta(\rho - \rho')},$$

and we are done.

Now it is time to bring all the information together. Having started with a compact set F of finite Hausdorff measure and positive analytic capacity, we have constructed a bounded Borel measurable function h, a point $x_0 \in \mathbb{C}$ for which $h(x_0) \neq 0$, a measure μ (which is just $\mathcal{H}F$ restricted to some small disk centered at x_0), a large constant M, an open set H, a Lipschitz function $\widetilde{\Phi}$ and a (large) constant B (they are listed in the order one can choose them) such that the following properties hold:

1) Every non-Ahlfors disk is contained in H, in particular, $\mu(B(x,r)) > Mr \implies B(x,r) \subset H$ (recall that $\mu(B(x,r)) \leqslant \mathcal{H}F(B(x,r))$);

2) $h(x) = h(x_0) + g(x)$ with $\int_{\mathbb{C}} |g|^2 d\mu \leqslant \delta^4 \mu(\mathbb{C})$;

3) $\widetilde{\Phi}(x) \geqslant \operatorname{dist}(x, \mathbb{C} \setminus H)$, $\qquad \mu\{x \in \mathbb{C} : \widetilde{\Phi}(x) > 0\} \leqslant \frac{2\delta}{3}\mu(\mathbb{C})$

4) For every Lipschitz function $\Theta \geqslant \delta\widetilde{\Phi}$ and for every point $x \in \mathbb{C}$, one has $K^\sharp_\Theta h(x) \leqslant B|h(x_0)|$.

We recommend the reader to reread this list of objects and their properties attentively several times. They are all completely natural, but a little too many to grasp at first glance.

XI. Construction of perfect hair

Given δ, M, B, h, H and $\widetilde{\Phi}$ as above, let us construct perfect hair. In order not to drag x_0 and $h(x_0)$ along all the time, assume that $x_0 = 0$ and $h(x_0) = 1$. Clearly, the problem is shift-invariant, and we specially wrote all the above conditions in such a way that division of f and h by the same constant would change nothing in them.

First we should construct a perfect dyadic lattice $\mathcal{D}$. Our construction will be surprisingly simple (compared to Guy David's decomposition, say): we will just take the standard dyadic lattice and consider its random shifts.

Pick any point $\omega \in [-\frac{1}{4}, \frac{1}{4})^2$ and take the square $Q^0(\omega) := \omega + [-\frac{1}{2}, \frac{1}{2})^2$ as the "starting" square of the dyadic lattice $\mathcal{D} = \mathcal{D}(\omega)$. Recall that $\operatorname{supp}\mu \subset B(0, \frac{1}{8})$ and therefore $\operatorname{supp}\mu \subset Q^0(\omega)$ for every such ω.

We are going to assign equal probability to every ω; so our probability P will be just 4 times the Lebesgue measure restricted to $[-\frac{1}{4}, \frac{1}{4})^2$.

Once we have fixed the starting square Q^0, we have no choice of how to *position* the smaller squares of $\mathcal{D}$: we just split Q^0 into four equal subsquares (of the same kind $[a,b) \times [c,d)$), then split each new square etc. Nevertheless, we still have the freedom of *how far down to go* at every point. Now we are going to use this freedom.

We will call a square *terminal* in the following two cases: *1) $Q \subset H$; or 2) $\int_Q |g|^2\, d\mu \geqslant \delta^2 \mu(Q)$.* Note that in particular, (2) holds if $\mu(Q) = 0$. If a square is not terminal, we will call it *transit*.

Now start the construction of $\mathcal{D}$ with the square Q^0, which is always transit. It has size (side length) $l(Q^0) = 2^{-0} = 1$. Split it into four equal subsquares. Some of them may be terminal and we will not touch those any more. But we will further split each transit square of size 2^{-1} into four subsquares of size 2^{-2} and so on.

XII. Projections Λ and Δ_Q

Let $\mathcal{D}$ be one of the dyadic lattices constructed above. For a function $\psi \in L^1(\mu)$ and for a square $Q \subset \mathbb{C}$, denote by $\langle\psi\rangle_Q$ the average value of ψ over Q with respect to the measure μ, i.e.,

$$\langle\psi\rangle_Q := \frac{1}{\mu(Q)} \int_Q \psi \, d\mu$$

(of course, $\langle\psi\rangle_Q$ makes sense only for squares Q with $\mu(Q) > 0$).

Put

$$\Lambda\varphi := \frac{\langle\varphi\rangle_{Q^0}}{\langle h\rangle_{Q^0}} h.$$

Clearly, $\Lambda\varphi \in L^2(\mu)$ for all $\varphi \in L^2(\mu)$, and $\Lambda^2 = \Lambda$, i.e., Λ is a projection. Note also, that actually Λ does not depend on the lattice $\mathcal{D}$, because the average is taken over the whole support of the measure μ regardless of the position of the square Q^0.

From now on, we will always denote by Q_j $(j = 1, 2, 3, 4)$ the four subsquares of a square Q enumerated in some "natural order" (to be chosen by the reader). In particular, that means that we will have to give up our idea to denote the squares in two copies $\mathcal{D}_1$ and $\mathcal{D}_2$ of the same random dyadic lattice $\mathcal{D}$ by Q_1 and Q_2, respectively. This is okay, because while above it was important to emphasize the symmetry between $\mathcal{D}_1$ and $\mathcal{D}_2$, below we will start almost every claim with "Assume (for definiteness) that $l(Q) \leqslant l(R)$...".

For every square $Q \in \mathcal{D}^{tr}$, define $\Delta_Q\varphi$ by

$$\Delta_Q\varphi\Big|_{\mathbb{C}\setminus Q} := 0, \qquad \Delta_Q\varphi\Big|_{Q_j} := \begin{cases} \left[\dfrac{\langle\varphi\rangle_{Q_j}}{\langle h\rangle_{Q_j}} - \dfrac{\langle\varphi\rangle_Q}{\langle h\rangle_Q}\right] h, & \text{if } Q_j \text{ is transit;} \\ \varphi - \dfrac{\langle\varphi\rangle_Q}{\langle h\rangle_Q} h, & \text{if } Q_j \text{ is terminal} \end{cases}$$

$(j = 1, 2, 3, 4)$. Observe that for every transit square Q, we have $\mu(Q) > 0$ and

$$\langle h\rangle_Q = 1 + \langle g\rangle_Q; \qquad |\langle g\rangle_Q| \leqslant \sqrt{\langle |g|^2\rangle_Q} \leqslant \delta,$$

so our definition makes sense: no zero can appear in the denominator.

Easy properties of $\Delta_Q\varphi$

For every $\varphi \in L^2(\mu)$ and $Q \in \mathcal{D}^{tr}$, *1)* $\Delta_Q\varphi \in L^2(\mu)$ *; 2)* $\int_{\mathbb{C}} \Delta_Q\varphi \, d\mu = 0$*; 3)* Δ_Q *is a projection, i.e.,* $\Delta_Q^2 = \Delta_Q$*; 4)* $\Delta_Q\Lambda = \Lambda\Delta_Q = 0$*; 5) If* $R \in \mathcal{D}^{tr}$ *and* $R \neq Q$*, then* $\Delta_Q\Delta_R = 0$.

To check these properties is left to the reader as an exercise.

Lemma:

For every $\varphi \in L^2(\mu)$ we have

$$\varphi = \Lambda\varphi + \sum_{Q \in \mathcal{D}^{tr}} \Delta_Q\varphi,$$

the series converges in $L^2(\mu)$ and, moreover,

$$2^{-1}\|\varphi\|^2_{L^2(\mu)} \leqslant \|\Lambda\varphi\|^2_{L^2(\mu)} + \sum_{Q\in\mathcal{D}^{tr}} \|\Delta_Q\varphi\|^2_{L^2(\mu)} \leqslant 2\|\varphi\|^2_{L^2(\mu)}.$$

Proof:

Note first of all that if one understands the sum $\sum_{Q\in\mathcal{D}^{tr}}$ as $\lim_{n\to\infty}\sum_{Q\in\mathcal{D}^{tr}:l(Q)>2^{-n}}$, then for μ-almost every $x\in\mathbb{C}$, one has

$$\varphi(x) = \Lambda\varphi(x) + \sum_{Q\in\mathcal{D}^{tr}} \Delta_Q\varphi(x).$$

Indeed, the claim is obvious if the point x lies in some terminal square. Suppose now that it is not the case. Observe that

$$\Lambda\varphi(x) + \sum_{Q\in\mathcal{D}^{tr}:l(Q)>2^{-n}} \Delta_Q\varphi(x) = \frac{\langle\varphi\rangle_{Q^n}}{\langle h\rangle_{Q^n}} h(x),$$

where Q^n is the dyadic square of size 2^{-n}, containing x. Therefore, the claim is true if

$$\langle\varphi\rangle_{Q^n} \to \varphi(x) \qquad \text{and} \qquad \langle h\rangle_{Q^n} \to h(x) \qquad \text{as } n\to\infty$$

(since for every transit square Q the average $\langle h\rangle_Q$ is close to 1, we surely have $h(x)\neq 0$ for such x). But the exceptional set for any of these conditions has μ-measure 0.

Now let us compare $\Lambda\varphi$ and $\Delta_Q\varphi$ to the corresponding terms in the standard martingale decomposition, i.e., to

$$\widetilde{\Lambda}\varphi := \langle\varphi\rangle_{Q^0}$$

and

$$\widetilde{\Delta}_Q\varphi\big|_{\mathbb{C}\setminus Q} := 0, \qquad \widetilde{\Delta}_Q\varphi\big|_{Q_j} := \begin{cases} \langle\varphi\rangle_{Q_j} - \langle\varphi\rangle_Q, & \text{if } Q_j \text{ is transit;} \\ \varphi - \langle\varphi\rangle_Q, & \text{if } Q_j \text{ is terminal} \end{cases}$$

($j=1,2,3,4$). It is well-known (and easy to prove) that

$$\|\widetilde{\Lambda}\varphi\|^2_{L^2(\mu)} + \sum_{Q\in\mathcal{D}^{tr}} \|\widetilde{\Delta}_Q\varphi\|^2_{L^2(\mu)} = \|\varphi\|^2_{L^2(\mu)}.$$

A direct computation yields

$$\|\widetilde{\Lambda}\varphi\|^2_{L^2(\mu)} = |\langle\varphi\rangle_{Q^0}|^2\mu(Q^0), \qquad \|\Lambda\varphi\|^2_{L^2(\mu)} = \frac{\langle|h|^2\rangle_{Q^0}}{|\langle h\rangle_{Q^0}|^2}|\langle\varphi\rangle_{Q^0}|^2\mu(Q^0),$$

i.e.,

$$\|\Lambda\varphi\|^2_{L^2(\mu)} = \frac{\langle|h|^2\rangle_{Q^0}}{|\langle h\rangle_{Q^0}|^2}\|\widetilde{\Lambda}\varphi\|^2_{L^2(\mu)}.$$

We are going to show that the ratio $\dfrac{\langle |h|^2\rangle_{Q^0}}{|\langle h\rangle_{Q^0}|^2}$ is close to 1. Indeed, we can write

$$\frac{\langle |h|^2\rangle_{Q^0}}{|\langle h\rangle_{Q^0}|^2}-1=\frac{\langle |h|^2\rangle_{Q^0}-|\langle h\rangle_{Q^0}|^2}{|\langle h\rangle_{Q^0}|^2}=\frac{\langle |g|^2\rangle_{Q^0}-|\langle g\rangle_{Q^0}|^2}{|\langle h\rangle_{Q^0}|^2}.$$

Now note that

$$|\langle h\rangle_{Q^0}|\geqslant 1-\langle |g|\rangle_{Q^0}\geqslant 1-\sqrt{\langle |g|^2\rangle_{Q^0}}\geqslant 1-\delta,$$

while the numerator is not less than 0 (Cauchy inequality) and not greater than $\langle |g|^2\rangle_{Q^0}\leqslant\delta^2$. Therefore the whole ratio lies between 0 and $\delta^2(1-\delta)^{-2}\leqslant\delta$. So, we finally get

$$\|\widetilde{\Lambda}\varphi\|^2_{L^2(\mu)}\leqslant\|\Lambda\varphi\|^2_{L^2(\mu)}\leqslant(1+\delta)\|\widetilde{\Lambda}\varphi\|^2_{L^2(\mu)}.$$

As to the terms $\Delta_Q\varphi$, we will represent each of them as the difference $\Delta'_Q\varphi-\dfrac{\langle\varphi\rangle_Q}{\langle h\rangle_Q}h_Q$, where

$$\Delta'_Q\varphi\Big|_{\mathbb{C}\setminus Q}:=0,\qquad \Delta'_Q\varphi\Big|_{Q_j}:=\begin{cases}\dfrac{\langle\varphi\rangle_{Q_j}-\langle\varphi\rangle_Q}{\langle h\rangle_{Q_j}}h, & \text{if } Q_j \text{ is transit;}\\ \varphi-\langle\varphi\rangle_Q, & \text{if } Q_j \text{ is terminal,}\end{cases}$$

and

$$h_Q\Big|_{\mathbb{C}\setminus Q}:=0,\qquad h_Q\Big|_{Q_j}:=\begin{cases}\dfrac{\langle h\rangle_{Q_j}-\langle h\rangle_Q}{\langle h\rangle_{Q_j}}h, & \text{if } Q_j \text{ is transit;}\\ h-\langle h\rangle_Q, & \text{if } Q_j \text{ is terminal}\end{cases}$$

($j=1,2,3,4$). Note that $\Delta'\varphi\equiv\widetilde{\Delta}\varphi$ on $\mathbb{C}\setminus Q$ and on every terminal square Q_j. Also, if Q_j is a transit subsquare of Q, then

$$\int_{Q_j}|\widetilde{\Delta}_Q\varphi|^2d\mu\leqslant\int_{Q_j}|\Delta'_Q\varphi|^2d\mu\leqslant(1+\delta)\int_{Q_j}|\widetilde{\Delta}_Q\varphi|^2d\mu$$

(the reasoning is exactly the same as we had for $\Lambda\varphi$ and $\widetilde{\Lambda}\varphi$). Using the elementary inequality

$$\frac{2}{3}|a|^2-2|b|^2\leqslant|a-b|^2\leqslant\frac{3}{2}|a|^2+3|b|^2\qquad(a,b\in\mathbb{C}),$$

we get

$$\tfrac{2}{3}\|\varphi\|^2_{L^2(\mu)}-2\sigma\leqslant\|\Lambda\varphi\|^2_{L^2(\mu)}+\sum_{Q\in\mathcal{D}^{tr}}\|\Delta_Q\varphi\|^2_{L^2(\mu)}\leqslant\tfrac{3}{2}(1+\delta)\|\varphi\|^2_{L^2(\mu)}+3\sigma,$$

where

$$\sigma:=\sum_{Q\in\mathcal{D}^{tr}}\frac{|\langle\varphi\rangle_Q|^2}{|\langle h\rangle_Q|^2}\|h_Q\|^2_{L^2(\mu)}\leqslant\frac{1+\delta}{(1-\delta)^2}\sum_{Q\in\mathcal{D}^{tr}}|\langle\varphi\rangle_Q|^2\|\widetilde{\Delta}_Q g\|^2_{L^2(\mu)}\leq$$

$$2\sum_{Q\in\mathcal{D}^{tr}}|\langle\varphi\rangle_Q|^2\|\widetilde{\Delta}_Q g\|^2_{L^2(\mu)},$$

because $|\langle h\rangle_Q|\geqslant 1-\delta$; the same reasoning as we used when comparing $\Delta'_Q\varphi$ to $\widetilde{\Delta}_Q\varphi$, allows us to conclude that $\|h_Q\|^2_{L^2(\mu)}\leqslant(1+\delta)\|\widetilde{\Delta}_Q h\|^2_{L^2(\mu)}$, and, at last, $\widetilde{\Delta}_Q h=\widetilde{\Delta}_Q g$.

Now let us remind the reader of the celebrated

Dyadic Carleson Imbedding Theorem

Assume that we have a dyadic lattice $\mathcal{D}$ as above and a family of nonnegative numbers $\{a_Q\}_{Q\in\mathcal{D}}$. Suppose also that for every square $R\in\mathcal{D}$, we have

$$\sum_{Q\in\mathcal{D}:Q\subset R}a_Q\leqslant A\mu(R).$$

Then for every function $\varphi\in L^2(\mu)$ we have

$$\sum_{Q\in\mathcal{D}:\mu(Q)\neq 0}a_Q|\langle\varphi\rangle_Q|^2\leqslant 4A\|\varphi\|^2_{L^2(\mu)}.$$

Now observe that for every transit square $R\in\mathcal{D}$, we have

$$\sum_{Q\in\mathcal{D}^{tr}:Q\subset R}\|\widetilde{\Delta}_Q g\|^2_{L^2(\mu)}=\sum_{Q\in\mathcal{D}^{tr}:Q\subset R}\|\widetilde{\Delta}_Q(g\chi_R)\|^2_{L^2(\mu)}\leqslant\|g\chi_R\|^2_{L^2(\mu)}=\int_R|g|^2d\mu\leqslant\delta^2\mu(R).$$

Thus, applying the Dyadic Carleson Imbedding Theorem to $a_Q=\|\widetilde{\Delta}_Q g\|^2_{L^2(\mu)}$, if Q is transit, and $a_Q=0$, if Q is terminal, we get

$$\sigma\leqslant 8\delta^2\|f\|_{L^2(\mu)}.$$

To finish the proof of the lemma, it remains only to note that

$$\frac{2}{3}-16\delta^2\geqslant\frac{1}{2}\qquad\text{and}\qquad\frac{3}{2}(1+\delta)+24\delta^2\leqslant 2.$$

XIII. Functions $\Phi_{\mathcal{D}}$

Recall that we already have the Lipschitz function $\widetilde{\Phi}$ and that $\widetilde{\Phi}(x)\geqslant\operatorname{dist}(x,\mathbb{C}\setminus H)$. In particular it follows that

$$\widetilde{\Phi}(x)\geqslant\operatorname{dist}(x,\partial Q)\qquad\text{for all } x\in Q,$$

if $Q\in\mathcal{D}^{term}$ and $Q\subset H$.

We would like to extend this property to *all* terminal squares in $\mathcal{D}$. So, let us define

$$\Phi_{\mathcal{D}}(x)=\sup\{\widetilde{\Phi}(x),\operatorname{dist}(x,\mathbb{C}\setminus Q)\,:\,Q\in\mathcal{D}^{term},\int_Q|g|^2d\mu\geqslant\delta^2\mu(Q)\}.$$

Clearly, $\Phi_{\mathcal{D}}$ is Lipschitz, $\Phi_{\mathcal{D}} \geqslant \widetilde{\Phi}$, and $\Phi_{\mathcal{D}}(x) \geqslant \operatorname{dist}(x, \partial Q)$ whenever $x \in Q \in \mathcal{D}^{term}$.

Now note that

$$\mu\{x \in \mathbb{C} : \Phi_{\mathcal{D}}(x) > 0\} \leqslant \mu\{x \in \mathbb{C} : \widetilde{\Phi}(x) > 0\} + \sum_{Q \in \mathcal{D}^{term}, \int_Q |g|^2 \geqslant \delta^2 \mu(Q)} \mu(Q).$$

The list in the end of Section X shows $\mu\{x \in \mathbb{C} : \widetilde{\Phi}(x) > 0\} \leqslant \frac{2\delta}{3}\mu(\mathbb{C})$. On the other hand, the squares in $\mathcal{D}^{term}$ are pairwise disjoint. Therefore the second sum does not exceed $\delta^{-2}\int_{\mathbb{C}} |g|^2 d\mu \leqslant \delta^2 \mu(\mathbb{C})$, and we finally get

$$\mu\{x \in \mathbb{C} : \Phi_{\mathcal{D}}(x) > 0\} \leqslant \left(\tfrac{2\delta}{3} + \delta^2\right)\mu(\mathbb{C}) \leqslant \delta\mu(\mathbb{C}).$$

XIV. Action on good functions

Now let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two dyadic lattices of the above kind. We need to show that for every two good functions $\varphi, \psi \in L^2(\mu)$ (they play the roles of the functions φ_1 and φ_2 in the definition of perfect hair, respectively) and for every Lipschitz function $\Theta \geqslant \delta \max\{\Phi_{\mathcal{D}_1}, \Phi_{\mathcal{D}_2}\}$ satisfying $\inf_{\mathbb{C}} \Theta > 0$, one has

$$|\langle \varphi, K_\Theta \psi \rangle| \leqslant N \|\varphi\|_{L^2(\mu)} \|\psi\|_{L^2(\mu)}.$$

The reader may be surprised by the fact that we are talking about good functions without defining the bad squares first. Actually, to tell the truth, the bad squares are those with which we do not know what to do. Almost all the statements below are very hard or even impossible to prove directly for *arbitrary* squares $Q \in \mathcal{D}_1$, $R \in \mathcal{D}_2$. But they become next to trivial, if we introduce some additional assumptions. All we need to do is to show that all our auxiliary assumptions hold with probability close to 1, and this can be postponed to the very end.

Note first of all, that it is enough to prove the desired inequality for functions φ and ψ such that $\Lambda\varphi = \Lambda\psi = 0$.

Indeed, for any $\varphi \in L^2(\mu)$, we have

$$\|K_\Theta \Lambda\varphi\|_{L^2(\mu)} = \frac{|\langle \varphi \rangle_{Q^0}|}{|\langle h \rangle_{Q^0}|} \|K_\Theta h\|_{L^2(\mu)} \leq$$

$$\frac{1}{1-\delta} |\langle \varphi \rangle_{Q^0}| \cdot \|K_\Theta h\|_{L^\infty(\mu)} \sqrt{\mu(Q^0)} \leqslant 2B |\langle \varphi \rangle_{Q^0}| \sqrt{\mu(Q^0)} \leqslant 2B \|\varphi\|_{L^2(\mu)}.$$

Taking into account that $\langle \varphi, K_\Theta \psi \rangle = -\langle K_\Theta \varphi, \psi \rangle$ for all $\varphi, \psi \in L^2(\mu)$, we get

$$\langle \varphi, K_\Theta \psi \rangle = -\langle K_\Theta \Lambda\varphi, \psi \rangle + \langle \varphi - \Lambda\varphi, K_\Theta \Lambda\psi \rangle + \langle \varphi - \Lambda\varphi, K_\Theta(\psi - \Lambda\psi) \rangle.$$

The first two terms do not exceed $2B\|\varphi\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)}$ and $4B\|\varphi\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)}$, correspondingly (because $\|\varphi - \Lambda\varphi\|_{L^2(\mu)} \leqslant 2\|\varphi\|_{L^2(\mu)}$). Meanwhile, the functions $\varphi' = \varphi - \Lambda\varphi$ and $\psi' = \psi - \Lambda\psi$ clearly satisfy the condition $\Lambda\varphi = \Lambda\psi = 0$ and their $L^2(\mu)$-norms are bounded by $2\|\varphi\|_{L^2(\mu)}$ and $2\|\varphi\|_{L^2(\mu)}$, respectively. So, if we prove the desired inequality for all good φ and ψ satisfying $\Lambda\varphi = \Lambda\psi = 0$ with some constant N_1, then we will get it for two arbitrary good functions with the constant $N = 4N_1 + 6B$.

We would like to write

$$\langle \varphi, K_\Theta \psi\rangle = \sum_{Q\in\mathcal{D}_1^{tr},\, R\in\mathcal{D}_2^{tr}} \langle \Delta_Q \varphi, K_\Theta \Delta_R \psi\rangle.$$

The question arises of why this series converges in any reasonable sense. But let us observe that, since $\inf_{\mathbb{C}} \Theta > 0$, the operator K_Θ is bounded in $L^2(\mu)$ and therefore we can restrict ourselves to the good functions φ and ψ that have only finitely many non-zero terms in their decompositions (clearly, if φ is good, then any partial sum of the series $\Lambda\varphi + \sum_{Q\in\mathcal{D}_1} \Delta_Q \varphi$ is good as well). This not only removes any questions about the convergence, but also allows us to rearrange and to group the terms in the sum in any way we want.

Due to this observation and due to the (anti)symmetry, it is enough to estimate the sum over $Q \in \mathcal{D}_1^{tr}$ and $R \in \mathcal{D}_2^{tr}$, for which $l(Q) \le l(R)$. For the sake of notational simplicity, everywhere below instead of

$$\sum_{Q\in\mathcal{D}_1^{tr},\, Q \text{ is good},\, R\in\mathcal{D}_2^{tr},\, l(Q)\le l(R),\, \text{other conditions}},$$

we will write

$$\sum_{Q,R:\ \text{other conditions}}.$$

Also we will always reduce $\sum_{Q\in\mathcal{D}_1^{tr}:\, Q \text{ is good, other conditions}}$ to $\sum_{Q\,:\,\text{other conditions}}$ and $\sum_{R\in\mathcal{D}_2^{tr}\,:\,\text{other conditions}}$ to $\sum_{R\,:\,\text{other conditions}}$.

Note, that (unless otherwise specified) we will always think that the summation over Q goes only over *good* squares $Q \in \mathcal{D}_1^{tr}$, while the summation over R goes over *all* $R \in \mathcal{D}_2^{tr}$.

Of course, formally it doesn't matter, because, since the functions φ and ψ are good, it is merely a business of adding or omitting several zeros. But it will allow us (and the reader) to see clearly where and what property is used. As the reader might have already guessed, for the sum over pairs Q, R with $l(Q) > l(R)$, this point of view should be changed to the opposite.

Pick some large positive integer m and write

$$\sum_{Q,R} \langle \Delta_Q \varphi, K_\Theta \Delta_R \psi\rangle = \sum_{Q,R: l(Q)\geqslant 2^{-m} l(R)} + \sum_{Q,R: l(Q)< 2^{-m} l(R)} =$$

$$\sum_{\substack{Q,R: l(Q)\geqslant 2^{-m} l(R),\\ \operatorname{dist}(Q,R)\leqslant l(R)}} + \left[\sum_{\substack{Q,R: l(Q)\geqslant 2^{-m} l(R),\\ \operatorname{dist}(Q,R)> l(R)}} + \sum_{\substack{Q,R: l(Q)< 2^{-m} l(R),\\ Q\cap R=\emptyset}} \right] + \sum_{\substack{Q,R: l(Q)< 2^{-m} l(R),\\ Q\cap R\neq\emptyset}}$$

$$=: \sigma_1 + \sigma_2 + \sigma_3.$$

Recall that the kernel k_Θ satisfies the estimates

$$|k_\Theta(x,y)| \leqslant \frac{1}{\max\{|x-y|, \Theta(x), \Theta(y)\}} \qquad \text{and} \qquad |\nabla_x k_\Theta(x,y) \leqslant \frac{4}{|x-y|^2}.$$

The second inequality implies that

$$|k_{\Theta}(x,y) - k_{\Theta}(x',y)| \leqslant \frac{16|x-x'|}{|x-y|^2},$$

provided that $|x-x'| \leqslant \frac{1}{2}|x-y|$. Actually, we do not need the kernel to be that smooth. We will see that the estimate

$$|k_{\Theta}(x,y) - k_{\Theta}(x',y)| \leqslant \frac{A|x-x'|^{\varepsilon}}{|x-y|^{1+\varepsilon}}$$

with some (fixed) $0 < \varepsilon \leqslant 1$ and $0 < A < +\infty$ is sufficient for all our tricks. The reader may ask: "Why introduce a special notation for the parameter, which is actually equal to 1; isn't it merely a generalization for the sake of generalization?" Well, first of all, we want to show that there is nothing very special about the Cauchy kernel $\frac{1}{x-y}$; it can be replaced by any other (antisymmetric) Calderon-Zygmund kernel. And secondly, it will allow the reader to check that our proof works not because of some "magic" numerical identities like $\frac{1}{3}-\frac{1}{2}+\frac{1}{6}=0$, but because we really have found a good way to go around the main drawback of the Haar system: the impossibility to make good estimates near jumps. And once this main drawback is removed, the old-fashioned Haar system becomes more elegant and powerful than any ultramodern and superfamous wavelets.

XV. Estimation of σ_2

Recall that the sum σ_2 is taken over pairs Q, R such that $Q \cap R = \emptyset$. If $l(Q) \geqslant 2^{-m} l(R)$, then the squares not only do not intersect, but are well-separated: $\operatorname{dist}(Q,R) \geqslant l(R)$. We would like to extend this property onto the case $l(Q) < 2^{-m} l(R)$. Though we cannot achieve exactly the same separation by the length of the larger square, we can get as close to it as we want. Namely, for any $\alpha > 0$ and for any $Q \in \mathcal{D}_1$, the probability

$$P_{\mathcal{D}_2}\{\text{there exists } R \in \mathcal{D}_2 : l(R) > 2^m l(Q), R \cap Q = \emptyset \text{ and } \operatorname{dist}(Q,R) < l(Q)^{\alpha} l(R)^{1-\alpha}\}$$

allows an estimate that does not depend on Q and tends to 0 as $m \to \infty$.

We shall need this result for $\alpha = \frac{\varepsilon}{2(1+\varepsilon)}$ ($\frac{1}{4}$ in the case of the Cauchy kernel). We will postpone the proof of this claim to the end of the paper, as we said before; and now let us observe that if we declare the corresponding squares Q bad and if φ is good, then for *every* pair Q, R, participating in σ_2, we have $\operatorname{dist}(Q,R) \geqslant l(Q)^{\alpha} l(R)^{1-\alpha}$.

Define the *long distance* $D(Q,R)$ between the squares Q and R by

$$D(Q,R) = l(Q) + l(R) + \operatorname{dist}(Q,R).$$

Far Interaction Lemma:

Suppose that Q and R are two squares on the complex plane $\mathbb{C}$, such that $l(Q) \leq l(R)$. Let $\varphi_Q, \psi_R \in L^2(\mu)$. Assume that φ_Q vanishes outside Q, ψ_R vanishes outside R; $\int_{\mathbb{C}} \varphi_Q = 0$ and, at last, $\operatorname{dist}(Q, \operatorname{supp} \psi_R) \geqslant l(Q)^{\alpha} l(R)^{1-\alpha}$.

Then

$$|\langle \varphi_Q, K_\Theta \psi_R\rangle| \leqslant 3^{1+\varepsilon} A \frac{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}{D(Q,R)^{1+\varepsilon}} \sqrt{\mu(Q)}\sqrt{\mu(R)} \|\varphi_Q\|_{L^2(\mu)} \|\psi_R\|_{L^2(\mu)}.$$

Remark

Note, that we require only that the support of the function ψ lies far from Q; the squares Q and R themselves may intersect! We will really have such a situation when estimating σ_3.

Proof:

Let x_Q be the center of the square Q. Note that for all $x \in Q$, $y \in \operatorname{supp}\psi_R$, we have

$$|x_Q - y| \geqslant \frac{l(Q)}{2} + \operatorname{dist}(Q, \operatorname{supp}\psi_R) \geqslant \frac{3l(Q)}{2} \geqslant \sqrt{2} l(Q) \geqslant 2|x - x_Q|.$$

Therefore,

$$|\langle \varphi_Q, K_\Theta \psi_R\rangle| = \Big| \iint k_\Theta(x,y)\varphi_Q(x)\psi_R(y)\, d\mu(x)\, d\mu(y)\Big| =$$

$$\Big| \iint [k_\Theta(x,y) - k_\Theta(x_Q,y)]\varphi_Q(x)\psi_R(y)\, d\mu(x)\, d\mu(y)\Big| \leq$$

$$A \frac{l(Q)^\varepsilon}{\operatorname{dist}(Q, \operatorname{supp}\psi_R)^{1+\varepsilon}} \|\varphi_Q\|_{L^1(\mu)} \|\psi_R\|_{L^1(\mu)}.$$

There are two possible cases:

Case 1: $\operatorname{dist}(Q, \operatorname{supp}\psi_R) \geqslant l(R)$

Then

$$D(Q,R) = l(Q) + l(R) + \operatorname{dist}(Q,R) \leqslant 3 \operatorname{dist}(Q, \operatorname{supp}\psi_R)$$

and therefore

$$\frac{l(Q)^\varepsilon}{\operatorname{dist}(Q, \operatorname{supp}\psi_R)^{1+\varepsilon}} \leqslant 3^{1+\varepsilon} \frac{l(Q)^\varepsilon}{\mathcal{D}(Q,R)^{1+\varepsilon}} \leqslant 3^{1+\varepsilon} \frac{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}{D(Q,R)^{1+\varepsilon}}.$$

Case 2: $l(Q)^\alpha l(R)^{1-\alpha} \leqslant \operatorname{dist}(Q, \operatorname{supp}\psi_R) \leqslant l(R)$

Then $D(Q,R) \leqslant 3l(R)$ and we get

$$\frac{l(Q)^\varepsilon}{\operatorname{dist}(Q, \operatorname{supp}\psi_R)^{1+\varepsilon}} \leqslant \frac{l(Q)^\varepsilon}{[l(Q)^\alpha l(R)^{1-\alpha}]^{1+\varepsilon}} = \frac{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}{l(R)^{1+\varepsilon}} \leqslant 3^{1+\varepsilon} \frac{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}{D(Q,R)^{1+\varepsilon}}.$$

Now, to finish the proof of the lemma, it remains only to note that

$$\|\varphi_Q\|_{L^1(\mu)} \leqslant \sqrt{\mu(Q)} \|\varphi_Q\|_{L^2(\mu)} \qquad \text{and} \qquad \|\psi_R\|_{L^1(\mu)} \leqslant \sqrt{\mu(R)} \|\psi_R\|_{L^2(\mu)}.$$

Applying this lemma to $\varphi_Q = \Delta_Q \varphi$ and $\psi_R = \Delta_R \psi$, we obtain

$$|\sigma_2| \leqslant 3^{1+\varepsilon} A \sum_{Q,R} \frac{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}{D(Q,R)^{1+\varepsilon}} \sqrt{\mu(Q)}\sqrt{\mu(R)} \|\Delta_Q \varphi\|_{L^2(\mu)} \|\Delta_R \psi\|_{L^2(\mu)} \qquad (**)$$

We are going to show that the matrix $T_{Q,R}$ defined by

$$T_{Q,R} := \frac{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}{D(Q,R)^{1+\varepsilon}} \sqrt{\mu(Q)}\sqrt{\mu(R)} \qquad (Q \in \mathcal{D}_1^{tr},\ R \in \mathcal{D}_2^{tr},\ l(Q)l(R)\,)$$

generates a bounded operator in l^2.

Lemma:

For any two "sequences" $\{a_Q\}_{Q\in\mathcal{D}_1^{tr}}$ and $\{b_R\}_{R\in\mathcal{D}_2^{tr}}$ of nonnegative numbers, one has

$$\sum_{Q,R} T_{Q,R}a_Q b_R \leqslant \frac{3^{1+\varepsilon}(3+\varepsilon^{-1})M}{1-2^{-\frac{\varepsilon}{2}}}\Big[\sum_Q a_Q^2\Big]^{\frac{1}{2}}\Big[\sum_R b_R^2\Big]^{\frac{1}{2}}.$$

Remark:

Note that $T_{Q,R}$ are defined for all Q, R with $l(Q) \leqslant l(R)$ and that the condition $\operatorname{dist}(Q,R) \geqslant l(Q)^\alpha l(R)^{1-\alpha}$ (or even the condition $Q \cap R = \emptyset$) no longer appears in the summation!

Proof:

Let us "slice" the matrix $T_{Q,R}$ according to the ratio $\frac{l(Q)}{l(R)}$. Namely, let

$$T^{(n)}_{Q,R} = \begin{cases} T_{Q,R}, & \text{if } l(Q) = 2^{-n} l(R); \\ 0, & \text{otherwise} \end{cases}$$

$(n = 0, 1, 2, \dots)$. To prove the lemma, it is enough to show that for every $n \geqslant 0$,

$$\sum_{Q,R} T^{(n)}_{Q,R} a_Q b_R \leqslant 2^{-\frac{\varepsilon}{2}n} 3^{1+\varepsilon}(3+\varepsilon^{-1})M\Big[\sum_Q a_Q^2\Big]^{\frac{1}{2}}\Big[\sum_R b_R^2\Big]^{\frac{1}{2}}.$$

The matrix $\{T^{(n)}_{Q,R}\}$ has a "block" structure: the variables b_R corresponding to the squares $R \in \mathcal{D}_2^{tr}$, for which $l(R) = 2^j$, can interact only with variables a_Q corresponding to the squares $Q \in \mathcal{D}_1^{tr}$, for which $l(Q) = 2^{j-n}$. Thus, to get the desired inequality, it is enough to estimate each block separately, i.e., to demonstrate that

$$\sum_{Q,R\,:\,l(Q)=2^{j-n},\,l(R)=2^j} T^{(n)}_{Q,R} a_Q b_R \leq$$

$$2^{-\frac{\varepsilon}{2}n} 3^{1+\varepsilon}(3+\varepsilon^{-1})M\Big[\sum_{Q\,:\,l(Q)=2^{j-n}} a_Q^2\Big]^{\frac{1}{2}}\Big[\sum_{R\,:\,l(R)=2^j} b_R^2\Big]^{\frac{1}{2}}.$$

Let us introduce the functions

$$F := \sum_{Q\,:\,l(Q)=2^{j-n}} \frac{a_Q}{\sqrt{\mu(Q)}}\chi_Q \qquad \text{and} \qquad G := \sum_{R\,:\,l(R)=2^j} \frac{b_R}{\sqrt{\mu(R)}}\chi_R.$$

Note that the squares of a given size in one dyadic lattice do not intersect, and therefore at each point $x \in \mathbb{C}$, at most one term in the sum can be non-zero. Also observe that

$$\|F\|_{L^2(\mu)} = \Big[\sum_{Q\,:\,l(Q)=2^{j-n}} a_Q^2\Big]^{\frac{1}{2}} \qquad \text{and} \qquad \|G\|_{L^2(\mu)} = \Big[\sum_{R\,:\,l(R)=2^j} b_R^2\Big]^{\frac{1}{2}}.$$

Then the estimate we need can be rewritten as

$$\iint k_{j,n}(x,y)F(x)G(y)\,d\mu(x)\,d\mu(y) \leqslant 2^{-\frac{\varepsilon}{2}n}3^{1+\varepsilon}(3+\varepsilon^{-1})M\|F\|_{L^2(\mu)}\|G\|_{L^2(\mu)},$$

where

$$k_{j,n}(x,y) = \sum_{Q,R\,:\,l(Q)=2^{j-n},l(R)=2^j} \frac{l(Q)^{\frac{\varepsilon}{2}}l(R)^{\frac{\varepsilon}{2}}}{D(Q,R)^{1+\varepsilon}}\chi_Q(x)\chi_R(y).$$

Again, for every pair of points $x, y \in \mathbb{C}$, only one term in the sum can be nonzero. Since $|x-y| + l(R) \leqslant 3D(Q,R)$ for any $x \in Q$, $y \in R$, we obtain

$$k_{j,n}(x,y) = 2^{-\frac{\varepsilon}{2}n}\frac{l(R)^\varepsilon}{D(Q,R)^{1+\varepsilon}} \leqslant 2^{-\frac{\varepsilon}{2}n}3^{1+\varepsilon}\frac{2^{j\varepsilon}}{[2^j+|x-y|]^{1+\varepsilon}} =: 2^{-\frac{\varepsilon}{2}n}3^{1+\varepsilon}k_j(x,y).$$

So, it is enough to check that

$$\iint k_j(x,y)F(x)G(y)\,d\mu(x)\,d\mu(y) \leqslant (3+\varepsilon^{-1})M\|F\|_{L^2(\mu)}\|G\|_{L^2(\mu)}.$$

According to the Schur test, it would suffice to prove that for every $y \in \mathbb{C}$, one has the estimate $\int_{\mathbb{C}} k_j(x,y)\,d\mu(x) \leqslant (3+\varepsilon^{-1})M$ and vice versa (i.e., for every $x \in \mathbb{C}$, one has $\int_{\mathbb{C}} k_j(x,y)\,d\mu(y) \leqslant (3+\varepsilon^{-1})M$). Then the norm of the integral operator with kernel k_j in $L^2(\mu)$ would be bounded by the same constant $(3+\varepsilon^{-1})M$, and the story would be over.

If we assumed a priori that $\mathcal{R}(y) \leqslant 2^{j+1}$, then the needed estimate would be next to trivial: we could write

$$\int_{\mathbb{C}} k_j(x,y)\,d\mu(x) = \int_{B(y,2^{j+1})} k_j(x,y)\,d\mu(x) + \int_{\mathbb{C}\setminus B(y,2^{j+1})} k_j(x,y)\,d\mu(x) \leq$$

$$2^{-j}\mu(B(y,2^{j+1})) + \int_{\mathbb{C}\setminus B(y,2^{j+1})} \frac{2^{j\varepsilon}}{|x-y|^{1+\varepsilon}}\,d\mu(x) \leq$$

$$M\Big(2+1+\int_{2^j}^{+\infty}\frac{2^{j\varepsilon}}{t^{1+\varepsilon}}dt\Big) = (3+\varepsilon^{-1})M$$

(we applied Comparison Lemma to estimate the integral over $\mathbb{C}\setminus B(y,2^{j+1})$, and again we used the possibility to switch from the radius 2^{j+1} to the smaller number 2^j)

The problem is that we cannot guarantee that $\mathcal{R}(y) \leqslant 2^{j+1}$ for *every* $y \in \mathbb{C}$. So, generally speaking, we are unable to show that the integral operator with kernel $k_j(x,y)$ acts in $L^2(\mu)$. But we *do not need* that much! We only need to check that the corresponding bilinear form is bounded on two *given* functions F and G. So, we are not interested in the points $y \in \mathbb{C}$ for which $G(y) = 0$ (or in the points $x \in \mathbb{C}$, for which $F(x) = 0$). But, by definition, G can be non-zero on transit squares in $\mathcal{D}_2$ of size 2^j only. Now let us notice that if $R \in \mathcal{D}_2^{tr}$, then $\mathcal{R}(y) \leqslant 2l(R)$ for every $y \in R$. Indeed, otherwise there exists a non-Ahlfors disk $B(y,r)$ of radius $r > 2l(R)$. But then $R \subset B(y,r) \subset H$, which is impossible for a transit square!

The same reasoning shows that $\mathcal{R}(x) \leqslant 2^{j-n+1} \leqslant 2^{j+1}$ whenever $F(x) \neq 0$, and we are done with $|\sigma_2|$.

Now, we hope, the reader will agree that the decision to declare the squares contained in H terminal was a good one: not only does the fact that the measure μ is not Ahlfors not put us in any real trouble, but we just hardly have a chance to notice this fact at all. Also, it is clear why the squares with large average of $|g|^2$ have been declared terminal: this allowed us to treat h like an accretive function all the time.

But it still remains unexplained why we were so eager to suppress the Cauchy kernel on every terminal square. The answer is in the next two sections.

XVI. Estimation of σ_3

Recall that the sum σ_3 is taken over the pairs Q, R, for which $l(Q) < 2^{-m} l(R)$ and $Q \cap R \neq \emptyset$. We would like to improve this condition to the demand that Q lie "deep inside" one of the four subsquares R_j ($j = 1, 2, 3, 4$).

Define the *skeleton* R of the square R by

$$\mathrm{sk} R := \bigcup_{j=1}^{4} \partial R_j.$$

We will declare a square $Q \in \mathcal{D}_1$ bad if there exists a square $R \in \mathcal{D}_2$ such that $l(R) > 2^m l(Q)$ and $\operatorname{dist}(Q, \mathrm{sk} R) \leqslant 8 l(Q)^{\alpha} l(R)^{1-\alpha}$. Note that any square bad in the sense of the previous section is bad in this new sense as well.

Now, for every good square $Q \in \mathcal{D}_1$, the conditions $l(Q) < 2^{-m} l(R)$ and $Q \cap R \neq \emptyset$ together imply that Q lies inside one of the four subsquares R_j. We will denote this subsquare by R_Q. The sum σ_3 can now be split into

$$\sigma_3^{term} := \sum_{\substack{Q, R \,:\, Q \subset R,\, l(Q) < 2^{-m} l(R), \\ \mathcal{R}_Q \text{ is terminal}}} \langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle$$

and

$$\sigma_3^{tr} := \sum_{\substack{Q, R \,:\, Q \subset R,\, l(Q) < 2^{-m} l(R), \\ \mathcal{R}_Q \text{ is transit}}} \langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle.$$

XVII. Estimation of σ_3^{term}

First of all, write (recall that R_j denote the children of R):

$$\sigma_3^{term} = \sum_{j=1}^{4} \sum_{\substack{Q, R \,:\, l(Q) < 2^{-m} l(R), \\ Q \subset R_j \in \mathcal{D}_2^{term}}} \langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle.$$

Clearly, it is enough to estimate the inner sum for every fixed j. Let us do it for $j = 1$. We have

$$\sum_{\substack{Q, R \,:\, l(Q) < 2^{-m} l(R), \\ Q \subset R_1 \in \mathcal{D}_2^{term}}} \langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle = \sum_{R : R_1 \in \mathcal{D}_2^{term}} \sum_{\substack{Q : l(Q) < 2^{-m} l(R), \\ Q \subset R_1}} \langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle.$$

Roughly speaking, our main idea here is the following. If $R_1 \in \mathcal{D}_2^{term}$, then for all $x \in R_1$, one has

$$\Theta(x) \geqslant \delta \Phi_{\mathcal{D}_2}(x) \geqslant \delta \operatorname{dist}(x, \partial R_1).$$

For the points x that lie in the "central part" of R_1, the right hand side is at least $\frac{\delta(R)}{8}$. Assume that it is so for *every* point $x \in R_1$. Then

$$k_\Theta(x,y) \leqslant \frac{1}{\Theta(x)} \leqslant \frac{8}{\delta(R)} \qquad \text{for all } x \in R_1, y \in \mathbb{C}.$$

Hence

$$|K_\Theta \Delta_R \psi(x)| \leqslant \frac{8\|\Delta_R \psi\|_{L^1(\mu)}}{\delta(R)} \qquad \text{for all } x \in R_1,$$

and therefore

$$\|\chi_{R_1} \cdot K_\Theta \Delta_R \psi\|_{L^2(\mu)} \leqslant 8\|\Delta_R \psi\|_{L^1(\mu)} \frac{\sqrt{\mu(R_1)}}{\delta(R)} \leqslant \frac{8\mu(R)}{\delta(R)} \|\Delta_R \psi\|_{L^2(\mu)} \leqslant \frac{8M}{\delta} \|\Delta_R \psi\|_{L^2(\mu)},$$

because $\mu(R_1) \leqslant \mu(R)$, $\|\Delta_R \psi\|_{L^1(\mu)} \leqslant \sqrt{\mu(R)}\|\Delta_R \psi\|_{L^2(\mu)}$, and $\mu(R) \leqslant Ml(R)$ (otherwise the disk of radius $l(R)$, centered at the same point as R, would be non-Ahlfors, and we would have $R \subset H$, which is impossible).

Now, recalling the remark from Section II, and taking into account that $\Delta_Q \varphi \equiv 0$ outside Q, we get

$$\sum_{Q:\, Q \subset R_1} |\langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle| = \sum_{Q:\, Q \subset R_1} |\langle \Delta_Q \varphi, \chi_{R_1} \cdot K_\Theta \Delta_R \psi \rangle| \leq$$

$$\sqrt{2}\|\chi_{R_1} \cdot K_\Theta \Delta_R \psi\|_{L^2(\mu)} \Big[\sum_{Q:\, Q \subset R_1} \|\Delta_Q \varphi\|^2_{L^2(\mu)} \Big]^{\frac{1}{2}} \leq$$

$$\frac{16M}{\delta} \|\Delta_R \psi\|_{L^2(\mu)} \Big[\sum_{Q:\, Q \subset R_1} \|\Delta_Q \varphi\|^2_{L^2(\mu)} \Big]^{\frac{1}{2}}.$$

So, we obtain

$$\sum_{R:\, R_1 \in \mathcal{D}_2^{term}} \sum_{Q:\, Q \subset R_1} |\langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle| \leqslant$$

$$\frac{16M}{\delta} \sum_{R:\, R_1 \in \mathcal{D}_2^{term}} \|\Delta_R \psi\|_{L^2(\mu)} \Big[\sum_{Q:\, Q \subset R_1} \|\Delta_Q \varphi\|^2_{L^2(\mu)} \Big]^{\frac{1}{2}} \leqslant$$

$$\frac{16M}{\delta} \Big[\sum_{R:\, R_1 \in \mathcal{D}_2^{term}} \|\Delta_R \psi\|^2_{L^2(\mu)} \Big]^{\frac{1}{2}} \Big[\sum_{R:\, R_1 \in \mathcal{D}_2^{term}} \sum_{Q:\, Q \subset R_1} \|\Delta_Q \varphi\|^2_{L^2(\mu)} \Big]^{\frac{1}{2}}.$$

But the terminal squares in $\mathcal{D}_2$ do not intersect! Therefore every $\Delta_Q \varphi$ can appear at most once in the last double sum, and we get the bound

$$\sum_{R:\, R_1 \in \mathcal{D}_2^{term}} \sum_{Q:\, Q \subset R_1} |\langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle| \leqslant$$

$$\frac{16M}{\delta}\Big[\sum_R \|\Delta_R\psi\|^2_{L^2(\mu)}\Big]^{\frac12}\Big[\sum_Q \|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac12} \leqslant \frac{32M}{\delta}\|\varphi\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)}.$$

The problem is that we cannot guarantee the estimate $\Theta(x) \geqslant \frac{\delta(R)}{8}$ for *every* point $x \in R_1$. So, the kernel k_Θ can grow near the boundary. Nevertheless, due to our definition of good squares, we need only to consider the squares $Q \subset R_1$, for which $\operatorname{dist}(Q, \partial R_1) \geqslant 8l(Q)^\alpha l(R)^{1-\alpha}$. So, if such a square Q lies close to the boundary of R_1, the size $l(Q)$ has to be very small and the corresponding function $\Delta_Q\varphi$ should oscillate very fast. We may hope that this fast oscillation will compensate for the growth of the kernel. To show that it is really the case, we need one more standard technical tool.

XVIII. Whitney decomposition

Let S^0 be an arbitrary square on the complex plane $\mathbb{C}$. Consider the standard dyadic lattice starting with the square S^0, and denote by $W(S^0)$ the family of all maximal subsquares S in this lattice, for which $\operatorname{dist}(S, \partial S^0) \geqslant l(S)$ (see Picture 2). The Whitney decomposition $W(S^0)$ has the following remarkable properties: *1) The squares $S \in W(S^0)$ are pairwise disjoint and cover the interior of S^0; 2)* $\operatorname{dist}(S, \partial S^0) = l(S)$ *for every $S \in W(S^0)$; 3) The expanded squares $\widetilde{S} := 2S$ ($S \in W(S^0)$) still lie "deep inside" S, namely,* $\operatorname{dist}(\widetilde{S}, \partial S^0) = \frac{l(S)}{2} = \frac{l(\widetilde{S})}{4}$*, and every point $x \in \mathbb{C}$ belongs to at most* 6 *squares $\widetilde{S}$.* Denote again the center of a square Q by x_Q. For $S \in W(R_1)$ put

$$\psi_{R,S} := \chi_{\widetilde{S}}\,\Delta_R\psi \qquad \text{and} \qquad \widetilde{\psi}_{R,S} := \chi_{R\setminus\widetilde{S}}\,\Delta_R.$$

We have

$$\sum_{\substack{Q:\, l(Q)<2^{-m}l(R),\\ Q\subset R_1}} \langle \Delta_Q\varphi, K_\Theta\Delta_R\psi\rangle = \sum_{S\in W(R_1)} \sum_{\substack{Q:\, l(Q)<2^{-m}l(R),\\ Q\subset R_1,\, x_Q\in S}} \langle \Delta_Q\varphi, K_\Theta\Delta_R\psi\rangle =$$

$$\sum_{S\in W(R_1)} \sum_{\substack{Q:\, l(Q)<2^{-m}l(R),\\ Q\subset R_1,\, x_Q\in S}} \langle \Delta_Q\varphi, K_\Theta\psi_{R,S}\rangle + \sum_{S\in W(R_1)} \sum_{\substack{Q:\, l(Q)<2^{-m}l(R),\\ Q\subset R_1,\, x_Q\in S}} \langle \Delta_Q\varphi, K_\Theta\widetilde{\psi}_{R,S}\rangle.$$

Note now that for every good $Q \subset R_1$ such that $x_Q \in S \in W(R_1)$, one has

$$8l(Q) \leqslant 8l(Q)^\alpha l(R)^{1-\alpha} \leqslant \operatorname{dist}(Q, \partial R_1) \leqslant \operatorname{dist}(x_Q, \partial R_1) \leqslant 2l(S),$$

and therefore

$$\operatorname{dist}(Q, \operatorname{supp}\widetilde{\psi}_{R,S}) \geqslant \operatorname{dist}(Q, \partial\widetilde{S}) \geqslant \frac{l(S)-l(Q)}{2} \geqslant \frac{l(S)}{4} \geqslant l(Q)^\alpha l(R)^{1-\alpha}.$$

Now the Far Interaction Lemma yields

$$|\langle \Delta_Q\varphi, K_\Theta\widetilde{\psi}_{R,S}\rangle| \leqslant 3^{1+\varepsilon}A\frac{l(Q)^{\frac{\varepsilon}{2}}l(R)^{\frac{\varepsilon}{2}}}{D(Q,R)^{1+\varepsilon}}\sqrt{\mu(Q)}\sqrt{\mu(R)}\|\Delta_Q\varphi\|_{L^2(\mu)}\|\widetilde{\psi}_{R,S}\|_{L^2(\mu)}.$$

Taking into account that $\|\widetilde{\psi}_{R,S}\|_{L^2(\mu)} \leqslant \|\Delta_R\psi\|_{L^2(\mu)}$ and summing over all $R \in \mathcal{D}_2^{tr}$, we arrive at the same sum as in the long term interaction of Section XV (actually, we arrive at the part of that sum which *has not been used yet*, but *has already been estimated* there).

So, it remains to find a good upper bound for

$$\sum_{S\in W(R_1)} \sum_{\substack{Q:\, l(Q)<2^{-m}l(R),\\ Q\subset R_1,\, x_Q\in S}} \langle \Delta_Q\varphi, K_\Theta\psi_{R,S}\rangle.$$

Observe once more that the conditions $Q \in \mathcal{D}_1^{tr}$, Q is good, $l(Q) < 2^{-m}l(R)$, $Q \subset R_1$ and $x_Q \in S$ together imply $Q \subset \widetilde{S}$ (as we have seen above, they even imply that Q lies deep inside $\widetilde{S}$). So, it is enough to estimate the sum

$$\sum_{S\in W(R_1)} \sum_{Q:\, Q\subset\widetilde{S},\, x_Q\in S} |\langle \Delta_Q\varphi, K_\Theta\psi_{R,S}\rangle|.$$

Note now that for *every* $x \in \widetilde{S}$, we have

$$\Theta(x) \geqslant \delta\,\mathrm{dist}(x, \partial R_1) \geqslant \frac{\delta(\widetilde{S})}{4}.$$

Recall that the "naive" reasoning from Section XVII could not be used for the whole R_1. But it can be used for $\widetilde{S}$. Repeating our "naive" reasoning from Section XVII for the square $\widetilde{S}$ instead of the whole R_1, we obtain

$$\sum_{Q:\, Q\subset\widetilde{S},\, x_Q\in S} |\langle \Delta_Q\varphi, K_\Theta\psi_{R,S}\rangle| \leqslant \|\chi_{\widetilde{S}}\cdot K_\Theta\psi_{R,S}\|_{L^2(\mu)}\Big[\sum_{Q:\, Q\subset\widetilde{S},\, x_Q\in S} \|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac12} \leqslant$$

$$\frac{4\mu(\widetilde{S})}{\delta(\widetilde{S})}\|\psi_{R,S}\|_{L^2(\mu)}\Big[\sum_{Q:\, Q\subset\widetilde{S},\, x_Q\in S} \|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac12}.$$

We would like to say again that $\mu(\widetilde{S}) \leqslant Ml(\widetilde{S})$. If not, then, of course, we can conclude that $\widetilde{S} \subset H$, but this *does not* yield a contradiction immediately, because $\widetilde{S}$ *is not* a transit square in $\mathcal{D}_2$ (actually, it is not in $\mathcal{D}_2$ at all!). Note, nevertheless, that if we have at least one good square $Q \in \mathcal{D}_1^{tr}$ such that $Q \subset \widetilde{S}$ (otherwise the sum is 0, and we have nothing to worry about), then we can extend the above chain of inclusions to $Q \subset \widetilde{S} \subset H$, which *is* a contradiction! So, as before, despite the fact that we cannot use the Ahlfors condition *whenever we want to*, we can use it *whenever we need to*.

Thus, we get

$$\sum_{S\in W(R_1)} \sum_{\substack{Q:\, Q\subset\widetilde{S},\\ x_Q\in S}} |\langle \Delta_Q\varphi, K_\Theta\psi_{R,S}\rangle| \leqslant \frac{4M}{\delta} \sum_{S\in W(R_1)} \|\psi_{R,S}\|_{L^2(\mu)}\Big[\sum_{\substack{Q:\, Q\subset\widetilde{S},\\ x_Q\in S}} \|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac12} \leqslant$$

$$\frac{4M}{\delta}\Big[\sum_{S\in W(R_1)}\|\psi_{R,S}\|^2_{L^2(\mu)}\Big]^{\frac12}\Big[\sum_{S\in W(R_1)}\ \sum_{Q:\,Q\subset R_1,\,x_Q\in S}\|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac12}$$

(we relaxed the condition $Q\subset\widetilde{S}$ in the last sum to $Q\subset R_1$; it causes no harm now). But

$$\sum_{S\in W(R_1)}\|\psi_{R,S}\|^2_{L^2(\mu)}=\sum_{S\in W(R_1)}\int_{\widetilde{S}}|\Delta_R\psi|^2\,d\mu\leqslant 6\int_{\mathbb{C}}|\Delta_R\psi|^2\,d\mu=6\|\Delta_R\psi\|^2_{L^2(\mu)}$$

(because every point lies in not more than 6 squares $\widetilde{S}$).

Meanwhile,

$$\sum_{S\in W(R_1)}\ \sum_{Q:\,Q\subset R_1,\,x_Q\in S}\|\Delta_Q\varphi\|^2_{L^2(\mu)}=\sum_{Q:\,Q\subset R_1}\|\Delta_Q\varphi\|^2_{L^2(\mu)}.$$

Hence, summing over all $R\in\mathcal{D}_2^{tr}$, for which $R_1\in\mathcal{D}_2^{term}$, we get

$$\sum_{R:\,R_1\in\mathcal{D}_2^{term}}\ \sum_{S\in W(R_1)}\ \sum_{\substack{Q:\,Q\subset\widetilde{S},\\ x_Q\in S}}|\langle\Delta_Q\varphi,K_\Theta\psi_{R,S}\rangle|\leqslant$$

$$\frac{4\sqrt{6}M}{\delta}\sum_{R:\,R_1\in\mathcal{D}_2^{term}}\|\Delta_R\psi\|_{L^2(\mu)}\Big[\sum_{Q:\,Q\subset R_1}\|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac12}\leqslant$$

$$\frac{10M}{\delta}\Big[\sum_{R:\,R_1\in\mathcal{D}_2^{term}}\|\Delta_R\psi\|^2_{L^2(\mu)}\Big]^{\frac12}\Big[\sum_{R:\,R_1\in\mathcal{D}_2^{term}}\ \sum_{Q:\,Q\subset R_1}\|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac12}\leqslant$$

$$\frac{10M}{\delta}\Big[\sum_{R}\|\Delta_R\psi\|^2_{L^2(\mu)}\Big]^{\frac12}\Big[\sum_{Q}\|\Delta_Q\varphi\|_{L^2(\mu)}\Big]^{\frac12}\leqslant\frac{20M}{\delta}\|\varphi\|_{L^2(\mu)}\|\psi\|_{L^2(\mu)},$$

finishing the story with σ_3^{term}.

XIX. Estimation of σ_3^{tr}

Recall that

$$\sigma_3^{tr}=\sum_{\substack{Q,R\,:\,Q\subset R,\,l(Q)<2^{-m}l(R),\\ \mathcal{R}_Q\text{ is transit}}}\langle\Delta_Q\varphi,K_\Theta\Delta_R\psi\rangle.$$

Split every term in the sum as

$$\langle\Delta_Q\varphi,K_\Theta\Delta_R\psi\rangle=\langle\Delta_Q\varphi,K_\Theta(\chi_{R_Q}\Delta_R\psi)\rangle+\langle\Delta_Q\varphi,K_\Theta(\chi_{R\setminus R_Q}\Delta_R\psi)\rangle.$$

Observe that since Q is good, $Q\subset R$ and $l(Q)<2^{-m}l(R)$, we have

$$\operatorname{dist}(Q,\operatorname{supp}\chi_{R\setminus R_Q}\Delta_R\psi)\geqslant\operatorname{dist}(Q,\operatorname{sk}R)\geqslant l(Q)^\alpha l(R)^{1-\alpha}.$$

Using the Far Interaction Lemma and taking into account that the norm $\|\chi_{R\setminus R_Q}\Delta_R\psi\|_{L^2(\mu)}$ does not exceed $\|\Delta_R\psi\|_{L^2(\mu)}$, we conclude that the sum

$$\sum_{\substack{Q,R:\,Q\subset R,\,l(Q)<2^{-m}l(R),\\ \mathcal{R}_Q\text{ is transit}}} |\langle \Delta_Q\varphi, K_\Theta(\chi_{R\setminus R_Q}\Delta_R\psi)\rangle|$$

can be estimated by the sum $(**)$ from Section XV.

Thus, our task is to find a good bound for the sum

$$\sum_{\substack{Q,R:\,Q\subset R,\,l(Q)<2^{-m}l(R),\\ \mathcal{R}_Q\text{ is transit}}} \langle \Delta_Q\varphi, K_\Theta(\chi_{R_Q}\Delta_R\psi)\rangle.$$

Recalling the definition of $\Delta_R\psi$ and recalling that R_Q is a *transit* square, we get

$$\chi_{R_Q}\Delta_R\psi = c_{R,Q}\chi_{R_Q}h,$$

where

$$c_{R,Q} = \frac{\langle\psi\rangle_{R_Q}}{\langle h\rangle_{R_Q}} - \frac{\langle\psi\rangle_R}{\langle h\rangle_R}$$

is a *constant*. So, our sum can be rewritten as

$$\sum_{\substack{Q,R:\,Q\subset R,\,l(Q)<2^{-m}l(R),\\ \mathcal{R}_Q\text{ is transit}}} c_{R,Q}\langle \Delta_Q\varphi, K_\Theta(\chi_{R_Q}h)\rangle.$$

Our next aim will be to extend the function $\chi_{R_Q}h$ to the whole function h in every term (which is exactly the opposite of the idea of the previous section, where, in a similar situation, we tried to "shrink" the function $\Delta_R\psi$ to $\psi_{R,S}$).

Let us observe that

$$\langle \Delta_Q\varphi, K_\Theta(\chi_{\mathbb{C}\setminus R_Q}h)\rangle = \int_{\mathbb{C}\setminus R_Q} k_\Theta(x,y)\Delta_Q\varphi(x)h(y)\,d\mu(x)\,d\mu(y) =$$

$$\int_{\mathbb{C}\setminus R_Q} [k_\Theta(x,y)-k_\Theta(x_Q,y)]\Delta_Q\varphi(x)h(y)\,d\mu(x)\,d\mu(y).$$

Note again that for every $x\in Q$, $y\in\mathbb{C}\setminus R_Q$, we have

$$|x_Q-y| \geqslant \frac{l(Q)}{2} + \operatorname{dist}(Q,\mathbb{C}\setminus R_Q) \geqslant \frac{3l(Q)}{2} \geqslant \sqrt{2}l(Q) \geqslant 2|x-x_Q|.$$

Therefore

$$|k_\Theta(x,y)-k_\Theta(x_Q,y)| \leqslant \frac{A|x-x_Q|^\varepsilon}{|x_Q-y|^{1+\varepsilon}} \leqslant \frac{Al(Q)^\varepsilon}{|x_Q-y|^{1+\varepsilon}},$$

and

$$|\langle \Delta_Q \varphi, K_\Theta(\chi_{\mathbb{C}\setminus R_Q} h)\rangle| \leqslant A l(Q)^\varepsilon \|\Delta_Q \varphi\|_{L^1(\mu)} \int_{\mathbb{C}\setminus R_Q} \frac{|h(y)|\, d\mu(y)}{|x_Q - y|^{1+\varepsilon}}.$$

Now let us consider the sequence of squares $R^{(j)} \in \mathcal{D}_2$, beginning with $R^{(0)} = R_Q$ and gradually ascending ($R^{(j)} \subset R^{(j+1)}$, $l(R^{(j+1)}) = 2l(R^{(j)})$) to the starting square $R^0 = R^{(N)}$ of the lattice $\mathcal{D}_2$. Clearly, all the squares $R^{(j)}$ are transit.

We have

$$\int_{\mathbb{C}\setminus R_Q} \frac{|h(y)|\, d\mu(y)}{|x_Q - y|^{1+\varepsilon}} = \int_{R^0\setminus R_Q} \frac{|h(y)|\, d\mu(y)}{|x_Q - y|^{1+\varepsilon}} = \sum_{j=1}^{N} \int_{R^{(j)}\setminus R^{(j-1)}} \frac{|h(y)|\, d\mu(y)}{|x_Q - y|^{1+\varepsilon}} =: \sum_{j=1}^{N} I_j.$$

Note now that, since Q is good and $l(Q) < 2^{-m} l(R) \leqslant 2^{-m} l(R^{(j)})$ for all $j = 1, \dots, N$, we have

$$\operatorname{dist}(Q, R^{(j)} \setminus R^{(j-1)}) \geqslant \operatorname{dist}(Q, R^{(j)}) \geqslant l(Q)^\alpha l(R^{(j)})^{1-\alpha}.$$

Hence

$$I_j \leqslant \frac{1}{[l(Q)^\alpha l(R^{(j)})^{1-\alpha}]^{1+\varepsilon}} \int_{R^{(j)}} |h|\, d\mu.$$

Recalling that $\alpha = \frac{\varepsilon}{2(1+\varepsilon)}$, we see that the first factor equals $\dfrac{1}{l(Q)^{\frac{\varepsilon}{2}} l(R^{(j)})^{1+\frac{\varepsilon}{2}}}$.

Since $R^{(j)}$ is transit, we have

$$\int_{R^{(j)}} |h|\, d\mu \leqslant \int_{R^{(j)}} (1 + |g|)\, d\mu \leqslant (1+\delta)\mu(R^{(j)}) \leqslant (1+\delta) M l(R^{(j)}).$$

Thus,

$$I_j \leqslant \frac{(1+\delta)M}{l(Q)^{\frac{\varepsilon}{2}} l(R^{(j)})^{\frac{\varepsilon}{2}}} = 2^{-(j-1)\frac{\varepsilon}{2}} \frac{(1+\delta)M}{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}.$$

Summing over $j \geqslant 1$, we get

$$\int_{\mathbb{C}\setminus R_Q} \frac{|h(y)|\, d\mu(y)}{|x_Q - y|^{1+\varepsilon}} = \sum_{j=1}^{N} I_j \leqslant \frac{(1+\delta)M}{1 - 2^{-\frac{\varepsilon}{2}}} \frac{1}{l(Q)^{\frac{\varepsilon}{2}} l(R)^{\frac{\varepsilon}{2}}}.$$

Now let us note that, since $R_Q \in \mathcal{D}_2^{tr}$, we have

$$\|\Delta_R \psi\|^2_{L^2(\mu)} \geqslant \int_{R_Q} |\Delta_R \psi\|^2\, d\mu = |c_{Q,R}|^2 \int_{R_Q} |h|^2\, d\mu \geqslant$$

$$|c_{Q,R}|^2 |\langle h\rangle|^2_{R_Q} \mu(R_Q) \geqslant (1-\delta)^2 |c_{Q,R}|^2 \mu(R_Q).$$

So,

$$|c_{Q,R}| \leqslant \frac{1}{1-\delta} \frac{\|\Delta_R \psi\|_{L^2(\mu)}}{\sqrt{\mu(R_Q)}}.$$

Combining this estimate with the Cauchy inequality $\|\Delta_Q\varphi\|_{L^1(\mu)} \leqslant \sqrt{\mu(Q)}\|\Delta_Q\varphi\|_{L^2(\mu)}$, we finally obtain

$$|\langle\Delta_Q\varphi, K_\Theta(\chi_{\mathbb{C}\setminus R_Q} h)\rangle| \leqslant \frac{(1+\delta)MA}{(1-\delta)(1-2^{-\frac{\varepsilon}{2}})}\left[\frac{l(Q)}{l(R)}\right]^{\frac{\varepsilon}{2}}\sqrt{\frac{\mu(Q)}{\mu(R_Q)}}\|\Delta_Q\varphi\|_{L^2(\mu)}\|\Delta_R\psi\|_{L^2(\mu)}$$

and

$$\sum_{\substack{Q,R:\,Q\subset R,\,l(Q)<2^{-m}l(R),\\ \mathcal{R}_Q \text{ is transit}}} |c_{R,Q}|\cdot|\langle\Delta_Q\varphi, K_\Theta(\chi_{\mathbb{C}\setminus R_Q} h)\rangle| \leqslant$$

$$\frac{(1+\delta)MA}{(1-\delta)(1-2^{-\frac{\varepsilon}{2}})}\sum_{j=1}^{4}\sum_{Q,R:\,Q\subset R_j}\left[\frac{l(Q)}{l(R)}\right]^{\frac{\varepsilon}{2}}\sqrt{\frac{\mu(Q)}{\mu(R_j)}}\|\Delta_Q\varphi\|_{L^2(\mu)}\|\Delta_R\psi\|_{L^2(\mu)}.$$

So, it is enough to demonstrate that, say, the matrix $\{T_{Q,R}\}$ defined by

$$T_{Q,R} := \left[\frac{l(Q)}{l(R)}\right]^{\frac{\varepsilon}{2}}\sqrt{\frac{\mu(Q)}{\mu(R_1)}} \qquad (Q\subset R_1),$$

generates a bounded operator in l^2 in the sense that for every two "sequences" $\{a_Q\}_{Q\in\mathcal{D}_1^{tr}}$ and $\{b_R\}_{R\in\mathcal{D}_2^{tr}}$ of nonnegative numbers, one has

$$\sum_{Q,R:Q\subset R_1} T_{Q,R}a_Qb_R \leqslant \frac{1}{1-2^{-\frac{\varepsilon}{2}}}\left[\sum_Q a_Q^2\right]^{\frac{1}{2}}\left[\sum_R b_R^2\right]^{\frac{1}{2}}.$$

Again let us "slice" the matrix $T_{Q,R}$ according to the ratio $\frac{l(Q)}{l(R)}$. Namely, let

$$T_{Q,R}^{(n)} = \begin{cases} T_{Q,R}, & \text{if } Q\subset R_1,\ l(Q)=2^{-n}l(R);\\ 0, & \text{otherwise}\end{cases}$$

$(n=1,2,\dots)$. It is enough to show that for every $n\geqslant 0$,

$$\sum_{Q,R} T_{Q,R}^{(n)}a_Qb_R \leqslant 2^{-\frac{\varepsilon}{2}n}\left[\sum_Q a_Q^2\right]^{\frac{1}{2}}\left[\sum_R b_R^2\right]^{\frac{1}{2}}.$$

The matrix $\{T_{Q,R}^{(n)}\}$ has a very good "block" structure: every a_Q can interact with *only one* variable b_R. So, it is enough to estimate each block separately, i.e., to show that for every fixed $R\in\mathcal{D}_2^{tr}$,

$$\sum_{Q:\,Q\subset R_1,\,l(Q)=2^{-n}l(R)} 2^{-\frac{\varepsilon}{2}n}\sqrt{\frac{\mu(Q)}{\mu(R_1)}}a_Qb_R \leqslant 2^{-\frac{\varepsilon}{2}n}\left[\sum_Q a_Q^2\right]^{\frac{1}{2}}b_R.$$

But, reducing both parts by the non-essential factor $2^{-\frac{\varepsilon}{2}n}b_R$, we see that this estimate is equivalent to the trivial estimate

$$\sum_{Q:\,Q\subset R_1,\,l(Q)=2^{-n}l(R)}\sqrt{\frac{\mu(Q)}{\mu(R_1)}}a_Q\leqslant\Big[\sum_{Q:\,Q\subset R_1,\,l(Q)=2^{-n}l(R)}\frac{\mu(Q)}{\mu(R_1)}\Big]^{\frac12}\Big[\sum_Q a_Q^2\Big]^{\frac12}\leqslant\Big[\sum_Q a_Q^2\Big]^{\frac12},$$

(since squares $Q\in\mathcal{D}_1$ of fixed size do not intersect, $\sum_{Q:\,Q\subset R_1,\,l(Q)=2^{-n}l(R)}\mu(Q)\leqslant\mu(R_1)$).

So, the extension of $\chi_{R_Q}h$ to the whole h does not cause much harm, and we get the sum

$$\sum_{\substack{Q,R:\,Q\subset R,\,l(Q)<2^{-m}l(R),\\ \mathcal{R}_Q\text{ is transit}}} c_{R,Q}\langle\Delta_Q\varphi,K_\Theta h\rangle$$

to estimate. Note that the inner product $\langle\Delta_Q\varphi,K_\Theta h\rangle$ *does not depend* on R at all, so it seems to be a good idea to sum over R for fixed Q first. Recalling that

$$c_{R,Q}=\frac{\langle\psi\rangle_{R_Q}}{\langle h\rangle_{R_Q}}-\frac{\langle\psi\rangle_R}{\langle h\rangle_R}$$

and that $\Lambda\psi=0\Longleftrightarrow\langle\psi\rangle_{R^0}=0$, we conclude that for every $Q\in\mathcal{D}_1^{tr}$ that really appears in the above sum,

$$\sum_{\substack{R:\,R\supset Q,\,l(R)>2^m l(Q),\\ \mathcal{R}_Q\text{ is transit}}} c_{R,Q}=\frac{\langle\psi\rangle_{R(Q)}}{\langle h\rangle_{R(Q)}},$$

where $R(Q)$ is the smallest *transit* square $R\in\mathcal{D}_2$ containing Q and such that $l(R)\geqslant 2^m l(Q)$. So, we obtain the sum

$$\sum_{Q:\,l(Q)<2^{-m}l(R)}\frac{\langle\psi\rangle_{R(Q)}}{\langle h\rangle_{R(Q)}}\langle\Delta_Q\varphi,K_\Theta h\rangle$$

to take care of.

Actually, the range of summation should be $Q\in\mathcal{D}_1^{tr}$, Q is good (default); there exists a square $R\in\mathcal{D}_2^{tr}$ such that $l(Q)<2^{-m}l(R)$, $Q\subset R$ and R_Q is transit, so the last sum we wrote includes some extra terms compared to the original one, namely, the terms corresponding to the squares Q, for which $R(Q)=R^0$. But first, we remember that $\langle\psi\rangle_{R^0}=0$, and second, now (but not before!) we are going to put the absolute value bars around each term, so we may *add* as many terms as we want; the point is not to *lose* any of them. In this respect everything is obviously fine.

Clearly, the squares with $\|\Delta_Q\varphi\|_{L^2(\mu)}=0$ do not contribute anything to the sum. Also, since $R(Q)$ is transit, $|\langle h\rangle_{R(Q)}|\geqslant 1-\delta$. So, we can write

$$\sum_{Q:\,l(Q)<2^{-m}l(R)}\left|\frac{\langle\psi\rangle_{R(Q)}}{\langle h\rangle_{R(Q)}}\langle\Delta_Q\varphi,K_\Theta h\rangle\right|\leqslant$$

$$\frac{1}{1-\delta}\sum_{\substack{Q:\,l(Q)<2^{-m}l(R),\\ \|\Delta_Q\varphi\|_{L^2(\mu)}>0}} |\langle\psi\rangle_{R(Q)}|\frac{|\langle\Delta_Q\varphi, K_\Theta h\rangle|}{\|\Delta_Q\varphi\|_{L^2(\mu)}}\cdot\|\Delta_Q\varphi\|_{L^2(\mu)} \leqslant$$

$$\frac{1}{1-\delta}\Bigg[\sum_{\substack{Q:\,l(Q)<2^{-m}l(R),\\ \|\Delta_Q\varphi\|_{L^2(\mu)}>0}} |\langle\psi\rangle_{R(Q)}|^2\frac{|\langle\Delta_Q\varphi, K_\Theta h\rangle|^2}{\|\Delta_Q\varphi\|^2_{L^2(\mu)}}\Bigg]^{\frac{1}{2}}\Big[\sum_Q\|\Delta_Q\varphi\|^2_{L^2(\mu)}\Big]^{\frac{1}{2}}.$$

The last factor does not exceed $\sqrt{2}\|\varphi\|_{L^2(\mu)}$. So, it is sufficient to show that the middle factor squared is bounded by some constant times $\|\psi\|^2_{L^2(\mu)}$. Switching to the summation over R, we see that the middle factor squared equals

$$\sum_R|\langle\psi\rangle_R|^2\sum_{Q\in\mathcal{F}(R)}\frac{|\langle\Delta_Q\varphi, K_\Theta h\rangle|^2}{\|\Delta_Q\varphi\|^2_{L^2(\mu)}} =: \sum_R a_R|\langle\psi\rangle_R|^2,$$

where

$$\mathcal{F}(R) := \{Q:\ R(Q)=R,\ \|\Delta_Q\varphi\|_{L^2(\mu)}>0\}\,.$$

So, in order to finish the story with σ_3^{tr}, it is enough to show that the numbers a_R satisfy the Carleson condition. Note that for every $Q\in\mathcal{F}(R)$, one has $Q\subset R$ and that the families $\mathcal{F}(R)$ are pairwise disjoint (one could say much more, but these two trivial observations are the only ones that will matter). Now, for every $S\in\mathcal{D}_2$, we have

$$\sum_{R:\,R\subset S} a_R \leqslant \sum_{\substack{Q:\,Q\subset S,\\ \|\Delta_Q\varphi\|_{L^2(\mu)}>0}}\frac{|\langle\Delta_Q\varphi, K_\Theta h\rangle|^2}{\|\Delta_Q\varphi\|^2_{L^2(\mu)}} = \sum_{\substack{Q:\,Q\subset S\\ \|\Delta_Q\varphi\|_{L^2(\mu)}>0}}\frac{|\langle\Delta_Q\varphi, \chi_S\cdot K_\Theta h\rangle|^2}{\|\Delta_Q\varphi\|^2_{L^2(\mu)}} \leqslant$$

$$2\|\chi_S\cdot K_\Theta h\|^2_{L^2(\mu)} = \int_S|K_\Theta h|^2\,d\mu \leqslant 2B^2\mu(S)$$

(because $\Theta\geqslant\delta\Phi_{\mathcal{D}_2}\geqslant\delta\widetilde{\Phi}$), and we are through.

XX. Estimation of σ_1

Recall that

$$\sigma_1 = \sum_{\substack{Q,R:\,l(Q)\geqslant 2^{-m}l(R),\\ \mathrm{dist}(Q,R)l(R)}}\langle\Delta_Q\varphi, K_\Theta\Delta_R\psi\rangle.$$

We are going to put the absolute value signs around every term and to restore the symmetry between Q and R (so, we will add the corresponding part from the sum over pairs Q,R, for which $l(Q)\geqslant l(R)$). Thus, we have to estimate the sum

$$\sigma_1' = \sum_{\substack{Q\in\mathcal{D}_1^{tr},\,R\in\mathcal{D}_2^{tr}:\\ Q,R\text{ are good},\\ 2^{-m}\leqslant\frac{l(Q)}{l(R)}\leqslant 2^m,\\ \mathrm{dist}(Q,R)\leqslant\max\{l(Q),l(R)\}}}|\langle\Delta_Q\varphi, K_\Theta\Delta_R\psi\rangle|$$

(now all the conditions for the range of summation are written explicitely).

The key observation about this sum is that every square Q can interact with not more than $2^{2m}(4\cdot 2^m+1)^2(2m+1)$ squares R and vice versa (the estimate is quite rough, of course, and is obtained as follows: one has $2m+1$ possible values for $l(R)$; once the size $l(R)\in[2^{-m}l(Q),2^m l(Q)]$ is fixed, the corresponding squares R are contained in the square of size $(4\cdot 2^m+1)l(Q)$, centered at the same point as Q, are pairwise disjoint, and the area of each of them is not less than $2^{-2m}l(Q)^2$). Therefore, it is enough to show that for some large constant $U>0$, not depending on φ, ψ and Θ, one has

$$|\langle\Delta_Q\varphi,K_\Theta\Delta_R\psi\rangle|\leqslant U\|\Delta_Q\varphi\|_{L^2(\mu)}\|\Delta_R\psi\|_{L^2(\mu)},$$

provided that $Q\in\mathcal{D}_1^{tr}$, $R\in\mathcal{D}_2^{tr}$, Q,R are good, $2^{-m}\leqslant\frac{l(Q)}{l(R)}\leqslant 2^m$ and $\operatorname{dist}(Q,R)\leqslant\max\{l(Q),l(R)\}$.

XXI. Negligible contours

Let G be a contour on the complex plane $\mathbb{C}$. Let $\widetilde{M}$ be some large positive number. We will call G *negligible* (the full name should be $\widetilde{M}$-negligible with respect to the measure μ), if for every $r>0$,

$$\mu\{x\in\mathbb{C}\,:\,\operatorname{dist}(x,G)\leqslant r\}\leqslant\widetilde{M}r.$$

Lemma: Let G be a negligible contour splitting the complex plane $\mathbb{C}$ into two (open) regions Ω_1 and Ω_2. Then for any two functions $\eta_1,\eta_2\in L^2(\mu)$ such that η_j vanishes outside $\Omega_j\cup G$, one has

$$|\langle\eta_1,K_\Theta\eta_2\rangle|\leqslant 4\widetilde{M}\|\eta_1\|_{L^2(\mu)}\|\eta_2\|_{L^2(\mu)}.$$

Proof: Note that the condition that G is negligible immediately implies that $\mu(G)=0$. So, we may assume that η_j vanishes outside Ω_j. We have

$$|\langle\eta_1,K_\Theta\eta_2\rangle|\leqslant\iint|k_\Theta(x_1,x_2)|\cdot|\eta_1(x_1)|\cdot|\eta_2(x_2)|\,d\mu(x_1)\,d\mu(x_2).$$

Clearly, the integrand can be non-zero only if $x_1\in\Omega_1$ and $x_2\in\Omega_2$. According to the Schur test (full L^2-version), it is enough to find a function $\lambda:\mathbb{C}\setminus G\to(0,+\infty)$, such that

$$\int_{\Omega_1}|k_\Theta(x_1,x_2)|\lambda(x_1)\,d\mu(x_1)\leqslant 4\widetilde{M}\lambda(x_2)\qquad\text{for every } x_2\in\Omega_2,$$

and vice versa, i.e.,

$$\int_{\Omega_2}|k_\Theta(x_1,x_2)|\lambda(x_2)\,d\mu(x_2)\leqslant 4\widetilde{M}\lambda(x_1)\qquad\text{for every } x_1\in\Omega_1.$$

We will check that these properties hold for

$$\lambda(x)=\frac{1}{\sqrt{\operatorname{dist}(x,G)}}\,.$$

Indeed, for $x_1 \in \Omega_1$ and $x_2 \in \Omega_2$, we have

$$|k_\Theta(x_1, x_2)| \leqslant \frac{1}{|x_1 - x_2|} \leqslant \frac{1}{\max\{\operatorname{dist}(x_1, G), \operatorname{dist}(x_2, G)\}} .$$

Thus, according to the Comparison lemma,

$$\int_{\Omega_1} |k_\Theta(x_1, x_2)| \lambda(x_1)\, d\mu(x_1) \leqslant \int_{\Omega_1} \frac{1}{\max\{\operatorname{dist}(x_1, G), \operatorname{dist}(x_2, G)\}} \frac{1}{\sqrt{\operatorname{dist}(x_1, G)}} d\mu(x_1) \leqslant$$

$$\widetilde{M} \int_0^{+\infty} \frac{1}{\max\{t, \operatorname{dist}(x_2, G)\}} \frac{1}{\sqrt{t}} dt = \frac{4\widetilde{M}}{\sqrt{\operatorname{dist}(x_2, G)}} = 4\widetilde{M}\lambda(x_2).$$

Now observe that

$$\langle \Delta_Q \varphi, K_\Theta \Delta_R \psi \rangle = \sum_{i,j=1}^{4} \langle \varphi_Q^{(i)}, K_\Theta \psi_R^{(j)} \rangle,$$

where $\varphi_Q^{(i)} := \chi_{Q_i} \Delta_Q \varphi$, and $\psi_R^{(j)} := \chi_{R_j} \Delta_R \psi$.

Assume that the boundaries of all the subsquares Q_i and R_j are $\widetilde{M}$-negligible contours. Then it makes sense to write

$$\langle \varphi_Q^{(i)}, K_\Theta \psi_R^{(j)} \rangle =$$

$$\langle \chi_{Q_i \setminus R_j} \cdot \varphi_Q^{(i)}, K_\Theta \psi_R^{(j)} \rangle + \langle \chi_{Q_i \cap R_j} \cdot \varphi_Q^{(i)}, K_\Theta (\chi_{R_j \setminus Q_i} \cdot \psi_R^{(j)}) \rangle +$$

$$\langle \chi_{Q_i \cap R_j} \cdot \varphi_Q^{(i)}, K_\Theta (\chi_{Q_i \cap R_j} \cdot \psi_R^{(j)}) \rangle.$$

In the first two terms the supports of the functions are separated by negligible contours (∂R_j and ∂Q_i, respectively). So, the corresponding inner products are bounded by

$$4\widetilde{M} \|\varphi_Q^{(i)}\|_{L^2(\mu)} \|\psi_R^{(j)}\|_{L^2(\mu)} \leqslant 4\widetilde{M} \|\Delta_Q \varphi\|_{L^2(\mu)} \|\Delta_R \psi\|_{L^2(\mu)} .$$

As to the inner product $\langle \chi_{Q_i \cap R_j} \cdot \varphi_Q^{(i)}, K_\Theta (\chi_{Q_i \cap R_j} \cdot \psi_R^{(j)}) \rangle$, there are two possibilities:

Case 1: one of the squares (say, Q_i) is terminal
Then we have the estimate

$$|k_\Theta(x, y)| \leqslant \frac{1}{\delta \max\{\operatorname{dist}(x, \partial Q_i), \operatorname{dist}(y, \partial Q_i)\}}$$

for all $x, y \in Q_i \cap R_j$ and, repeating our above reasoning with the Schur test, we obtain

$$|\langle \chi_{Q_i \cap R_j} \cdot \varphi_Q^{(i)}, K_\Theta (\chi_{Q_i \cap R_j} \cdot \psi_R^{(j)}) \rangle| \leqslant$$

$$\frac{4\widetilde{M}}{\delta} \|\varphi_Q^{(i)}\|_{L^2(\mu)} \|\psi_R^{(j)}\|_{L^2(\mu)} \leqslant \frac{4\widetilde{M}}{\delta} \|\Delta_Q \varphi\|_{L^2(\mu)} \|\Delta_R \psi\|_{L^2(\mu)}.$$

Case 2: both squares Q_i and R_j are transit

Then both functions $\chi_{Q_i\cap R_j}\cdot\varphi_Q^{(i)}$ and $\chi_{Q_i\cap R_j}\cdot\psi_R^{(j)}$ are constant multiples of *the same* function $\eta := \chi_{Q_i\cap R_j}\cdot h$. But the kernel k_Θ is antisymmetric, and therefore $\langle \eta, K_\Theta\eta\rangle = 0$.

What if the boundary of some square Q_i (or R_j) is not negligible? We do not know how to get a good estimate in this case; instead, we will try to rule it out by declaring the corresponding squares bad. But we should be very careful here: the temptation to declare a square $Q\in\mathcal{D}_1^{tr}$ bad if ∂Q is not negligible should be severely suppressed, because, as we remember, "badness" of the square Q should depend rather on $\mathcal{D}_2$, than on Q itself. So, we are going to use a little bit less straightforward definition.

Namely, we will call a transit square $Q\in\mathcal{D}_1$ bad if there exists a transit square $R\in\mathcal{D}_2$ such that $2^{-m}l(Q)\leqslant l(R)\leqslant 2^m l(Q)$, $\operatorname{dist}(R,Q)\leqslant 2^m l(Q)$ and for at least one $j=1,2,3,4$, the boundary ∂R_j is not $\widetilde{M}$-negligible (we do not care about the terminal squares, so let them all be "good by the definition"). Then for every pair of squares Q,R appearing in the sum σ_1', the assumption that Q is good allows to conclude that all the four subsquares R_j of the square R are negligible and vice versa! Now it remains only to show that we can choose m and $\widetilde{M}$ (in this order) so that $P_{\mathcal{D}_2}\{Q \text{ is bad}\}\leqslant\delta$ for every $Q\in\mathcal{D}_1^{tr}$.

XXII. Estimation of probability

Let $Q\in\mathcal{D}_1^{tr}$. Consider the "extended lattice"

$$\widetilde{\mathcal{D}}_2=\widetilde{\mathcal{D}}_2(\omega)=\left\{\omega+\left[\tfrac{j}{2^n},\,\tfrac{j+1}{2^n}\right)\times\left[\tfrac{k}{2^n},\,\tfrac{k+1}{2^n}\right)\ :\ j,k,n\in\mathbb{Z}, n\geqslant 1\right\}.$$

Clearly, $\widetilde{D}_2$ contains every square $R\in\mathcal{D}_2$ of size $\frac12$ or less. Note that when ω runs over $\left[-\frac14,\,\frac14\right)\times\left[-\frac14,\,\frac14\right)$, the lattice $\widetilde{D}_2$ runs over its whole period.

Starting now, we will declare a square $Q\in\mathcal{D}_1^{tr}$ bad if either

1) *there exists a square* $R\in\widetilde{\mathcal{D}}_2$ *such that* $\operatorname{dist}(Q,\partial R)\leqslant 16 l(Q)^\alpha l(R)^{1-\alpha}$ *and* $l(R)\geqslant 2^m l(Q)$, or **2)** *there exists a square* $R\in\widetilde{\mathcal{D}}_2$ *such that* $R\subset(4\cdot 2^m+1)Q$, $l(R)\geqslant 2^{-(m+1)}l(Q)$ *and* ∂R *is not* $\widetilde{M}$*-negligible.* We leave it to the reader to check that every square Q bad in the sense of Section XV, XVI or XXI is bad according to this new definition as well.

Choice of m

Fix $k\geqslant m$. Let us estimate the probability that there exists a square $R\in\widetilde{\mathcal{D}}_2$ of size $l(R)=2^k l(Q)$ such that $\operatorname{dist}(Q,\partial R)\leqslant 16 l(Q)^\alpha l(R)^{1-\alpha}$. Since the lattice $\widetilde{\mathcal{D}}_2$ runs over its whole period, we can find this probability exactly: it equals to the ratio of the area of the dashed rim on Picture 3 to the area of the whole square with side $2^k l(Q)$ (just look at where the center x_Q should lie with respect to the lattice $\mathcal{D}_2$). Observing that

$$16 l(Q)^\alpha l(R)^{1-\alpha}+\tfrac{l(Q)}{2}\leqslant 17 l(Q)^\alpha l(R)^{1-\alpha},$$

we conclude that this ratio is less than $68\left[\frac{l(Q)}{l(R)}\right]^\alpha = 68\cdot 2^{-k\alpha}$.

Therefore the probability that the square Q is bad according to the first part of our definition does not exceed

$$68\sum_{k=m}^\infty 2^{-k\alpha}=\frac{68\cdot 2^{-m\alpha}}{1-2^{-\alpha}}\leqslant\frac{\delta}{3},$$

provided that m is taken large enough.

Choice of $\widetilde{M}$

Now let us look at how large the probability that Q is bad according to the second part of our definition may be. Recall that ∂R is $\widetilde{M}$-negligible if $\mu\{x \in \mathbb{C} : \operatorname{dist}(x, \partial R) \leqslant r\} \leqslant \widetilde{M}r$ for all $r > 0$. Note first of all, that we do not have any trouble with $r \geqslant l(Q)$. Indeed, since $R \subset (4 \cdot 2^m + 1)Q$, we have

$$\{x \in \mathbb{C} : \operatorname{dist}(x, \partial R) \leqslant r\} \subset B(x_Q, (4 \cdot 2^m + 1)l(Q) + r)) \subset B(x_Q, (4 \cdot 2^m + 2)r).$$

But $\mu(B(x_Q, (4 \cdot 2^m + 2)r) \leqslant (4 \cdot 2^m + 2)Mr$, because Q is a transit square, $r \geqslant l(Q)$ and therefore $\mathcal{R}(x_Q) \leqslant l(Q) \leqslant r < (4 \cdot 2^m + 2)r$. So, everything is okay with such r, provided that $\widetilde{M} \geqslant (4 \cdot 2^m + 2)M$.

Now observe that for $r < l(Q)$ we have

$$\{x \in \mathbb{C} : \operatorname{dist}(x, \partial R) \leqslant r\} \subset B(x_Q, (4 \cdot 2^m + 2)l(Q)\,).$$

So, the part of the measure μ that lies outside the disk $B(x_Q, (4 \cdot 2^m + 2)l(Q)\,)$ does not matter and we can replace the whole measure μ by its restriction $\widetilde{\mu}$ to this disk, defined as

$$\widetilde{\mu}(E) := \mu(E \cap B(x_Q, (4 \cdot 2^m + 2)l(Q)\,)\,).$$

Though we do not know *much* about $\widetilde{\mu}$, there is one thing we can say for certain:

$$\widetilde{\mu}(\mathbb{C}) \leqslant (4 \cdot 2^m + 2)Ml(Q);$$

and this will be enough for us.

Consider the grid $\mathcal{L} = \mathcal{L}(\omega)$ consisting of all vertical lines serving as boundaries of squares in $\widetilde{\mathcal{D}}_2$ of size $2^{-(m+1)}l(Q)$. We are going to show that if $\widetilde{M}$ is sufficiently large, then, with probability $1 - \frac{\delta}{3}$ or more, this entire grid is $\frac{\widetilde{M}}{2}$-negligible with respect to the measure $\widetilde{\mu}$. Of course (together with the same estimate for horizontal lines), this will imply that the probability that the square Q is bad according to the second part of our definition does not exceed $\frac{2\delta}{3}$, finishing the story.

Note that the grid $\mathcal{L}$ runs (several times) over its whole period when ω runs over $\left[-\frac{1}{4}, \frac{1}{4}\right) \times \left[-\frac{1}{4}, \frac{1}{4}\right)$. So, we can change the random parameter ω to another random parameter $\tau \in [0, 2^{-(m+1)}l(Q)\,)$ (which is just the real part of $\omega \mod 2^{-(m+1)}l(Q)$, of course) and reformulate our problem as the following: we should demonstrate that the one-dimensional Lebesgue measure of such $\tau \in [0, 2^{-(m+1)}l(Q)\,)$ that the grid $\mathcal{L}(\tau)$ consisting of all vertical lines intersecting the real axis at the points of the kind $\tau + \frac{k}{2^{m+1}}$, $k \in \mathbb{Z}$, is not $\frac{\widetilde{M}}{2}$-negligible with respect to the measure $\widetilde{\mu}$, does not exceed $\frac{\delta}{3}2^{-(m+1)}l(Q)$.

Consider the $2^{-(m+1)}l(Q)$-periodic sweeping ν of the measure $\widetilde{\mu}$, i.e. the measure defined on Borel subsets E of the real line $\mathbb{R}$ by

$$\nu(E) = \widetilde{\mu}\Big(\bigcup_{k \in \mathbb{Z}} \big(k \cdot 2^{-(m+1)}l(Q) + E\big) \times \mathbb{R}\Big).$$

Note that $\mathcal{L}(\tau)$ is not $\frac{\widetilde{M}}{2}$-negligible if and only if $\mathcal{M}\nu(\tau) > \frac{\widetilde{M}}{2}$, where

$$\mathcal{M}\nu(\tau) = \sup_{r>0} \frac{\nu([\tau - r, \tau + r])}{2r}$$

is the Hardy-Littlewood maximal function. But the standard estimate for the maximal function of a periodic measure yields

$$m_1\{\tau \in [0, 2^{-(m+1)}l(Q)) \,:\, \mathcal{M}\nu(\tau) > \frac{\widetilde{M}}{2}\} \leqslant \frac{4\nu([0, 2^{-(m+1)}l(Q))\,)}{\widetilde{M}} =$$

$$\frac{4\widetilde{\mu}(\mathbb{C})}{\widetilde{M}} \leqslant \frac{4(4\cdot 2^m + 2)M}{\widetilde{M}} l(Q).$$

So, if $\widetilde{M} \geqslant 12\delta^{-1}(4\cdot 2^m + 2)M$, we are okay.

XXIII. Quantitative pulling ourselves up by the hair

We are going to present the succession of "fancy" Tb theorems in the nonhomogeneous setting.

The first one is the least "fancy" because b will be accretive in it, but it solves a problem of P. Mattila about an analytic characterization of Besicovitch rectifiable sets.

The middle one is the theorem proved in the previous sections; it gives an alternative proof of the result of Guy David [D1], thus solving the analytic part of Vitushkin's conjecture.

The last one—and the most difficult—gives a quantitative information in the solution of Vitushkin's conjecture. Namely, given a set E of positive analytic capacity γ and length M, this last theorem allows us to say quantitatively what portion of the length is rectifiable, and "how" rectifiable it is.

In what follows μ is a positive measure on $\mathbb{C}$ satisfying the following *non-uniform linear growth condition*:

$$\limsup_{r\to 0} \frac{\mu(B(x,r))}{r} < \infty \text{ for } \mu \text{ a.e } x\,.$$

The truncated Cauchy integral is

$$(C_\mu^\varepsilon b)(\zeta) = \int_{|z-\zeta|>\varepsilon} \frac{b(z)\, d\,\mu(z)}{\zeta - z}\,.$$

The maximal Cauchy integral is

$$(C_\mu^* b)(\zeta) = \sup_{\varepsilon>0} |(C_\varepsilon^\mu b)(\zeta)|\,.$$

Recall that for any 1-Lipschitz function Φ on $\mathbb{C}$ the following Calderón-Zygmund kernel was introduced

$$k_\Phi(x,y) = \frac{\overline{x-y}}{|x-y|^2 + \Phi(x)\Phi(y)}$$

and let K_Φ be the canonical Calderón-Zygmund operator with this antisymmetric kernel. Recall that

$$k_\Phi(x,y) \le \min[\frac{1}{\Phi(x)}, \frac{1}{\Phi(y)}]\,.$$

Consider another truncation of the Cauchy integral:

$$(C_\Phi b)(\zeta) = \int_{|z-\zeta|\ge\Phi(\zeta)} \frac{b(z)\,d\,\mu(z)}{\zeta - z}\,.$$

As usual M_1 denotes the following maximal function

$$(M_1 f)(\zeta) = \sup_{r>0} \frac{1}{r}\int_{B(\zeta,r)} |f(z)|\,d\,\mu(z)\,.$$

Lemma 1:

$$|(K_\Phi f)(x) - (C_\Phi f)(x)| \leqslant A(M_1 f)(x)\,.$$

Proof Fix x and consider the absolute value of the difference of the kernels. For $y \in B(x,\Phi(x))$ it is at most $\frac{1}{\Phi(x)}$. For y such that $|y-x| \geqslant \Phi(x)$ it is

$$\leqslant \frac{(\Phi(x)\Phi(y)|x-y|}{[|x-y|^2 + \Phi(x)\Phi(y)]|x-y|^2} \leqslant |k_\Phi|\frac{\Phi(x)\Phi(y)}{|x-y|^2} \leqslant \frac{\Phi(x)}{|x-y|^2}\,.$$

Splitting $\{y : |y-x| \geqslant \Phi(x)\}$ into annuli $\{y : 2^{k+1}\Phi(x) > |y-x| \geqslant 2^k\Phi(x)\}$ finishes the proof.

Recall that we have assumption of non-uniform linear growth on μ.

Let us also normalize μ and think (if otherwise not stated) that $\|\mu\| = 1$. Recall that M-non-Ahlfors disc is a $B(x,R)$ such that

$$\mu(B(x,R)) \geqslant MR,\; x \in \operatorname{supp}\mu\,.$$

In this case the point x is called an M-non-Ahlfors point.

Lemma 2: There exists $\varepsilon = \varepsilon(M)$, $\varepsilon \to 0$ if $M \to \infty$, such that the union of all M-non-Ahlfors discs has μ-measure at most ε.

Proof. It follows from non-uniform linear growth condition that

$$\mu\{x \in \operatorname{supp}\mu : \sup_r \frac{\mu(B(x,r))}{r} \geqslant \sqrt{M}\} = \delta(M) \to 0,\text{ when } M \to \infty\,.$$

Denote this set by G_M. For $x \in \operatorname{supp}\mu \setminus G_M$ we choose the maximal M-non-Ahlfors disc centered at x (if any). Their union will be called O. By Vitali's lemma, O is covered by $\cup B(x_j, 5r_j)$, where $x_j \in \operatorname{supp}\mu \setminus G_M$, and $B(x_j,r_j)$ are disjoint and M-non-Ahlfors. Thus,

$$\Sigma\, r_i \leqslant \frac{1}{M}\Sigma\,\mu(B(x_i,r_i)) \leqslant \frac{1}{M}$$

On the other hand, $\mu(B(x_j,5r_j)) \leqslant 5\sqrt{M}\,r_j$. Thus,

$$\mu(O) \leqslant \Sigma\, \mu(B(x_j, 5r_j)) \leqslant 5\sqrt{M}\Sigma\, r_j \leqslant \frac{5}{\sqrt{M}}$$

M-non-Ahlfors points can be only in $O \cup G_M$. So we see that

$$\mu(\text{M-non-Ahlfors points}) \to 0, \text{ when } M \to \infty$$

But we want a bit more—the smallness of measure of the union of all M-non-Ahlfors discs. To get this, consider points in G_M, and consider maximal M-non-Ahlfors disc centered at each of them. Call their union G. The set G is covered by $\cup_j B(y_j, 5R_j)$, where $B(y_j, R_j)$ are disjoint M-non-Ahlfors discs. Consider $y \in B(y_j, 5R_j)$. Then $\frac{\mu(B(y,10R_j))}{10R_j} \geqslant \frac{\mu(B(y_j,R_j))}{10R_j} \geqslant \frac{\sqrt{M}}{10}$. In our notations this means that $y \in G_{M/100}$. Thus, $G \subset G_{M/100}$. So $\mu(G)$ is small if M is large. The lemma is proved.

Lemma 3: Let $\|\mu\| = 1$, let μ be a positive measure with non-uniform linear growth, and let $H = H_M$ be the union of all M-non-Ahlfors discs. Let Φ be a 1-Lipschitz function such that $\Phi(x) \geqslant \operatorname{dist}(x, \mathbb{C} \setminus H)$. Then K_Φ and C_Φ are bounded or unbounded simultaneously on $L^2(\mu)$.

Proof. In Lemma 1 we saw that $|(K_\Phi - C_\Phi)(f)(x)| \leqslant A(M_1 f)(x)$. Actually, the proof says more, namely

$$|(K_\Phi - C_\Phi)(f)(x)| \leqslant A(M_{1,\Phi} f)(x) := A \sup_{r \geqslant \Phi(x)} \frac{1}{r} \int_{B(x,r)} |f(y)|\, d\,\mu(y)\,.$$

But for $r \geqslant \Phi(x)$ we have $\mu(B(x,r)) \leqslant Mr$. Therefore,

$$(M_{1,\Phi} f)(x) \leqslant M \sup \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y)|\, d\,\mu(y)\,.$$

It is well-known that this maximal operator is bounded in $L^2(\mu)$. Lemma is proved.

Now we are ready to present several conditions for K_Φ (Φ is a 1-Lipschitz function) to be bounded on $L^2(\mu)$. While doing that we are interested in such Φ's that $F_\Phi := \{x \in \mathbb{C} : \Phi(x) = 0\}$ has positive measure (or, if circumstances permit, even measure close to 1). This interest is easy to explain: for such Φ we have

$$(K_\Phi\, f,\, g) = (Cf, g) \qquad (*)$$

for f, g supported on F_Φ. And after all, we are interested in estimates of the Cauchy operator C.

Theorem 1: Let μ be a measure with non-uniform linear growth, let H_M be the union of all M-non-Ahlfors discs,

$$\Phi(x) \geqslant \operatorname{dist}(x, \mathbb{C} \setminus H_M)$$

and let Φ be a 1-Lipschitz function. Consider $(K_{\Phi,\varepsilon} f)(x) := \int_{|y-x| \geqslant \varepsilon} k_\Phi(x,y) f(y)\, d\,\mu(y)$. If there exists a constant B such that

$$(K^*_\Phi 1)(x) := \sup_{\varepsilon > 0} |(K_{\Phi,\varepsilon} 1)(x)| \leqslant B < \infty \text{ for } \mu \text{ a. e. } x$$

then

$$\|K_\Phi\|_{L^2(\mu)\to L^2(\mu)} \leqslant ABM\,.$$

An assumption on *the maximal singular function* $(K^*_\Phi 1)$ can be conveniently modified. Let us consider the following assumption of *a.e. finiteness of the maximal singular function*:

$$(K^*_\Phi 1)(x) < \infty \text{ for } \mu \text{ a. e. } x\,.$$

Fix a large M and $L > 100M$. Fix a bounded measurable function b. If x is such that $(K^*_\Phi b)(x) > L$, then there exists a maximal $\varepsilon_0(x)$ such that $|(K_{\Phi,\varepsilon_0} b)(x)| \geqslant L$ (the function $\varepsilon \to (K_{\Phi,\varepsilon_0} b)(x)$ is right continuous). Consider

$$G_L(b) := \cup_{x\in \operatorname{supp}\mu} B(x, 2\varepsilon_0(x))\,.$$

Lemma 4: Let us assume the a.e. finiteness of the maximal singular function. Then $\mu(G_L(b)\setminus H_M) \to 0$ if $L\to\infty$.

Remind that Calderón-Zygmund constants of kernel k_Φ are bounded by C.

Proof. Let $y \in B(x, 2\varepsilon_0(x)$. Let us prove first that

$$K^*_\Phi 1(y) \geq L - ACM$$

for an absolute constant A. In fact, let us consider two cases: a) $\Phi(x) \geq \frac{1}{5}\varepsilon_0(x)$, b) $\Phi(x) < \frac{1}{5}\varepsilon_0(x)$. In the first case let $\varepsilon = 20\Phi(x)$. Then

$$|K^\varepsilon_\Phi(y) - K^{\varepsilon_0(x)}_\Phi(x)| \leq \int_{z:|z-y|\geq\varepsilon} |k_\Phi(y,z) - k_\Phi(x,z)|\;, d\mu(z) + \int_{z:|z-y|\leq 20\Phi(x)} |k_\Phi(x,z)|\, d\mu(z)\,.$$

The first integral can be estimated as usual using the Calderón-Zygmund property of the kernel k_Φ and the fact that all "large" disks centered at y are contained in discs centered at x of "almost" the same radii. These radii will be larger than $\Phi(x)$, and, hence, they will be M-Ahlfors. The second integral is bounded by AM because $k_\Phi(x,z) \leq \frac{1}{\Phi(x)}$ and $\mu(B(y, 20\Phi(x)) \leq \mu(B(x, 40\Phi(x)) \leq 40M\Phi(x)$ (the first inequality holds because we are in the first case).

Let us consider case b) now. Put $\varepsilon = 4\varepsilon_0(x)$. Then

$$K^\varepsilon_\Phi(y) - K^{\varepsilon_0(x)}_\Phi(x)| \leq \int_{z:|z-y|\geq\varepsilon} |k_\Phi(y,z) - k_\Phi(x,z)|\, d\mu(z) +$$

$$\int_{z:|z-y|\leq 4\varepsilon_0(x)} |k_\Phi(x,z)|\, d\mu(z)\,.$$

The first integral can be estimated exactly as in the case a). The second integral is bounded by $\frac{C}{\varepsilon_0(x)}\,\mu(B(x, 6\varepsilon_0(x)))$, where C is the constant from Calderón-Zygmund properties of our kernel. The disc $B(x, 6\varepsilon_0(x))$ is M-Ahlfors because we are in case b). Thus the second integral is also bounded by ACM.

Now it is clear that the assumption

$$K^*_\Phi 1(x) < \infty \text{ for } \mu \text{ a.e } x$$

implies that

$$\mu(G_L(1)) \to 0 \text{ when } L \to \infty\,.$$

Lemma 4 is proved.

Recall that for a given M, H_M denotes the union of all M-non-Ahlfors disks.

Theorem 1a: Let μ satisfy the non-uniform linear growth condition, let Φ be a 1-Lipschitz function such that $\Phi(x) \geqslant \text{dist}(x, \mathbb{C} \setminus H_M)$, and assume the a.e. finiteness of the maximal singular function $(K^*_\Phi 1)(x)$. Let Ψ be a 1-Lipschitz function such that $\Psi(x) \geqslant \max[\text{dist}(x, \mathbb{C} \setminus (G_L(1)), \Phi(x)]$. Then
1) $(K^*_\Psi 1)(x) \leqslant AC(L+M)$ uniformly, and
2) $\|K_\Psi\|_{L^2(\mu)\to L^2(\mu)} \leqslant AC(L+M)M$.

Remark. In the first claim of Theorem 1a one can replace 1 by any bounded function $b, \|b\|_\infty \leq 1$.

Proof. The second claim of the Theorem follows from the first claim and from Theorem 1. The first claim is a simple calculation using Lemma 1. Let us do it for the sake of completeness.

Let $C^\varepsilon_\Psi f(x) := \int_{y:|y-x|\geqslant \max[\varepsilon, \Psi(x)]} k(x,y)\, d\mu(y)$.

Step 1. For any 1-Lipschitz $\Psi \geqslant \Phi$ and any ε,

$$|K^\varepsilon_\Psi 1(x) - C^\varepsilon_\Psi 1(x)| \leqslant AM_{1,\Psi}1(x) \leqslant AM_{1,\Psi}1(x) \leqslant AM\,.$$

In fact, if $\varepsilon \leqslant \Psi(x)$, then

$$|K^\varepsilon_\Psi 1(x) - C^\varepsilon_\Psi 1(x)| \leqslant \int_{\varepsilon\leqslant|y-x|\leqslant\Psi(x)} \frac{d\mu(y)}{\Psi(x)} + \int_{y:|y-x|>\Psi(x)} \frac{\Psi(x)\Psi(y)|x-y|}{(|x-y|^2+\Psi(x)\Psi(y))|x-y|^2}\,.$$

The first term is bounded by $\frac{\mu(B(x,\Psi(x))}{\Psi(x)} \leqslant M$.

The second term can be estimated precisely as in Lemma 1 if we use that $\Psi \geqslant \Phi$. So it is also bounded by AM. If now $\varepsilon > \Psi(x)$, then only the second term will appear. We are done with the first step.

Step 2. Recall that $\varepsilon_0(x) = \max\{\varepsilon : |K^\varepsilon_\Phi 1(x)| \leqslant L\}$. Fix x_0 and let $\varepsilon \leqslant \varepsilon_0(x_0)$. Then $\Psi(x_0) \geqslant 2\varepsilon_0(x_0) > \varepsilon$. Then $|K^\varepsilon_\Psi 1(x_0)| \leqslant \int_{\varepsilon\leqslant|y-x|\leqslant 2\varepsilon_0} |k_\Psi(x_0,y)\, d\mu(y) + |K^{2\varepsilon_0}_\Psi 1(x_0)|$. The first term is bounded by $\frac{\mu(B(x_0,2\varepsilon_0)}{\Psi(x_0)} \leqslant \frac{\mu(B(x_0,\Psi(x_0))}{\Psi(x_0)} \leqslant M$ since $\Psi \geqslant \Phi$. Using step 1 we can reduce the estimate of the second term to the estimate of $|C^{2\varepsilon_0}_\Psi 1(x_0)|$ (with the error bounded by AM). Let us now use the fact that $\Psi(x_0) \geqslant 2\varepsilon_0(x_0)$, $\Psi(x_0) \geqslant \Phi(x_0)$. This means that $C^{2\varepsilon_0}_\Psi 1(x_0) = C^{\Psi(x_0)}_\Phi 1(x_0)$. By another application of step 1 we can see that the last quantity is within AM of $K^{\Psi(x_0)}_\Phi 1(x_0)$. The absolute value of this expression is bounded by L by the

definition of ε_0 and the fact that $\Psi(x_0) \geqslant 2\varepsilon_0$. In particular, our second term is bounded by $L + AM$.
Step 3. $\varepsilon_0(x_0) < \varepsilon \leqslant \Psi(x_0)$. Then $|K^\varepsilon_\Psi 1(x_0) - K^\varepsilon_\Phi 1(x_0)| \leqslant \int_{\varepsilon \leqslant |y-x_0| \leqslant \Psi(x_0)} |k_\Psi(x_0,y)|\, d\mu(y) + |\int_{\varepsilon \leqslant |y-x_0| \leqslant \Psi(x_0)} k_\Phi(x_0,y)\, d\mu(y)| + \int_{y:|y-x_0| \geqslant \Psi(x_0)} |k_\Phi(x_0,y) - k_\Psi(x_0,y)|\, d\mu(y)$. The first term is bounded by $\frac{\mu(B(x_0,\Psi(x_0)))}{\Psi(x_0)} \leqslant M$. The second term is bounded by $|K^\varepsilon_\Phi 1(x_0)| + |K^{\Psi(x_0)}_\Phi 1(x_0)| \leqslant 2L$, because $\varepsilon_0(x_0) < \varepsilon \leqslant \Psi(x_0)$, just by the definition of ε_0. The third term can be estimated precisely as in Lemma 1 if we notice that the integrand is bounded by $|k_\Phi(x_0,y) - k(x_0,y)| + |k(x_0,y) - k_\Psi(x_0,y)| \leqslant \frac{2\Psi(x_0)}{|x-x_0|^2}$. The integral then is bounded by AM.
Step 4. $\varepsilon > \Psi(x_0)$. We use the first step to write $|K^\varepsilon_\Psi 1(x_0) - C^\varepsilon_\Psi(x_0)| \leqslant AM$ and also $|K^\varepsilon_\Psi 1(x_0) - C^\varepsilon_\Psi(x_0)| \leqslant AM$. Therefore, we are left to estimate $|C^\varepsilon_\Phi 1(x_0) - C^\varepsilon_\Psi(x_0)|$. But this quantity vanishes because $\varepsilon > \Psi(x_0) \geqslant \Phi(x_0)$.
The first claim of Theorem 1a is completely proved. We have already made a remark that the second claim follows from the first one combined with Theorem 1.

Before proving Theorem 1, we would like to present its beautiful application found by Xavier Tolsa [XT2].

Recall that $R(x,y,z)$ denotes the radius of the circle passing through $x,y,z \in \mathbb{C}$. We will call a measure $\mu \geqslant 0$ on $\mathbb{C}$ Besicovitch-Melnikov-Verdera rectifiable if $\mu = \sum_{n=0}^\infty \mu|E_n$, $E_n, n = 1,2,3,...$ are compact sets, and

$$c^2(\mu|E_n) := \iiint_{E_n^3} R^{-2}(x,y,z)\, d\,\mu(x)\, d\,\mu(y)\, d\,\mu(z) < \infty,\ n = 1,2,3,...; \mu(E_0) = 0\,.$$

The curvature $c^2(\mu)$ was widely used by Melnikov and Verdera (see, for example [MV]). The name is natural because if $\mu = \mathcal{H}^1|E$, E being a compact set, then μ turns out to be a Besicovitch-Melnikov-Verdera rectifiable measure if and only if E is a Besicovitch rectifiable set. This is a difficult geometric result proved by G. David and J.-C. Léger. This result becomes especially difficult because of the nonhomogeneity of the measure μ, namely because $\liminf_{r\to 0} \frac{\mu(B(x,r))}{r}$ may apriori easily become 0.

In his paper [XT2] Xavier Tolsa gave the following application of Theorem 1a. We use the notations of Theorem 1.

Theorem (Xavier Tolsa): If μ satisfies non-uniform linear growth condition and if for any M the assumption of a.e. finiteness of the maximal singular function $(K^*_\Phi 1)(x)$ holds for $\Phi(x) := \mathrm{dist}(x, \mathbb{C} \setminus H_M)$, then μ is Besicovitch-Melnikov-Verdera rectifiable. If in addition for μ a.e. x, $\limsup_{r\to 0} \frac{\mu(B(x,r))}{r} > 0$, then $\mathrm{supp}\,\mu$ is Besicovitch rectifiable. If E is a compact set such that $\mathcal{H}^1(E) < \infty$, then E is Besicovitch rectifiable if and only if the principal value of the Cauchy integral $C_{\mathcal{H}^1|E}(x)$ exists for $\mathcal{H}^1$ a.e. $x \in E$.

The last claim completely solves the conjecture of Mattila [Ma]. Mattila proved this result with the extra assumption of "non-uniform homogeneity":

$$\liminf_{r\to 0} \frac{\mathcal{H}^1(B(x,r))}{r} > 0 \text{ for } \mathcal{H}^1 \text{ a.e. } x \in E\,.$$

Proof. We will prove the first assertion. The rest is not difficult to deduce. We choose L, M so large that

$$\mu(G_L(1) \cup H_M) < \frac{1}{2}\,.$$

Theorem 1a says that with $\Psi(x) := \mathrm{dist}(x, \mathbb{C}\setminus(G_L(1)\cup H_M))$ the operator K_Ψ is bounded on $L^2(\mu)$ with norm at most ALM. Consider $G_{L,M} := G_L(1) \cup H_M$, $f = 1_{\mathbb{C}\setminus G_{L,M}}$. Then (*) implies

$$c^2(\mu|\mathbb{C}\setminus G_{L,M})) = \int_{\mathbb{C}\setminus G_{L,M})} |C 1_{\mathbb{C}\setminus G_{L,M})}|^2\, d\,\mu = \int_{\mathbb{C}\setminus G_{L,M})} |K_\Psi f|^2\, d\,\mu \leqslant ALM\,.$$

The first equality is the famous formula of Melnikov-Verdera from [MV]. Notice that $\mu(\mathbb{C}\setminus G_{L,M}) > \frac{1}{2}$. But choosing larger L, M we can make $\mu(\mathbb{C}\setminus G_{L,M})$ as close to 1 as we wish (recall that our convention is that $\|\mu\| = 1$). So we can scoop the measure μ by pieces with finite curvature c^2. This proves the first claim of the Theorem.

The other claims now follow easily. For example, the a.e. existence of the principal value of the Cauchy integral $C_{\mathcal{H}^1|E}(x)$ implies the a.e. finiteness of the maximal singular integral $C^*_{\mathcal{H}^1|E}(x)$. This and the non-uniform linear growth of $\mathcal{H}^1$ (it always has this property) imply that a.e. finiteness of the maximal singular integral $(K^*_\Phi 1)(x)$ holds for $\Phi(x) := \mathrm{dist}(x, \mathbb{C}\setminus H_M)$ and any M (see Section VIII). Then $\mu = \mathcal{H}^1|E$ is a Besicovitch-Melnikov-Verdera rectifiable measure (by the first claim). The result of David and Léger now shows that E is Besicovitch rectifiable.

To prove Theorems 1 (and, so, to prove the second claim of Theorem 1a) we need to use our decomposition into good and bad functions. Recall that we used the probability space (Ω, P) of *pairs* of random dyadic lattices, $\omega = (\omega_1, \omega_2)$, here ω_i "enumerates" the i-th $(i = 1, 2)$ dyadic lattice D_i. These lattices D_1, D_2 are independent. We used also the notion of "good" and "bad" squares in D_1 and D_2. We also used the decomposition of sure functions $f, g \in L^2(\mu)$ to random functions

$$f = f_{good} + f_{bad}, \quad g = g_{good} + g_{bad}\,,$$
$$f_{bad} = \Sigma_{Q\in D_1, Q \text{ is bad}} \Delta_Q f, \quad g_{bad} = \Sigma_{R\in D_2, R \text{ is bad}} \Delta_R g\,.$$

The proof of Theorem 1 (and, so, of 1a) is based on the following lemma.

Lemma 5: Let μ be a measure with non-uniform linear growth, let H_M be the union of all M-non-Ahlfors discs, let

$$\Phi(x) \geqslant \mathrm{dist}(x, \mathbb{C}\setminus H_M)\,,$$

and let Φ is a 1-Lipschitz function. Consider $(K_{\Phi,\varepsilon} f)(x) := \int_{|y-x|\geqslant\varepsilon} k_\Phi(x,y) f(y)\, d\,\mu(y)$. Let B be a finite constant such that

$$(K^*_\Phi 1)(x) := \sup_{\varepsilon>0} |(K_{\Phi,\varepsilon} 1)(x)| \leqslant B \text{ for } \mu \text{ a. e. } x\,.$$

Then

$$|(K_\Phi f_{good}, g_{good})| \leqslant ABM\|f\|\|g\|\,,$$

$$|(K_\Phi f, g)| \leqslant ABM\|f\|\|g\| + \|K_\Phi\| R(\omega, f, g),$$

where the expectation of the remainder $R(\omega, f, g)$ has the following estimate: $\mathbb{E}R(\omega, f, g) \leqslant \frac{1}{2}\|f\|\|g\|$.

The inequalities of the lemma imply immediately Theorem 1. In its turn, the last inequality follows from the first one and the fact proved in the previous sections:

$$\mathbb{E}\|f_{bad}\|^2 \leqslant 45^{-239}\|f\|^2\,,$$
$$\mathbb{E}\|g_{bad}\|^2 \leqslant 45^{-239}\|g\|^2\,.$$

The proof of the first inequality of the lemma takes a good part of previous sections. So, Theorem 1 and 1a are proved.

What if we replace the function 1 by a complex valued function b (even, say, real valued but not always positive) in one of our main assumptions:

$$(K^*_\Phi b)(x) := \sup_{\varepsilon>0} |(K_{\Phi,\varepsilon} b)(x)| \leqslant B < \infty \text{ for } \mu \text{ a. e. } x\,?$$

This is equivalent to still having the function 1 but having *complex* measure μ. We prefer to denote by μ only positive measures, and to use the symbol ν for $b\,d\,\mu$. So now b is an $L^\infty(\mu)$-function of norm 1, and we assume that

$$(C^* b)(x) := \sup_{\varepsilon>0} |(C^\varepsilon b)(x)| < \infty \text{ for } \mu \text{ a. e. } x\,.$$

We do not write the subscript μ because it will be always the same μ.

We still assume everywhere below that μ has the non-uniform linear growth condition (unless it is stated otherwise).

Now we are in the framework of the Tb theorem rather than the $T1$ theorem. The main problem we encounter is that our b will *not be accretive*. The second problem (we always have it in this paper) is that μ has no doubling property.

We start again by considering the set H_M of all M-non-Ahlfors discs for μ. Again we can see that our assumption on $(C^*b)(x)$ implies (see Lemma 1) *the a.e finiteness of the maximal singular operator*:

$$(K^*_\Phi b)(x) < \infty \text{ for } \mu \text{ a. e. } x$$

for every 1-Lipschitz Φ such that $\Phi(x) \geqslant \operatorname{dist}(x, \mathbb{C}\setminus H_M))$. Exactly as before we can introduce the sets $G_L = G_L(b) =$ the union of $B(x, 2\varepsilon_0(x))$, where ε_0 is the maximal radius for which $|(K_{\Phi,\varepsilon_0} b)(x)| \geqslant L$, and $G_{L,M} = G_L \cup H_M$.

Lemma 6: Let $(K^*_\Phi b)(x) < \infty$ for μ a. e. x hold. Then $\mu(G_L) \to 0$ when $L \to \infty$.

Let $\Psi(x) = \operatorname{dist}(x, \mathbb{C} \setminus G_{L,M})$. The set $F_\Psi = \{x \in \operatorname{supp}\mu : \Psi(x) = 0\}$ has measure close to the full measure of μ. Unfortunately, unlike in Theorem 1a, we cannot say that K_Ψ is bounded on $L^2(\mu)$. The place where the proof will break down is Lemma 5. The estimate for good functions will not work. This is because $\Delta_Q f$ is now adapted to the function b. On squares where accretivity of b becomes very bad (or non-existent) the pieces $\Delta_Q f$ will blow up because the accretivity constant lives in their denominators. This was impossible for $b = 1$—it is accretive in any scale. To deal with this problem of non-accretivity of b we need even more randomness: first let us assume that for a certain positive η the union of squares ("non-accretive squares") $Q \in D_1$ such that

$$|\int_Q b\, d\,\mu| < \eta\, \mu(Q)$$

has total measure less than δ, and this is uniformly true for every random lattice D_1 (so for D_2 also).

Let T_i be the family of "non-accretive" squares of D_i, $i = 1, 2$, in the above sense. Let $\omega \in \Omega$. Let $\mathcal{T}_i^\omega = \cup_{Q\in T_i} Q$.

We have the (strange) assumption that

$$\mu(\mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega) \leqslant \delta \text{ for all } \omega \in \Omega\,. \qquad (**)$$

Lemma 7: Consider any 1-Lipschitz function Φ_ω such that $\Phi_\omega(x) \geqslant \text{dist}(x, \mathbb{C} \setminus (G_{L,M} \cup \mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega))$. Then

$$|(K_{\Phi_\omega} f_{good}, g_{good})| \leqslant ALM\, \eta^{-2} \|f\| \|g\|\,.$$

This lemma is the result of our previous sections. Using the last inequality we can obviously write

$$|(K_{\Phi_\omega} f, g)| \leqslant ALM\, \eta^{-2} \|f\| \|g\| + \|K_{\Phi_\omega}\| R(\omega, f, g)$$

with $R(\omega, f, g)$ having small average (exactly as in Lemma 5). But now it is not clear what to do with the random norm $\|K_{\Phi_\omega}\|$. We can consider a sure function $\Phi = \sup \Phi_\omega$. It is again 1-Lipschitz and again

$$|(K_\Phi f, g)| \leqslant ALM\, \eta^{-2} \|f\| \|g\| + \|K_\Phi\| R(\omega, f, g)$$

with small $\mathbb{E} R(\omega, f, g)$. So the bound for the norm of $\|K_\Phi\|$ will follow by averaging the previous inequality.

But this is useless because our "nice" set

$$F_\Phi = \{x : \Phi(x) = 0\} = \cap_\omega F_{\Phi_\omega} = \cap_\omega \{x : \Phi_\omega(x) = 0\}$$

and these random sets could easily have empty intersection.

We have, however, an extra "strange" assumption (**): $\mu(\mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega) \leqslant \delta$ for all $\omega \in \Omega$. Then for sufficiently large L, M we have

$$\mu(G_{L,M} \cup \mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega) \leqslant 2\delta \text{ for all } \omega \in \Omega\,.$$

Notice that this means (by Fubini's theorem and Tchebyshov's inequality) that

$$\mu\{x : P\{\omega : x \in G_{L,M} \cup \mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega\} \leqslant \sqrt{2\delta}\} \geqslant 1 - \sqrt{2\delta}\,.$$

We can now consider the expectation of Φ_ω rather than maximum. Moreover, as we have done in Section IV, let us now consider *the truncated mathematical expectation*:

$$\Psi(x) := \inf\{\mathbb{E}(\Phi_\omega(x)\, 1_S(\omega)) : S \subset \Omega,\ P(S) = 1 - \sqrt{2\delta}\}\,.$$

Now we have the good estimate for the zero set F_Ψ:

$$\mu(F_\Psi) \geqslant 1 - \sqrt{2\delta}\,.$$

On the other hand, Lemma 7 can leads us to

Theorem 2: Let μ have a non-uniform linear growth condition. Assume the a.e. finiteness of maximal singular operator, namely:

$$(C^* b)(x) := \sup_{\varepsilon>0} |(C^\varepsilon b)(x)| < \infty \ \text{ for } \mu \text{ a. e. } x$$

We also assume that μ has the non-uniform linear growth condition. Assume also (**). Let $\Phi_\omega(x) = \operatorname{dist}(x, \mathbb{C}\setminus(G_{L,M}\cup\mathcal{T}_1^\omega\cup\mathcal{T}_2^\omega))$, and let Ψ be the truncated mathematical expectation of Φ_ω defined above. Then
1)

$$|(K_\Psi f, g)| \leqslant ALM\eta^{-2}\|f\|\|g\| + (\|K_\Psi^*\| + AM)\, R(\omega, f, g)$$

where $\mathbb{E}R(\omega, f, g) \leqslant A\delta\|f\|\|g\|$.
2) In particular, $\|K_\Psi\|_{L^2(\mu)\to L^2(\mu)} \leqslant ALM\eta^{-2}$.
Automatically, for the set F_Ψ (whose measure $\mu(F_\Psi) \geqslant 1 - \sqrt{2\delta}$) we have

$$\|C\|_{L^2(F_\Psi,\mu)\to L^2(F_\Psi,\mu)} \leqslant ALM\eta^{-2}\,.$$

This theorem was proved by all the previous sections. However, the second part of the theorem requires the estimate of $\|K_\Psi^*\|$ via $\|K_\Psi\|$. This is done in [NTV2] for Ahlfors measures (i.e. measures having a uniform linear growth condition). Completely similar reasoning for non-uniformly Ahlfors measures (i.e. measures having a non-uniform linear growth condition) can be found in Section XXV of the present paper.

Theorem 2 gives the analytic part of Vitushkin's conjecture but without the estimate of how large the rectifiable part of Vitushkin's compact is, and how rectifiable it is. This is because the assumption (**) does not happen very often. In fact, why should an arbitrary non-zero complex function b (and in applications we usually do not know anything else about b) be accretive except for a small set? In our previous sections we achieve (**) by localizing our considerations to an unspecified small disc around a Lebesgue point x_0 of b, where $b(x_0) \neq 0$. Clearly, this way will not lead us to the proof of quantitative version of Vitushkin's conjecture.

However, there is one piece of information which is usually available about b, and which has not been used so far. Namely, we have the accretivity of b in one—the highest—scale:

$$\|b\|_\infty = 1,\ \ |\int_{\mathbb{C}} b\, d\mu| = \gamma > 0\,. \qquad (\gamma)$$

This brings us to the quantitative version of Tb theorem, where b has accretivity only at the highest scale. We do not assume (**), but we assume (γ). As always $\|\mu\| = 1$.

Theorem 3: Assume the a.e. finiteness of the maximal singular operator:

$$(C^* b)(x) := \sup_{\varepsilon>0} |(C^\varepsilon b)(x)| < \infty \text{ for } \mu \text{ a. e. } x\,.$$

Also assume that μ has the non-uniform linear growth condition. Assume also (γ). Then there exists a set F, $\mu(F) \geqslant \frac{3\gamma}{16}$, such that

$$\|C\|_{L^2(F,d\mu)\to L^2(F,d\mu)} \leqslant A L(\gamma) M(\gamma) \gamma^{-20}$$

where $M(\gamma) = \inf\{M : \mu(H_M) < \frac{\gamma}{32}\}$ and $L(\gamma) = \inf\{L : \mu(G_L \setminus H_{M(\gamma)})) < \frac{\gamma}{32}\}$.

The next theorem is the promised quantitative version of Vitushkin's conjecture. We will obtain it (easily) as a corollary of Theorem 3.

Theorem 4 (quantitative version of Vitushkin's conjecture): Let $E \subset \mathbb{C}$ be a compact set such that its length $\mathcal{H}^1(E) = M < \infty$ and its analytic capacity $\gamma(E) = \gamma > 0$. Then there exists a set F, $\mathcal{H}^1(F) \geqslant \frac{\gamma}{16}$, such that

$$c^2(\mathcal{H}^1|F) \leqslant A\,(\frac{\text{diam} E}{\gamma})(\frac{M}{\gamma})^{42} \mathcal{H}^1(F)\,.$$

Proof of Theorem 3: Consider T_i = maximal squares from D_i such that

$$|\int_Q b\, d\,\mu| \leqslant \frac{\gamma}{2}\mu(Q)\,.$$

Put $\mathcal{T}_i = \cup_{Q\in T_i} Q$, $i = 1, 2$. For brevity, let $E = \operatorname{supp}\mu$. Using ($\gamma$) we have

$$|\int_E b\, d\,\mu| = |\int_{\mathcal{T}_1} b\, d\,\mu| + |\int_{E\setminus\mathcal{T}_1} b\, d\,\mu| =$$

$$|\Sigma_{Q\in T_1} \int_Q b\, d\,\mu| + |\int_{E\setminus\mathcal{T}_1} b\, d\,\mu| \leqslant \frac{\gamma}{2}\Sigma_{Q\in\mathcal{T}_1}\mu(Q) + \mu(E\setminus\mathcal{T}_1) \leqslant$$

$$\frac{\gamma}{2} + \mu(E\setminus\mathcal{T}_1)\,.$$

Therefore,

$$\mu(E\setminus\mathcal{T}_1^\omega),\ \mu(E\setminus\mathcal{T}_2^\omega) \geqslant \frac{\gamma}{2}\,.$$

We wrote the superscript ω to emphasize that these are random sets. We want to show that for some detectable (=not very small) set of $x \in E$ the probability $p(x) := P\{\omega : x \in E \setminus (\mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega)\}$ is not too small. Denote $p_1(x) := P\{\omega : x \in E \setminus \mathcal{T}_1^\omega\}$. Notice that the sets

$E \setminus \mathcal{T}_1^\omega$, $E \setminus \mathcal{T}_2^\omega$ are independent and that $E \setminus (\mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega)\} = (E \setminus \mathcal{T}_1^\omega) \cap (E \setminus \mathcal{T}_2^\omega)$. Therefore, $p(x) = p_1(x)^2$. Also

$$\int_E p_1(x)\, d\,\mu = \mathbb{E} \int 1_{E \setminus \mathcal{T}_1^\omega}\, d\,\mu = \mathbb{E}\mu(E \setminus \mathcal{T}_1^\omega) \geqslant \frac{\gamma}{2}\,.$$

Now let us split $E = S \cup L$, where $S := \{x \in E : p_1(x) \leqslant \frac{\gamma}{4}\}$ and $L := \{x \in E : p_1(x) > \frac{\gamma}{4}\}$. Then we have $\mu(L) \geqslant \frac{\gamma}{4}$. For $x \in L$, $p(x) = p_1^2(x) > \frac{\gamma^2}{16}$. For the sake of brevity we denote $\beta = \frac{\gamma^2}{16}$. So

$$\mu\{x \in E : P\{\omega : x \in E \setminus (\mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega)\} > \beta\} \geqslant \frac{\gamma}{4}\,.$$

Now let us choose $M = M(\gamma), k = L(\gamma)$ to be smallest numbers such that

$$\mu(H_M) \leqslant \frac{\gamma}{32},\ \ \mu(G_L \setminus H_M) \leqslant \frac{\gamma}{32}\,.$$

Consider $O^\omega := G_{L,M} \cup \mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega$. Put $\Phi_\omega(x) := \text{dist}(x, \mathbb{C} \setminus O^\omega)$. Thus,

$$\mu\{x \in E : P\{\omega : \Phi_\omega(x) = 0\} > \beta\} > \frac{3\gamma}{16}\,.$$

Let us introduce sure 1-Lipschitz function

$$\Phi_0(x) := \inf_{S \subset \Omega,\, P(S) = \beta} \sup_{\omega \in S} \Phi_\omega(x)\,.$$

Let us also fix a small positive number τ and put

$$\Phi(x) := \Phi_0(x) + \tau\,.$$

All estimates in the future will not depend on τ.

We know that the zero set $F := F_{\Phi_0}$ has detectable measure:

$$\mu(F) > \frac{3\gamma}{16}\,.$$

We will need a small modification of Lemma 7 of this section.

Lemma 7a: Consider any 1-Lipschitz function Φ_ω such that $\Phi_\omega(x) \geqslant \text{dist}(x, \mathbb{C} \setminus (G_{L,M} \cup \mathcal{T}_1^\omega \cup \mathcal{T}_2^\omega))$. Fix a small positive number ε. Then we can decompose $f = f_{good} + f_{bad}, g = g_{good} + g_{bad}$ in such a way that

$$\mathbb{E}\|f_{bad}\| \leq \varepsilon\|f\|,\ \mathbb{E}\|g_{bad}\| \leq \varepsilon\|g\|\,,$$

and

$$|(K_{\Phi_\omega} f_{good}, g_{good})| \leqslant ALMC(\varepsilon)\, \eta^{-2}\|f\|\|g\|, \text{ where } C(\varepsilon) \leq A\varepsilon^{-8}\,.$$

All the previous sections were devoted to the proof of such a statement (called Lemma 7 in this section) with a fixed small absolute constant ε (it has been chosen to be 45^{-239}). But the same proof gives Lemma 7a because in our calculations in Section XXII we can choose a very large m and a very large $\widetilde{M}$ in accordance with the smallness of ε. They can be chosen to achieve our first inequality of Lemma 7a. Then the second inequality of Lemma 7a follows from the bookkeeping of the estimate of the bilinear form of the operator K_{Φ_ω} on good functions.

Main Lemma: Operator C_Φ is bounded on $L^2(\mu)$ by $AL(\gamma)M(\gamma)\gamma^{-20}$ (and the bound does not depend on τ).

Proof. Fix $\varepsilon = a\gamma^2$. Here a is a small positive absolute constant. Recall that the splitting into good and bad functions can be made dependent on a number ε. Then

$$\mathbb{E}\|f_{bad}\| \leqslant \varepsilon\|f\|,\ \ \mathbb{E}\|f_{bad}\| \leqslant \varepsilon\|f\|\,.$$

Lemma 7a (with $\eta = \gamma/2$) then states the following:

$$|(K_{\Phi\vee\Phi_\omega} f_{good}, g_{good})| \leqslant ALMC(\varepsilon)\gamma^{-2},\ \text{ with } C(\varepsilon) \leqslant A\,\varepsilon^{-8}\,.$$

We used the notations $\Phi\vee\Phi_\omega = \max(\Phi,\Phi_\omega)$. We use now Lemma 1.

$$|(K_{\Phi\vee\Phi_\omega} f, g)| \leqslant |(K_{\Phi\vee\Phi_\omega} f_{good}, g_{good})| + |(C_{\Phi\vee\Phi_\omega} f_{bad}, g_{good})| +$$

$$|(C_{\Phi\vee\Phi_\omega} f_{good}, g_{bad})| + |(C_{\Phi\vee\Phi_\omega} f_{bad}, g_{bad}) + A\|M_{1,\Phi} f\|\|g\|\,.$$

Notice that $\Phi(x) \geqslant \mathrm{dist}(x, \mathbb{C}\setminus G_{L,M})$. Using Lemma 3 we make an estimate in the last term:

$$\|M_{1,\Phi} f\| \leqslant AM\|f\|\,.$$

The estimate of $|(C_{\Phi\vee\Phi_\omega} f_{bad}, g_{good})| + ...$ involves an important lemma and several notations. Let $k_\omega(x,y)$ denote the kernel of $K_{\Phi\vee\Phi_\omega}$. Let $c_\omega(x,y)$ denote the kernel of $C_{\Phi\vee\Phi_\omega}$.

Notice that

$$p_\omega(x,y) := |k_\omega(x,y) - c_\omega(x,y)|$$

is a "Poisson" type kernel. In particular,

$$\int p_\omega(x,y)|f(y)|\,d\,\mu(y) \leqslant A\,(M_{1,\Phi} f)(x)$$

Consider the averaging of the kernels: $k = \mathbb{E}k_\omega$, $c = \mathbb{E}c_\omega$, $p = \mathbb{E}p_\omega$. The same "Poisson" property holds then for the average $p = \mathbb{E}p_\omega$):

$$\int p(x,y)|f(y)|\,d\,\mu(y) \leqslant A\,(M_{1,\Phi} f)(x)\,.$$

Let us also introduce operators c^*, k^* as follows:

$$(c^* f)(x) := \sup_{r>0} |\int_{|y-x|\geqslant r} c(x,y)f(y)\,d\,\mu(y)|,\ \ (k^* f)(x) := \sup_{r>0} |\int_{|y-x|\geqslant r} k(x,y)f(y)\,d\,\mu(y)|\,.$$

The same "Poisson" property holds then for the comparison of k^* and c^* (notice that k, c are defined in such a way that $|k(x,y)|, |c(x,y)| \le \frac{1}{\Phi(x)}$:

$$(c^* f)(x) \leqslant (k^* f)(x) + (M_{1,\Phi} f)(x) \,.$$

We are ready to formulate the main inequalities:

$$|(C_\Phi f)(x)| \leqslant \frac{A}{\gamma^2}((c^* f)(x) + (M_{1,\Phi} f)(x)) \,, \tag{MI}$$

$$|(C_{\Phi\vee\Phi_\omega} f)(x)| \leqslant \frac{A}{\gamma^2}((c^* f)(x) + (M_{1,\Phi} f)(x)) \,. \tag{MI}$$

Let us use (MI) to estimate

$$|(C_{\Phi\vee\Phi_\omega} f_{bad}, g_{good})| + |(C_{\Phi\vee\Phi_\omega} f_{good}, g_{bad})| + |(C_{\Phi\vee\Phi_\omega} f_{bad}, g_{bad})| \,.$$

After that we will prove (MI). By (MI), Lemma 1 and the Poisson property for the comparison of k^* and c^*: $|(C_{\Phi\vee\Phi_\omega} f_{bad}, g_{good})| \leqslant \frac{A}{\gamma^2}(\|(c^* f_{bad})\|\|g\| + \|M_{1,\Phi} f)\|\|g\|) \leqslant \frac{A}{\gamma^2}\|(k^* f_{bad})\|\|g\| + \frac{A}{\gamma^2}\|M_{1,\Phi} f\|\|g\|$. We continue:

$$|(C_{\Phi\vee\Phi_\omega} f_{bad}, g_{good})| \leqslant \frac{A}{\gamma^2}\varepsilon\|k^*\|\|f\|\|g\| + \frac{A}{\gamma^2}M\|f\|\|g\| \,.$$

Collecting our estimates for the good and bad function together, we get

$$|(K_{\Phi\vee\Phi_\omega} f, g)| \leqslant ALM\gamma^{-2}\varepsilon^{-8}\|f\|\|g\| + \frac{A}{\gamma^2}\varepsilon\|k^*\|\|f\|\|g\| + \frac{A}{\gamma^2}M\|f\|\|g\| \,.$$

We already fixed $\varepsilon = a\gamma^2$. Thus (with very small absolute a)

$$|(K_{\Phi\vee\Phi_\omega} f, g)| \leqslant ALM\gamma^{-18}\|f\|\|g\| + Aa\|k^*\|\|f\|\|g\| + AM\gamma^{-2}\|f\|\|g\| \,.$$

Recall that k denotes the average of the kernel of $K_{\Phi\vee\Phi_\omega}$. Averaging the previous inequality we get

$$\|kf\| \leqslant ALM\gamma^{-18}\|f\| + Aa\|k^*\|\|f\| + \frac{A}{\gamma^2}M\|f\| \,. \tag{kI}$$

In Theorem 7.1 of [NTV2] it is proved that $\|k^* f\| \leqslant A_1 C\|f\| + A_2 C\|k\|\|f\|$, where C stands for the Calderón-Zygmund constant of the kernel. Theorem 7.1 of [NTV2] is valid for operators with Calderón-Zygmund kernels. This is the case here because the averaging k of the Calderón-Zygmund kernels k_ω is still a Calderón-Zygmund kernel.
However, there is a difference between the sitation in [NTV2] and the situation here. In [NTV2] one assumed that the measure μ has a uniform linear growth condition. Our μ, however, has only the non-uniform linear growth condition (we call such measures non-uniformly Ahlfors measures). We are going to formulate now an analog of Theorem 7.1

from [NTV2] that is valid for non-uniformly Ahlfors measures. First, recall that given a Calderón-Zygmund kernel and a measure μ we say that the operator T with kernel k (see [NTV3]) is a Calderón-Zygmund operator if it is bounded on $L^2(\mu)$. Also recall that

$$\tilde{M}_\beta g(x) := \sup_{r>0} \frac{1}{\mu(B(x,3r)} \left(\int_{B(x,r)} |g(y)|^\beta \, d\mu(y) \right)^{\frac{1}{\beta}} .$$

When $\beta = 1$ we write $\tilde{M}g(x)$ instead of $\tilde{M}_1 g(x)$.

Theorem 5. Let μ be a non-uniformly Ahlfors measure. Fix a positive number M, and let $\mathcal{R}(x) := \sup\{r > 0 : \mu(B(x,r)) > Mr\}$. Let $k(x,y)$ be a Calderón-Zygmund kernel having Calderrón-Zygmund constant C and such that

$$|k(x,y)| \leqslant \min\left[\frac{1}{\mathcal{R}(x)}, \frac{1}{\mathcal{R}(y)}\right] .$$

Let T be a Calderón-Zygmund operator with kernel k. Fix $\beta \in (1,2)$. Then the following Cotlar type inequality holds:

$$(T^* f)(x) \leqslant A_1 C[\tilde{M}Tf](x) + A_2 CM\tilde{M}_\beta f(x) + A_3 C\|T\|_{L^2(\mu)\to L^2(\mu)} \tilde{M}_\beta f(x) . \qquad (CI)$$

The proof follows exactly the lines of the proof of Theorem 7.1 of [NTV2]. But for the sake of completness we give a full proof in Section XXV.

Combining this result with inequality (kI), we get

$$\|k^* f\| \leqslant ALM\gamma^{-18}\|f\| + A\, a\|k^*\|\|f\| + AM\|f\| .$$

Finally, using the fact that a is very small we get the estimate for the maximal singular operator:

$$\|k^* f\| \leqslant 2ALM\gamma^{-18}\|f\| .$$

Now let us use again the “Poisson” property for the comparison of k^* and c^*: $(c^* f)(x) \leqslant (k^* f)(x) + (M_{1,\Phi} f)(x)$ to get

$$\|c^* f\| \leqslant ALM\gamma^{-18}\|f\| .$$

Let us use the first part of the main inequality (MI) to conclude now that

$$\|C_\Phi f\| \leqslant ALM\gamma^{-20}\|f\| .$$

The main Lemma is proved.

We are left to prove (MI).
The proof of (MI) is based on two ingredients—the calculation of the kernel $c(x,y)$ (average of $c_\omega(x,y)$) and on the following lemma.
As usual, given $R \geqslant 0$, we denote by $(M_{1,R} f)(x) = \sup_{r>R} \frac{1}{r} \int_{B(x,r)} |f(y)| \, d\,\mu(y)$.

Blanket Lemma: Let $b(x,y)$ be kernel such that $|b(x,y)| \leqslant \frac{1}{|x-y|}$. Then we have a well-defined $(b^*f)(x) := \sup_{r>0} |\int_{|y-x|>r} b(x,y)f(y)\,d\,\mu(y)|$. Let $R>0$ and let ϕ be a decreasing function on $[0,\infty)$, $0 \leqslant \phi \leqslant 1$. Consider

$$(b_R^\phi f)(x) := |\int_{|y-x|>R} b(x,y)\phi(|x-y|)f(y)d\mu(y)|\,.$$

Then

$$(b_R^\phi f)(x) \leqslant 2\,(b^*f)(x) + 2\,(M_{1,R}f)(x)\,.$$

Proof. Consider annuli $A_k(x) = \{y : 2^{k-1}R \leqslant |y-x| \leqslant 2^k R\}$. Then

$$(b_R^\phi f)(x) \approx \Sigma_{k\geqslant 1}\int_{A_k} b(x,y)\phi_k f(y)d\mu(y)$$

where ϕ_k are some values (say, left end point values) of $\phi(t)$ for $t \in [2^{k-1}R, 2^k R]$, $k = 1, 2, ...$ More precisely ($\phi_0 := 0$)

$$(b_R^\phi f)(x) = \Sigma_{k\geqslant 1}(\phi_k - \phi_{k-1})\int_{|y-x|\geqslant 2^{k-1}R} b(x,y)f(y)\,d\,\mu(y) + \text{Discrepancy}\,.$$

Thus, the monotonicity of ϕ implies

$$|\text{The first term}| \leqslant \phi_1|\int_{|y-x|\geqslant R} b(x,y)f(y)\,d\,\mu(y)| +$$

$$\Sigma_{k\geqslant 2}(\phi_{k-1} - \phi_k)|\int_{|y-x|\geqslant 2^{k-1}R} b(x,y)f(y)\,d\,\mu(y)| \leqslant 2\,(b^*f)(x)\sup\phi\,.$$

On the other hand, let us use the symbol J_k to denote the jump (the oscillation) of the monotone function ϕ on the interval $[a_k, a_{k+1}]$. Then

$$|\text{Discrepancy}| \leqslant \Sigma_{k\geqslant 1} J_k \frac{1}{2^{k-1}R}\int_{B(x,2^k R)} |f(y)|\,d\,\mu(y)\,.$$

We continue the previous estimate as follows:

$$|\text{Discrepancy}| \leqslant 2\,(M_{1,R}f)(x)\Sigma_{k\geqslant 1}J_k\,.$$

But ϕ was assumed to be monotone and $0 \leqslant \phi \leqslant 1$, so the sum of the jumps is bounded by 1. The lemma is proved.

We continue the proof of (MI). Let $t \geq \Phi(x)$. Then

$$v(t) := P\{\omega : \Phi \vee \Phi_\omega(x) \leqslant t\} \geq \gamma^2/16\,.$$

It is obvious that for $|x-y| < \Phi(x)$ we have $v(|x-y|) = 0$. Now let us compute the kernel $c(x,y) = \mathbb{E}c_\omega(x,y)$. Clearly,

$$c(x,y) = \frac{v(|x-y|)}{x-y} = \frac{\chi_{\mathbb{C}\setminus B(x,\Phi(x))}v(|x-y|)}{x-y}\,.$$

Put $\beta := \gamma^2/16$. To obtain (MI) we can apply the Blanket Lemma with $R(x) = \Phi(x)$ or $R(x) = \Phi \vee \Phi_\omega(x)$, with $b(x,y) = \frac{c(x,y)}{\beta}$ and $\phi(t) = \frac{\beta}{v(t)}$. Theorem 3 is completely proved.

XXIV. The proof of Theorem 4. The quantitative version of Vitushkin's conjecture.

Now let Γ be a compact on $\mathbb{C}$ whose $\mathcal{H}^1$ measure is L and whose analytic capacity is γ. We can think that Γ consists of finitely many circle arcs. Consider $x \in \Gamma$ and $R(x) > 0$ such that

$$\frac{\mathcal{H}^1(B(x,R)\cap\Gamma)}{R} > \frac{160\pi L}{\gamma}\,.$$

The union of such $B(x, R(x))$ is covered by $\cup B(x_j, 5R_j)$ and

$$\Sigma\,\mathcal{H}^1(\partial B(x_j, 5R_j)) \leqslant \frac{\gamma}{16}\,.$$

Let G be the boundary of the complement of $\cup_j B(x_j, 5R_j) \cup \Gamma$. Let $F = \Gamma \cap G$. It is now clear that

$$\mathcal{H}^1(G \setminus F) \leqslant \frac{\gamma}{16}\,.$$

It is easy to check that there is no $1000\,L/\gamma$-non-Ahlfors disc for G. On the other hand, there exists a function b on G such that its Cauchy integral is bounded by 1 outside of G (its Cauchy integral vanishes inside all $B(x_i, 5R_i)$), such that $\|b\|_\infty \leqslant 1$, and such that

$$|\int_G b\,d\,\mathcal{H}^1| = \gamma\,.$$

As b we can take just the Ahlfors function of Γ outside of $\cup B(x_i, 5R_i)$ and zero inside. In particular,

$$(C^*\,b\,d\,\mathcal{H}^1)(x) \leqslant A\frac{L}{\gamma} \text{ for } \mathcal{H}^1 \text{ a.e } x \in G\,.$$

Let us consider the normalized measure $\mu := \mathcal{H}^1/L$ restricted on Γ. Then we are under the assumptions of Theorem 3, where we can put $L := \frac{1}{\gamma}$, $M := \frac{1}{\gamma}$, $\gamma := \frac{\gamma}{L}$ and get a set $F_0 \subset E$ with $\mu(F_0) \geqslant \frac{\gamma}{8L}$, that is with $\mathcal{H}^1(F_0) \geqslant \frac{\gamma}{8}$, such that $\|C\|_{L^2(F_0,\mu)\to L^2(F_0,\mu)} \leqslant A\gamma^{-2}(\gamma/L)^{-20}$. That is $\|C\|_{L^2(F_0,\mathcal{H}^1)\to L^2(F_0,\mathcal{H}^1)} \leqslant A\gamma^{-1}(\gamma/L)^{-21}$.

Consider $F^* := F_0 \cap F$. Then $\mathcal{H}^1(F_0) \geqslant \frac{\gamma}{8}$ and $\mathcal{H}^1(G \setminus F) \leqslant \frac{\gamma}{16}$ imply that

$$\mathcal{H}^1(F^*) \geqslant \frac{\gamma}{16}\,.$$

The advantage of F^* is that it is contained in the original set Γ because F is. Also we have

$$\|C\|_{L^2(F^*,\mathcal{H}^1)\to L^2(F^*,\mathcal{H}^1)} \leqslant A\gamma^{-1}(\gamma/L)^{-21}$$

just because $F^* \subset F_0$. The last relationship and the formula of Melnikov-Verdera shows

$$c^2(\mathcal{H}^1|F^*) \leqslant (A\gamma^{-1}(\gamma/L)^{-21})^2\mathcal{H}^1(F^*) = A\gamma^{-2}(\gamma/L)^{-42}\mathcal{H}^1(F^*)\,.$$

We tacitly assumed $\mathrm{diam}\Gamma = 1$. Thus, we have in general

$$c^2(\mathcal{H}^1|F^*) \leqslant A(\frac{\mathrm{diam}\Gamma}{\gamma})^2(\gamma/L)^{-42}\mathcal{H}^1(F^*)\,.$$

Theorem 4 is proved.

XXV. The proof of Theorem 5. Cotlar's inequality for non-uniformly Ahlfors measures.

We start the proof by fixing $r > 0, x \in \operatorname{supp}\mu$, and putting $\hat{r} = \max[r, \mathcal{R}(x)]$. Consider $(T^r f)(x) := \int_{y:|y-x|\geqslant r} k(x,y) f(y)\,d\mu(y)$. Put $r_j := 3^j\hat{r}, \mu_j := \mu(B(x,r_j))$. Let k be the smallest index such that $\mu_{k+1} \leq 36\mu_{k-1}$. It exists, because otherwise, for every k, $\mu(B(x,\hat{r})) \leq 36^{-k}\mu_{2k} \leq 2M36^{-k}r_{2k}$. This is because our radii are greater than $\mathcal{R}(x) := \sup\{r > 0 : \mu(B(x,r)) > Mr\}$. We continue with $\mu(B(x,\hat{r})) \leq 2M36^{-k}3^{2k}\hat{r} = 2M2^{-2k}\hat{r}$. This contradicts the assumption $x \in \operatorname{supp}\mu$.

Let $R := r_{k-1}$. We estimate $|(T^r f)(x) - (T^{3R} f)(x)| \leq \int_{B(x,\hat{r})\setminus B(x,r)} |k(x,y)|\,|f(y)|\,d\mu(y) + \sum_{j=1}^k \int_{B(x,r_j)\setminus B(x,r_{j-1})} \ldots$. The first term vanishes if $\hat{r} > \mathcal{R}(x)$. Otherwise it is bounded by

$$\frac{1}{\mathcal{R}(x)}\int_{B(x,\hat{r})} |f(y)|\,d\mu(y) = \frac{1}{\hat{r}}\int_{B(x,\hat{r})} |f(y)|\,d\mu(y) \leq$$

$$\frac{\mu(B(x,3\hat{r})}{\hat{r}\mu(B(x,3\hat{r})}\int_{B(x,\hat{r})} |f(y)|\,d\mu(y)\,.$$

And this is less than $AM\,\widetilde{M}f(x)$. Similarly

$$\int_{B(x,r_j)\setminus B(x,r_{j-1})} |k(x,y)|\,|f(y)|\,d\mu(y) \leq \frac{\mu_{j+1}}{r_{j-1}\mu(B(x,r_{j+1})}\int_{B(x,r_j)} |f(y)|\,d\mu(y)\,.$$

But we know that $r_{j-1} = 3^{-k+j-1}r_k, \mu_{j+1} \leq 36(36)^{\frac{-k+j}{2}}\mu_k$. Hence $\frac{\mu_{j+1}}{r_{j-1}} \leq 36\cdot 3^{k-j+1}6^{-k+j}\frac{\mu_k}{r_k} \leq AM2^{-k+j}$. Therefore,

$$\sum_{j=1}^k \int_{B(x,r_j)\setminus B(x,r_{j-1})} |k(x,y)|\,|f(y)|\,d\mu(y) \leq AM\sum_{j=1}^k 2^{-k+j}\frac{1}{\mu(B(x,r_{j+1})}\int_{B(x,r_j)} |f(y)|\,d\mu(y)\,.$$

The last sum is obviously bounded by $AM\,\widetilde{M}f(x)$. We finally get

$$|(T^r f)(x) - (T^{3R} f)(x)| \leq AM\,\widetilde{M}f(x)\,.$$

Now we need to estimate $(T^{3R}f)(x)$. Consider the average $V_R(x) := \frac{1}{\mu(B(x,R)}\int_{B(x,R)} Tf\,d\mu$ First,

$$|V_R(x)| \leq \frac{\mu(B(x,3R))}{\mu(B(x,R))}\widetilde{M}[Tf](x) \leq 36\widetilde{M}[Tf](x)\,.$$

Second,

$$V_R(x) - (T^{3R}f)(x) = \int_{\mathbb{C}\setminus B(x,3R)} T'[\delta_x - \frac{1}{\mu(B(x,R))}\chi_{B(x,R)}\,d\mu]f\,d\mu -$$

$$\frac{1}{\mu(B(x,R))}\int_{B(x,R)} T[f\chi_{B(x,3R)}]\,d\mu = I + II\,.$$

Here T' denotes the operator with kernel $k(y,x)$.

Estimate of I. Put $\eta = \delta_x - \frac{1}{\mu(B(x,R))}\chi_{B(x,R)}\,d\mu$. All radii greater than $3R$ are M-Ahlfors for μ. This and the fact that $\eta(\mathbb{C}) = 0$ allows us to use the Calderón-Zygmund property of $k(y,x)$ to prove as usual (see [NTV2] for example) that $I \le AM\,\|\eta\|\,\widetilde{M}f(x) \le AM\,\widetilde{M}f(x)$.

Estimate of II. Fix $\beta \in (1,2)$. Let $1/\alpha + \beta = 1$.

$$|II| \le \frac{1}{\mu(B(x,R))}\|\chi_{B(x,R)}\|_{L^\alpha(\mu)}\|T(f\chi_{B(x,3R)})\|_{L^\beta(\mu)} \le \|T\|_\beta \frac{(\int_{B(x,3R)}|f|^\beta\,d\mu)^{\frac{1}{\beta}}}{\mu(B(x,R))^{\frac{1}{\beta}}}\,.$$

Here we abbreviate $\|T\|_\beta := \|T\|_{L^\beta(\mu)\to L^\beta(\mu)}$. We can continue

$$|II| \le \|T\|_\beta \frac{\mu(B(x,9R))^{\frac{1}{\beta}}\,(\widetilde{M}_\beta f)(x)}{\mu(B(x,R))^{\frac{1}{\beta}}} \le$$

$$36^{\frac{1}{\beta}}\|T\|_\beta(\widetilde{M}_\beta f)(x) \le A\|T\|_\beta(\widetilde{M}_\beta f)(x)\,.$$

To estimate $\|T\|_\beta$ via $\|T\|_2$ we need first

Estimate of weak type via $\|T\|_2$.

Lemma (G. David). For any measurable set F and any point $x \in \operatorname{supp}\mu$,

$$T^*\chi_F(x) \le A_1\widetilde{M}[T\chi_F](x) + A_2 M + A_3\|T\|_2\,.$$

Proof. Fix $x \in \operatorname{supp}\mu, r > 0$. Put $\hat{r} = \max[r, \mathcal{R}(x)]$, where $\mathcal{R}(x) := \sup\{r > 0 : \mu(B(x,r)) > Mr\}$. Consider $r_j = 3^j\hat{r}$. Put $\mu_j := \mu(B(x,r_j))$. Let k be the smallest index such that $\mu_k \le 6\mu_{k-1}$. It exists. Otherwise, for every k, $\mu(B(x,\hat{r})) \le 6^{-k}\mu_k \le 2M6^{-k}r_k$. This is because our radii are greater than $\mathcal{R}(x) := \sup\{r > 0 : \mu(B(x,r)) > Mr\}$. We continue with $\mu(B(x,\hat{r})) \le 2M6^{-k}3^k\hat{r} = 2M2^{-k}\hat{r}$. This contradicts the assumption $x \in \operatorname{supp}\mu$. Put $R = r_{k-1}$. We estimate $|(T^r f)(x) - (T^{3R}f)(x)| \le \int_{B(x,\hat{r})\setminus B(x,r)}|k(x,y)|\,|\chi_F(y)|\,d\mu(y) + \sum_{j=1}^k \int_{B(x,r_j)\setminus B(x,r_{j-1})} \ldots$. The first term vanishes if $\hat{r} > \mathcal{R}(x)$. Otherwise it is bounded by

$$\frac{1}{\mathcal{R}(x)}\int_{B(x,\hat{r})}|\chi_F(y)|\,d\mu(y) = \frac{1}{\hat{r}}\int_{B(x,\hat{r})}|\chi_F(y)|\,d\mu(y) \le 2M\,.$$

Similarly

$$\int_{B(x,r_j)\setminus B(x,r_{j-1})}|k(x,y)|\,|\chi_F(y)|\,d\mu(y) \le \frac{\mu_j}{r_{j-1}}\,.$$

But we know that $r_{j-1} = 3^{-k+j-1} r_k, \mu_j \leq 6(6)^{-k+j}\mu_k$. Hence $\frac{\mu_j}{r_{j-1}} \leq 6 \cdot 3^{k-j+1} 6^{-k+j} \frac{\mu_k}{r_k} \leq AM 2^{-k+j}$. Therefore,

$$\sum_{j=1}^{k} \int_{B(x,r_j)\setminus B(x,r_{j-1})} |k(x,y)|\,|\chi_F(y)|\,d\mu(y) \leq AM \sum_{j=1}^{k} 2^{-k+j} \leq AM\,.$$

We finally get

$$|(T^r f)(x) - (T^{3R} f)(x)| \leq AM\,.$$

Now we need to estimate $(T^{3R}\chi_F)(x)$. Consider the average $V_R(x) := \frac{1}{\mu(B(x,R))}\int_{B(x,R)} T\chi_F\,d\mu$ Firstly, by the choice of R, we have

$$|V_R(x)| \leq \frac{\mu(B(x,3R))}{\mu(B(x,R))}\widetilde{M}[T\chi_F](x) \leq 6\widetilde{M}[T\chi_F](x)\,.$$

Second,

$$V_R(x) - (T^{3R} f)(x) = \int_{\mathbb{C}\setminus B(x,3R)} T'[\delta_x - \frac{1}{\mu(B(x,R))}\chi_{B(x,R)}\,d\mu]\chi_F\,d\mu -$$

$$\frac{1}{\mu(B(x,R))}\int_{B(x,R)} T[\chi_{F\cap B(x,3R)}]\,d\mu = I + II\,.$$

Here T' denotes the operator with kernel $k(y,x)$.
Estimate of I. Put $\eta = \delta_x - \frac{1}{\mu(B(x,R))}\chi_{B(x,R)}\,d\mu$. All radii greater than $3R$ are M-Ahlfors for μ. This and the fact that $\eta(\mathbb{C}) = 0$ allows us to use the Calderón-Zygmund property of $k(y,x)$ to prove as usual (see [NTV2] for example) that $I \leq A\,\|\eta\|\,\sup_{\rho\geq R}\frac{\mu(B(x,\rho))}{\rho} \leq AM$.
Estimate of II.

$$|II| \leq \frac{1}{\mu(B(x,R))}\|\chi_{B(x,R)}\|_{L^2(\mu)}\|T(\chi_{F\cap B(x,3R)})\|_{L^2(\mu)} \leq \|T\|_2 \frac{(\int_{B(x,3R)}|\chi_F|^2\,d\mu)^{\frac12}}{\mu(B(x,R))^{\frac12}}\,.$$

We can continue

$$|II| \leq \|T\|_\beta \frac{\mu(B(x,3R))^{\frac12}}{\mu(B(x,R))^{\frac12}} \leq$$

$$6^{\frac12}\|T\|_2 \leq A\|T\|_2\,.$$

The lemma is completely proved.

Now we are ready to repeat the considerations of Theorem 5.1 of [NTV2] (with small modifications due to the fact that μ is a non-uniformly Ahlfors measure).

We are going to prove now that

$$\|T\|_{L^1(\mu)\to L^{1,\infty}} \leq A_1 CM + A_2 C\|T\|_2\,, \qquad (W)$$

where C depend only on Calderón-Zygmund constants of the kernel of T.

Let $\nu \in M(\mathbb{C})$ be a finite linear combination of unit point masses with positive coefficients, i.e.,

$$\nu = \sum_{i=1}^{N} \alpha_i \delta_{x_i}.$$

Theorem 6.

$$\|T\nu\|_{L^{1,\infty}(\mu)} \leqslant (A_1 C M + A_2 C \|T\|_2) \|\nu\|\,.$$

Here there is no problem with the definition of $T\nu$: it is just the finite sum $\sum_{i=1}^{N} \alpha_i K(x, x_i)$, which makes sense everywhere except at finitely many points.

Proof. In this proof $B(x,\rho)$ denotes closed ball, $B'(x,\rho)$ denotes open ball. Without loss of generality, we may assume that $\|\nu\| = \sum_i \alpha_i = 1$ (this is just a matter of normalization). Thus we have to prove that

$\|T\nu\|_{L^{1,\infty}(\mu)} \leqslant A_4$. Fix some $t > 0$ and suppose first that $\mu(\mathbb{C}) > \frac{1}{t}$. Let $B(x_1, \rho_1)$ be the smallest (closed) ball such that $\mu(B(x_1,\rho_1)) \geqslant \dfrac{\alpha_1}{t}$ (since the function $\rho \to \mu(B(x_1,\rho))$ is increasing and continuous from the right, tends to 0 as $\rho \to 0$, and is greater than $\dfrac{1}{t} \geqslant \dfrac{\alpha_1}{t}$ for sufficiently large $\rho > 0$, such ρ_1 exists and is strictly positive).

Note that for the corresponding *open* ball $B'(x_1,\rho_1) := \{y \in \mathbb{C} : \operatorname{dist}(x_1, y) < \rho_1\}$, we have $\mu(B'(x_1,\rho_1)) = \lim_{\rho\to\rho_1-0} \mu(B(x_1,\rho)) \leqslant \dfrac{\alpha_1}{t}$. Since the measure μ is σ-finite and non-atomic, one can choose a Borel set E_1 satisfying

$$B'(x_1,\rho_1) \subset E_1 \subset B(x_1,\rho_1) \qquad \text{and} \qquad \mu(E_1) = \frac{\alpha_1}{t}.$$

Let $B(x_2,\rho_2)$ be the smallest ball such that $\mu(B(x_2,\rho_2) \setminus E_1) \geqslant \dfrac{\alpha_2}{t}$ (since $\mu(\mathbb{C}) > \frac{1}{t}$, the measure of the remaining part $\mathbb{C} \setminus E_1$ is still greater than $\dfrac{1-\alpha_1}{t} \geqslant \dfrac{\alpha_2}{t}$). Again for the corresponding open ball $B'(x_2,\rho_2)$, we have $\mu(B'(x_2,\rho_2) \setminus E_1) \leqslant \dfrac{\alpha_2}{t}$, and therefore there exists a Borel set E_2 satisfying

$$B'(x_2,\rho_2) \setminus E_1 \subset E_2 \subset B(x_2,\rho_2) \setminus E_1 \qquad \text{and} \qquad \mu(E_2) = \frac{\alpha_2}{t}.$$

In general, for $i = 3, 4, \dots, N$, let $B(x_i,\rho_i)$ be the smallest ball such that

$$\mu\Big(B(x_i,\rho_i) \setminus \bigcup_{\ell=1}^{i-1} E_\ell\Big) \geqslant \frac{\alpha_i}{t},$$

and let E_i be a Borel set satisfying

$$B'(x_i,\rho_i) \setminus \bigcup_{\ell=1}^{i-1} E_\ell \subset E_i \subset B(x_i,\rho_i) \setminus \bigcup_{\ell=1}^{i-1} E_\ell \quad \text{and} \quad \mu(E_i) = \frac{\alpha_i}{t}.$$

Put $E := \bigcup_i E_i$. Clearly

$$\bigcup_i B'(x_i,\rho_i) \subset E \subset \bigcup_i B(x_i,\rho_i) \qquad \text{and} \qquad \mu(E) = \frac{1}{t}.$$

Now let us compare $T\nu$ to $t\sum_i \chi_{\mathbb{C}\setminus B(x_i,2\rho_i)} \cdot T\chi_{E_i} =: t\sigma$ outside E. We have

$$T\nu - t\sigma = \sum_i \varphi_i$$

where

$$\varphi_i = \alpha_i T\delta_{x_i} - t\,\chi_{\mathbb{C}\setminus B(x_i,2\rho_i)} \cdot T\chi_{E_i}.$$

Note now that

$$\int_{\mathbb{C}\setminus E} |\varphi_i| d\mu \leqslant \int_{\mathbb{C}\setminus B(x_i,2\rho_i)} \big|T[\alpha_i\delta_{x_i} - t\chi_{E_i} d\mu]\big| d\mu + \int_{B(x_i,2\rho_i)\setminus B'(x_i,\rho_i)} \alpha_i |T\delta_{x_i}| d\mu =: I + \alpha_i II\,.$$

To estimate I, notice that it has the form $\int_{\mathbb{C}\setminus B(x,2\rho)} |T\eta|\, d\mu$ with the measure η supported by $B(x,\rho)$ and $\eta(\mathbb{C}) = 0$. To estimate such an integral we put $\hat{r} := \max[2\rho, R(x)]$ and split $\int_{\mathbb{C}\setminus B(x,2\rho)} |T\eta|\, d\mu = \int_{B(x,\hat{r})\setminus B(x,2\rho)} |T\eta|\, d\mu + \int_{\mathbb{C}\setminus B(x,\hat{r})} |T\eta|\, d\mu =: I_1 + I_2$. The ntegral I_2 is estimated exactly as in Lemma 3.4 of [NTV2] because our measure is $2M$-Ahlfors for disks centered at x with radii larger than $\hat{r}$. Hence $I_2 \le ACM\|\eta\| \le ACM\alpha_i$. On the other hand using the properties of the kernel of T we see that

$$I_1 \le C\min[\frac{1}{2\rho}, \frac{1}{R(x)}]\mu(B(x,\hat{r}))\|\eta\| \le ACM\alpha_i\,.$$

Hence $I \le ACM\alpha_i$.

To estimate II we notice that it has the form $\int_{B(x,2\rho)\setminus B(x,\rho)} |T\delta_x|\, d\mu$. This is almost the same as I_1. Namely, $II \le AC\min[\frac{1}{\rho}, \frac{1}{R(x)}]\mu(B(x,2\rho)) \le \frac{AC\mu(B(x,2\max[R(x),\rho]}{\max[R(x),\rho]}$. This is bounded by ACM because our measure is $2M$-Ahlfors for disks centered at x with radii larger than $R(x)$. Finally $I + \alpha_i II \le ACM\alpha_i$.

Finally we conclude that

$$\int_{\mathbb{C}\setminus E} |T\nu - t\sigma| d\mu \leqslant ACM \sum_i \alpha_i = ACM,$$

and thereby $|T\nu - t\sigma| \leqslant ACMt$ everywhere on $\mathbb{C}\setminus E$, except, maybe, a set of measure $\frac{1}{t}$. To accomplish the proof of the theorem, we will show that for sufficiently large $B = B(C, M, \|T\|_2)$,

$$\mu\{|\sigma| > B\} \leqslant \frac{2}{t}.$$

Then, combining all the above estimates, we shall get

$$\mu\big\{x \in \mathbb{C} \,:\, |T\nu(x)| > (B + ACM)t\big\} \leqslant \frac{4}{t}.$$

We will apply the standard Stein-Weiss duality trick. Assume that the inverse inequality $\mu\{|\sigma| > B\} > \frac{2}{t}$ holds. Then either $\mu\{\sigma > B\} > \frac{1}{t}$, or $\mu\{\sigma < -B\} > \frac{1}{t}$. Assume for definiteness that the first case takes place and choose some set $F \subset \mathbb{C}$ of measure exactly $\frac{1}{t}$ such that $\sigma > B$ everywhere on F. Then, clearly,

$$\int_{\mathbb{C}} \sigma \chi_F d\mu > \frac{B}{t}.$$

On the other hand, this integral can be computed as

$$\sum_i \int_{\mathbb{C}} [T\chi_{E_i}] \cdot \chi_{F \setminus B(x_i, 2\rho_i)} \, d\mu = \sum_i \int_{\mathbb{C}} \chi_{E_i} \cdot [T^{'} \chi_{F \setminus B(x_i, 2\rho_i)}] \, d\mu.$$

Fix a point $x \in E_i \subset B(x_i, \rho_i)$. We will use again the property that $|K(x,y)| \le \frac{1}{R(x)}$.

$$|T^{'}\chi_{F\setminus B(x_i,2\rho_i)}(x) - T^{'}\chi_{F\setminus B(x,\rho_i)}(x)| \leqslant$$

$$\int_{B(x_i,2\rho_i)\setminus B(x,\rho_i)} |K(y,x)| \, d\mu(y) \le \frac{AC\mu(B(x, 3\max[\rho_i, R(x)]))}{\max[\rho_i, R(x)]} \le ACM,$$

because all disks centered at x and of radii greater than $R(x)$ are $2M$-Ahlfors, and therefore for every $x \in E_i \cap \operatorname{supp} \mu$,

$$|T^{'}\chi_{F\setminus B(x_i,2\rho_i)}(x)| \leqslant (T^{'})^{\sharp}\chi_F(x) + ACM \leqslant 2 \cdot A\,\widetilde{M}T^{'}\chi_F(x) + ACM$$

according to Guy David's lemma. Hence

$$\int_{\mathbb{C}} \sigma \chi_F d\mu \leqslant ACM\mu(E) + 2 \cdot A \int_{\mathbb{C}} \chi_E \cdot \widetilde{M}T^{'}\chi_F d\mu.$$

But the first term equals $\dfrac{ACM}{t}$ while the second one does not exceed

$$2 \cdot 3^n \, \|\chi_E\|_{L^2(\mu)} \|\widetilde{M}T^{'}\chi_F\|_{L^2(\mu)} \leqslant \frac{2 \cdot 3^n}{t} \|\widetilde{M}\|_{L^2(\mu)\to L^2(\mu)} \|T^{'}\|_{L^2(\mu)\to L^2(\mu)}.$$

Recalling that $\|T^{'}\|_{L^2(\mu)\to L^2(\mu)} = \|T\|_{L^2(\mu)\to L^2(\mu)}$, we see that one can take

$$B = ACM + 2 \cdot 3^n \, \|\widetilde{M}\|_{L^2(\mu)\to L^2(\mu)} \|T\|_{L^2(\mu)\to L^2(\mu)}$$

to get a contradiction. Since the norm $\|\widetilde{M}\|_{L^2(\mu)\to L^2(\mu)}$ is bounded by some absolute constant (the constant in the Marcinkiewicz interpolation theorem), we are done.

References

[CJS] R. R. Coifman, P. W. Jones and S. Semmes, Two elementary proofs of the L^2 boundedness of Cauchy integrals on Lipschitz curves, J. Amer. Math. Soc. 2, (1989), no. 3, 553–564.

[DM] G. David, P. Mattila, Removable sets for Lipschitz harmonic functions in the plane. Rev. Mat. Iberoamericana 16 (2000), no. 1, 137–215.

[D1] G. David, Unrectifiable 1-sets have vanishing analytic capacity. Rev. Mat. Iberoamericana 14 (1998), no. 2, 369–479.

[D2] G. David, Analytic capacity, Cauchy kernel, Menger curvature, and rectifiability. Harmonic analysis and partial differential equations (Chicago, IL, 1996), 183–197, Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 1999.

[D3] G. David, Analytic capacity, Calderón-Zygmund operators, and rectifiability. Publ. Mat. 43 (1999), no. 1, 3–25.

[Du] J. Doudziak, Vitushkin's conjecture for removal sets, Springer, 2010, 272 pp.

[GJ] J. B Garnett and P. W. Jones, BMO from dyadic BMO, Pacific J. Math., 99, (1982), no. 2, 351–371.

[L] J-C. Léger, Menger curvature and rectifiability. Ann. of Math. (2) 149 (1999), no. 3, 831–869.

[Ma] P. Mattila, Cauchy singular integrals and rectifiability in measures of the plane. Adv. Math. 115 (1995), no. 1, 1–34.

[MaP] P. Mattila, D. Preiss, David Rectifiable measures in R^n and existence of principal values for singular integrals. J. London Math. Soc. (2) 52 (1995), no. 3, 482–496.

[MTV] J. Matheu, X. Tolsa, J. Verdera, The planar Cantor sets of zero analytic capacity and the local $T(b)$-Theorem. Preprint, Univ. Auton. Barcelona, 2001, pp. 1–12.

[M] M. Melnikov, Analytic capacity: a discrete approach and the curvature of measure. (Russian) Mat. Sb. 186 (1995), no. 6, 57–76.

[MP] P. Mattila, P.V. Paramonov, On geometric properties of harmonic Lip_1-capacity, Pac. J. math., 171 (1991), No. 2, 469–491.

[MV] M.S. Melnikov, J. Verdera, A geometric proof of L^2 boundedness of the Cauchy integral on lipschitz graph. Intern. Math. Res. Notices, 7 (1995), 325–331.

[MMV] P. Mattila, M. Melnikov, J. Verdera, The Cauchy integral, analytic capacity, and uniform rectifiability. Ann. of Math. (2) 144 (1996), no. 1, 127–136.

[NTV1] F. Nazarov, S. Treil, A. Volberg, Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces, Intern. Math. Res. Notices, 15 (1997), 703–726.

[NTV2] F. Nazarov, S. Treil, A. Volberg, Weak type estimates and Cotlar's inequality for Calderòn-Zygmund operators on nonhomogeneous spaces, Intern. Math. Res. Notices, 9 (1998), 463–487.

[NTV3] F. Nazarov, S. Treil, A. Volberg, The Tb-theorem on non-homogeneous spaces. Acta Math. 190 (2003), no. 2, 151–239.

[To] X. Tolsa, Analytic capacity, the Cauchy transform, and non-homogeneous Calderón-Zygmund theory, Manuscript, 2012, http://www.mat.uab.cat/∼ xtolsa/llibreweb.pdf.

[XT1] X. Tolsa, *Painlevé's problem and analytic capacity*, Proceedings of the 7th International Conference on Harmonic Analysis and Partial Differential Equations El Escorial, Madrid (Spain), June 21–25, 2004, Collect. Math. (2006), 89–125.

[XT2] X. Tolsa, *Bilipschitz maps, analytic capacity, and the Cauchy integral.* Ann. of Math. (2) 162 (2005), no. 3, 12431304.

[XT3] X. Tolsa, *Painlevé's problem and the semiadditivity of analytic capacity*, Acta Math. 190:1 (2003), 105–149.

[Vo] A. Volberg, Calderón–Zygmund capacities and operators on non-homogeneous spaces. CBMS lecture series, AMS, v. 100, 2003, pp. 165.