V4B1 - Nonlinear Partial Differential Equations I

Christoph Thiele

WS 2012/13

1 The Poisson equation: basics and examples

1.1 Classical and weak Poisson equation

Let $n \geq 2$ and let $\Omega \subset \mathbf{R}^n$ be an open set. A twice continuously differentiable function $u : \Omega \to \mathbf{R}$ and a continuous function $f : \Omega \to \mathbf{R}$ are said to satisfy the classical Poisson equation if

$$\Delta u(x) = f(x)$$

for every $x \in \Omega$ where

$$\Delta u = \sum_{i=1}^{n} D_i^2 u$$

is the Laplace operator.

Given such a pair u and f and ϕ a smooth function with compact support in Ω, then

$$(\Delta u, \phi) = (f, \phi)$$

where the bracket $(., \phi)$ denotes integrating against the function ϕ or more generally evaluating a distribution at ϕ. By partial integration, using the compact support of ϕ, we obtain successively

$$-(Du, D\phi) = (f, \phi)$$

where D denotes the gradient operator and the bracket of two vector fields is the integral of the inner product, and

$$(u, \Delta\phi) = (f, \phi)$$

This is the weak Poisson equation with testing function ϕ.

Conversely, let two functions u, f have regularity as above and satisfy the weak Poisson equation for every smooth function ϕ with compact support in Ω. We intend to show u and f satisfy the classical Poisson equation. By a translation argument it suffices to assume $0 \in \Omega$ and show $\Delta u(0) = f(0)$. Moreover, it suffices to show for any given $\epsilon > 0$ that $|\Delta u(0) - f(0)| \leq \epsilon$. There exists a nonnegative smooth compactly supported function ϕ with

$$\int_{\mathbf{R}^n} \phi(x)\ dx = 1.$$

By scaling, that is by considering functions of the form $\phi_\lambda(x) = \lambda^n \phi(\lambda x)$ we may assume ϕ is supported in the ball $B_\delta(0)$ of radius δ about 0 where δ is sufficiently small so that

$|f(x) - f(0)| \leq \epsilon/2$ and $|\Delta u(x) - \Delta u(0)| \leq \epsilon/2$ for all $x \in B_\delta(0)$. Then we have

$$|\Delta u(0) - f(0)| = |\int (\Delta u(0) - f(0))\phi(x)\, dx|$$

$$\leq \int |f(0) - f(x)|\, \phi(x)\, dx + |\int (\Delta u(x) - f(x))\, \phi(x)\, dx| + \int |\Delta u(x) - \Delta u(0)|\, \phi(x)\, dx$$

The middle term vanishes as we can see by partial integration and the weak Poisson equation. The other two terms are bounded by $\epsilon/2$ each. This completes the proof that u and f satisfy the classical Poisson equation.

We shall therefore not distinguish between weak and classical Poisson equation and say that two distributions u and f satisfy Poisson's equation if they satisfy the weak Poisson equation.

Non-uniqueness of Poisson's equation

Given a distribution u, there is a unique distribution f satisfying the weak Poisson equation

$$(u, \Delta\phi) = (f, \phi)$$

Indeed, since Δ is a continuous map on the space of test functions, the left hand side defines a distribution

$$\phi \to (u, \Delta\phi)$$

and this distribution is the unique f satisfying the weak Poisson equation. In fact this distribution is denoted by Δu.

Given a distribution f, assume we have two distributions u_1 and u_2 satisfying the weak Poisson equation

$$(u_1, \Delta\phi) = (f, \phi)$$

$$(u_2, \Delta\phi) = (f, \phi)$$

Then we have for $u = u_2 - u_1$

$$(u, \Delta\phi) = 0$$

This is called the weak Laplace equation for $u := u_1 - u_2$., It implies (homework) that $u_1 - u_2$ is harmonic and in particular smooth and satisfies the classical Laplace equation

$$\Delta u(x) = 0$$

for every $x \in \Omega$. Thus the ambiguity in solving the Poisson equation rests in an undetermined additive harmonic function.

The theory of harmonic functions suggests additional constraints to force the solution to Poisson's equation to be unique.

For example, assume f is supported in the ball $B_r(0)$ and there exists a continuous function u on the closed ball $\overline{B_r(0)}$ that satisfies the Poisson equation on the open ball. Then given any continuous function on the boundary $\partial B_r(0)$ there is unique continuous function v on the closed ball satisfying the Poisson equation in the open ball and coinciding with the specified function on the boundary. This result rests of the existence and uniqueness of a continuous function on the closed ball which satisfies the Laplace equation in the open ball (homework) and coincides with a given continuous function on the boundary of the ball.

Note that this example of existence and uniqueness shows that a fair amount of regularity theory is required to achieve existence and uniqueness of solutions.

In the immediate sequel we shall focus on existence and regularity theory for solutions of Poisson's equation and put aside uniqueness questions. Since harmonic functions are very regular, namely infinitely often differentiable and the values of the derivative being controlled by averages over balls (e.g maximum principle), regularity theory of solutions to the Poisson equation is in the interior of the domain Ω.

The Poisson equation for the characteristic function of a ball

We seek a solution for the Poisson equation where f is the characteristic function of a ball. By symmetries we reduce the discussion to the case of the unit ball about the origin.

Elaborating on the symmetries, we define the following symmetries:

1. Translation. If f is a test function and $y \in \mathbf{R}^n$, define

$$T_y f(x) = f(x - y)$$

If f is a distribution, define

$$(T_y f, \phi) = (f, T_{-y}\phi)$$

2. Dilation of homogeneous degree $\alpha \in \mathbf{R}$: If f is a test function and $\lambda > 0$, define

$$\Lambda_\lambda^\alpha f(x) = \lambda^\alpha f(\lambda^{-1}x)$$

If f is a distribution, define

$$(\Lambda_\lambda^\alpha f, \phi) = (f, \Lambda_\lambda^{-n-\alpha}\phi)$$

3. Rotation: If f is a test function and A and orthogonal $n \times n$, define

$$O_A f(x) = f(A^T x)$$

If f is a distribution, define

$$(O_A f, \phi) = (f, O_{A^T}\phi)$$

Exercise: Show that uf f is a test function and we identify it with the distribution $\phi \to (f, \phi)$, then both definitions of the symmetries coincide.

Note that if u, f satisfy the Poisson equation in a domain Ω, then $T_y u$ and $T_y f$ satisfy the Poisson equation on the domain $\{x + y : x \in \Omega\}$. Moreover, $\tilde{u} = \Lambda_\lambda^{n-2} u$ and $\tilde{f} = \Lambda_\lambda^n f$ satisfy the Poisson equation in the domain $\{\lambda x : x \in \Omega\}$. To see the latter, we calculate

$$\Delta \tilde{u} = \sum_i D_i D_i \Lambda_\lambda^{n-2} u$$

$$\lambda^{-2} \Lambda_\lambda^{n-2} \sum_i D_i D_i u = \Lambda_\lambda^n f = \tilde{f}$$

Finally, $\tilde{u} = O_A u$ and $\tilde{f} = O_A f$ satisfy the Poisson equation on the domain $\{Ax : x \in \Omega\}$. We leave the verification of this as an exercise.

Using translation and dilation invariance, we may reduce the problem of finding a solution of the Poisson for f the characteristic function of a ball to the case when the ball has center at the origin and radius 1. By rotation symmetry, since f is roattion invariant, we may seek a solution u that is rotationally invariant.

We thus consider $f = 1_{B_1(0)}$. To avoid technical bifurcations we assume $n \geq 3$. Define a function

$$u(x) = -|x|^{2-n}$$

if $|x| \geq 1$ and

$$u(x) = a + b|x|^2$$

if $x \leq 1$, where we assume that the real numbers a and b are chosen so that both the value of and the gradient of the two defining functions for u coincide on the sphere $|x| = 1$. Since the derivative of $-r^{2-n}$ is $(n-2)r^{1-n}$ and the derivative of br^2 is $2br$, we see that the choice $b = (n-2)/2$ will assure continuity of the radial derivative across the unit sphere. All tangential derivatives are zero on the sphere. Since both defining functions are constant on the sphere, we can choose a to achieve continuity of u across the sphere. We will not need to explicitly calculate a. We then have for $|x| > 1$:

$$\Delta u(x) = 0$$

Namely,

$$D_i|x| = \frac{x_i}{|x|}$$

and

$$D_i u(x) = (n-2)x_i|x|^{-n}$$

$$D_i^2 u(x) = (n-2)(|x|^{-n} - n|x_i|^2|x|^{-n-2})$$

Summing over i and using $\sum_i |x_i|^2 = |x|^2$ gives $\Delta u(x) = 0$. Moreover, for $|x| < 1$

$$\Delta u(x) = 2nb = n(n-2)$$

since $D_i|x|^2 = 2x_i$ and $D_i^2|x|^2 = 2$. Using continuously differentiability of u we obtain by partial integration

$$(u, D_n D_n \phi) = \int_{\mathbf{R}^{n-1}} \int_{\mathbf{R}} -D_n u(x) D_n \phi(x)\, dx_n\, dx'$$

where x' denotes the vector $(x_1, \ldots, x_{n-1})$. Since $D_n u$ is piecewise smooth and continuous, we may do one more partial integration

$$(u, \mathbf{D}_n D_n \phi) = \int_{\mathbf{R}^{n-1}} \int_{\mathbf{R}} D_n D_n u(x) \phi(x)\, dx_n\, dx'$$

where $D_n D_n u$ is defined everywhere except at the boundary of $B_1(0)$ and the integral is understood in an L^1 sense. Similarly we may argue for the other second partial derivatives and we obtain

$$(u, \Delta\phi) = \int_{\mathbf{R}^n} \Delta u(x)\phi(x)\ dx$$

with the right hand side in an L^1 sense and hence

$$(u, \Delta\phi) = (n(n-2)1_{B_1(0)}, \phi)$$

Dividing by $n(n-2)|B_1(0)|$ shows that $\frac{1}{n(n-2)|B_1(0)|}u$ and $\frac{1}{B_1(0)}1_{B_1(0)}$ satisfy the Poisson equation.

To obtain the Poisson equation for the ball of radius r about the origin, we apply scaling as above to consider

$$f_r = \frac{1}{|B_r(0)|}1_{B_r(0)}$$

and

$$\Gamma_r(x) = \frac{-1}{n(n-2)|B_1(0)|}|x|^{2-n}$$

for $|x| \geq r$ and

$$\Gamma_r(x) = \frac{1}{n(n-2)|B_r(0)|}(ar^2 + b|x|^2)$$

for $|x| \leq r$. and observe that Γ_r and f_r solve the Poisson equation.

Note that both f and the Hessian $D_iD_j\Gamma_r$ are bounded measurable functions. In the sequel we will see that even for continuous compactly supported f it is not always true that the Hessian $D_iD_j\Gamma_r$ is bounded and measurable.

We turn to the case $n = 2$. The previous calculation does not hold since the degenerate term $n - 2$ occurs in the denominator in the calculation. Focusing on the ball $B_1(0)$ we make the ansatz

$$u(x) = \log|x|$$

for $|x| \geq 1$ and

$$u(x) = a + b|x|^2$$

for $|x| \leq 1$. Since the derivative of $\log(r) = r^{-1}$ we need to set $b = 2^{-1}$ to obtain continuity of radial derivatives. Beginning with these observations the further calculations proceed analoguosly to the above and we obtain the following. If we define

$$\Gamma_r(x) = \frac{1}{2|B_1(0)|}\log|x|$$

for $|x| \geq r$ and

$$\Gamma_r(x) = \frac{1}{2|B_r(0)|}(ar^2 + b|x|^2)$$

for $|x| \leq r$ then Γ_r and $\frac{1}{B_r(0)}1_{B_r(0)}$ solve the Poisson equation.

The fundamental solution of Poisson's equation

We again restrict attention first to $n \geq 3$. Now let f be in $L^1(\mathbf{R}^n)$ and assume f vanishes outside a bounded open set Ω. Define for $x \neq 0$

$$\Gamma(x) = \frac{-1}{n(n-2)|B_1(0)|}|x|^{2-n}$$

Then Γ is locally in L^1 (that is its restriction to any bounded open set is in L^1), in particular we verify

$$\int_{B_r(0)} \Gamma(x)\,dx \le C \sum_{k\ge 0} \int_{2^{-k-1}r\le|x|\le 2^{-k}r} (2^{-k}r)^{2-n}\,dx$$
$$\le C\sum_{k\ge 0}(2^{-k}r)^n(2^{-k}r)^{2-n} \le Cr^2$$

Here C is some constant depending on n that may change from occurrence to occurrence. Hence the convolution

$$\Gamma * f(x) := \int \Gamma(x-y)f(y)\,dy$$

is defiend almost everywhere and a local L^1 function (this is a consequence of Fubini's theorem).

Theorem 1 *If $f \in L^1(\omega)$ for some bounded open set Ω, then $\Gamma * f$ and f satisfy the Poisson equation on Ω.*

Proof: It is natural to compare Γ with Γ_r. Note that also

$$\int_{B_r(0)} \Gamma_r(x)\,dx \le C(2^{-k}r)^{-n}\sum_{k\ge 0}\int_{2^{-k-1}r\le|x|\le 2^{-k}r}(r^2+(2^{-k}r)^2)dx$$
$$\le Cr^2$$

Moreover $\Gamma-\Gamma_r$ is supported in $B_r(0)$. Hence Γ_r converges to Γ in L^1 as $r\to 0$ and we have

$$(\Gamma * f, \Delta\phi) = \lim_{r\to 0}(\Gamma_r * f, \Delta\phi) = \lim_{r\to 0}(f, \Gamma_r * (\Delta\phi))$$

Now to identify the value of the smooth function $\Gamma_r * (\Delta\phi)$ we note by symmetry of Γ_r

$$\Gamma_r * (\Delta\phi)(x) = (T_x\Gamma_r, \Delta\phi) = (\Gamma_r, \Delta T_{-x}\phi)$$
$$= (\frac{1}{B_r(0)}1_{B_r(0)}, T_{-x}\phi) = \frac{1}{B_r(0)}(1_{B_r(0)} * \phi)$$

Now the latter converges uniformly to ϕ, hence

$$(\Gamma * f, \Delta\phi) = \lim_{r\to 0}\frac{1}{|B_r(0)|}(f, 1_{B_r(0)} * \phi)) = (f,\phi)$$

This proves the theorem. □

Taking the gradient of Γ outside the origin gives

$$D\Gamma(x) = \frac{1}{n|B_1(0)|}x|x|^{-n}$$

This function is still locally L^1, in particular we have

$$|\int_{B_r(0)}\nabla\Gamma(x)\,dx|$$
$$\le C\sum_{k\ge 0}\int_{2^{-k-1}r\le|x|\le 2^{-k}r}(2^{-k}r)^{1-n}\,dx \le Cr$$

Theorem 2 *If $f \in L^1(\omega)$ for some bounded open set Ω, then $\Gamma * f$ is continuously differentiable and equals $D\Gamma * f$.*

Proof: This is a reprise of the previous argument. We have

$$|\int_{B_r(0)} \nabla\Gamma_r(x)\,dx| \le Cr$$

and $D(\Gamma - \Gamma_r)$ is supported in $B_r(0)$ $f * D\Gamma_r$ converges uniformly to $f * D\Gamma$. Since $f * \Gamma_r$ is continuously differentiable with gradient $f * D\Gamma_r$ we see that $f * \Gamma$ is continuously differentiable with gradient $f * D\Gamma$. □

Note that this argument fails for the Hessian matrix of second partial derivatives, which is homogeneous of degree $-n$ and hence $D_iD_j\Gamma$ is no longer locally in L^1.

Indeed, already the Poisson equation $\Delta u = f$ suggests that the second partial derivatives of u need not be continuous if f is merely locally L^1.

We leave the modifications of the above calculations in case $n = 2$ as an exercise.

1.2 The Poisson equation for 1_Q

We next discuss the solution $u = \Gamma * 1_Q$ of the Poisson equation for f the characteristic function of the cube $Q = I_1 \times I_2 \times \ldots \times I_n$ and we will draw some conclusions.

For specifity we assume each interval is half open of the form $[a_i, b_i)$ though this choice will not be of particular importance. We sall be particularly interested in studying the Hessian of $\Gamma * 1_Q$.

Set

$$u(x) = \int_Q \Gamma(x-y)\,dy$$

let $3Q$ the cube with same center as Q but 3 times the sidelength of Q.

Lemma 1 *The function u is smooth in $\mathbf{R}^n \setminus \partial Q$ and satisfies for every i, j, k*

1. *If $x \in 3Q$*

$$D_iD_ju(x) \le C(1 + |\log(\frac{\text{dist}(x,\partial Q)}{l(Q)}|)$$

2. *If $x \notin 3Q$*

$$D_iD_ju(x) \le C|\frac{\text{dist}(x,Q)}{l(Q)}|^{-n}$$

3. *If $x \in 3Q$, then*

$$D_kD_iD_ju(x) \le Cl(Q)^{-1}(\frac{\text{dist}(x,\partial Q)}{l(Q)})^{-1}$$

4. *If $x \in 3Q$, then*

$$D_kD_iD_ju(x) \le Cl(Q)^{-1}(\frac{\text{dist}(x,Q)}{l(Q)})^{-n-1}$$

Proof: We assume $n > 2$, the case $n = 2$ merely requiring notational changes. Assume $x \notin \partial Q$ and assume ϵ is small enough so that $B_{2\epsilon}(x)$ does not intersect Q.

Let $\phi : \mathbf{R}_{\geq 0} \to \mathbf{R}$ be a smooth function that vanishes in $[0, \epsilon/2)$ and is constant 1 on $[\epsilon, \infty)$. Such a function can for example be obtained by integrating a positive smooth function supported in $(\epsilon/2, \epsilon)$.

We define a smooth variants of the previously defined Γ_r, for simplicity of notation we use the same symbol. Thus define $\Gamma_r(x) = \Gamma(x)\phi(|x|)$ Note that by rotation symmetry of Γ we have for the integrable function $D_n\Gamma$:

$$\int_{B_r(0)} D_n\Gamma(x)\,dx = 0$$

More precisely, we only need reflection symmetry under reflection of the n-the component, namely

$$(x_1, \ldots, x_{n-1}, x_n) \to (x_1, \ldots, x_{n-1}, -x_n)$$

Which leaves Γ invariant but changes sign of $D_n\Gamma$. Since $B_r(0)$ is invariant under the reflection, the above integral is equal to its negative and thus vanishes. By the same token, we have for any i:

$$\int_{B_r(0)} D_i\Gamma_r(x)\,dx = 0$$

Since 1_Q is constant on the ball of radius ϵ about any point $y \in B_\epsilon(x)$, we obtain

$$D_n\Gamma * 1_Q(y) = D_n\Gamma_\epsilon * 1_Q(y)$$

But the latter is smooth sine Γ_ϵ is smooth. (use theorems about differentiation under the integral sign). Using the corresponding argument for the other partial derivatives gives that $\Gamma * 1_Q$ is smooth away from ∂Q.

To prove the first estimate, Pick $x \in 3Q$ and let ϵ be half the distance of x to ∂Q. Then

$$|D_iD_j(\Gamma * 1_Q)(x)| = |D_iD_j(\Gamma_\epsilon * 1_Q)(x)| \leq \|D_iD_j\Gamma_\epsilon\|_{L^1(5Q)}$$

But since $D_iD_j\Gamma$ is homogeneous of degree $-n$ we have

$$\int_{5Q \setminus B_{2\epsilon}(x)} D_iD_j\Gamma_\epsilon(x)\,dx \leq \sum_{\epsilon \leq 2^k \leq 10l(k)} \int_{2^{k-1} \leq |x| \leq 2^k} C|x|^{-n}\,dx$$

$$\leq \sum_{\epsilon \leq 2^k \leq 10l(k)} C \leq C(1 + |\log(\epsilon/l(Q))|)$$

Since also

$$\int_{B_{2\epsilon}(x)} |D_iD_j\Gamma_\epsilon(x)|\,dx \leq$$

$$\int_{B_{2\epsilon}(x) \setminus B_{\epsilon/2}(x)} |D_iD_j\Gamma(x)\eta(x)| + |D_i\Gamma(x)D_j\eta(x)| + |D_j\Gamma(x)D_i\eta(x)| + |\Gamma(x)D_iD_j\eta(x)|\,dx \leq C$$

We obtain the first estimate.

To prove the second estimate we simply integrate the estiamte

$$D_iD_j\Gamma(y) \leq C|y|^{-n}$$

and the fact that on the domain of integration for $\Gamma * 1_Q(x)$ the argument of $D_iD_j\Gamma_{l(Q)/2}$ is approximately $\text{dist}(x, Q)$.

To prove the third and foruth estimate we proceed similarly, using that

$$D_iD_jD_k\Gamma$$

is homogeneous of degree $-n-1$. $\square$

The estimate near the boundary of the cube can be improved. Recall that ∂Q is the set of all points in the closed cube for which at last one coordinate x_i is in the boundary of I_i. Let $\partial\partial Q$ denote the set of points in ∂Q for which at least two coordinates x_i, x_j are in the boundary of I_i and I_j respectively. Then

Lemma 2 *The function u satisfies for every i, j, k.*

1. *If $x \in 3Q$*

$$D_iD_ju(x) \leq C(1 + |\log(\frac{\text{dist}(x, \partial\partial Q)}{l(Q)})|)$$

2. *If $x \in 3Q$*

$$D_iD_jD_ku(x) \leq Cl(Q)^{-1}(\frac{\text{dist}(x, \partial\partial Q)}{l(Q)})^{-1}$$

The improvement rests in a stronger use of the full rotation symmetry of the kernel Γ. Assume that $\text{dist}(x_j, \partial I_j)$ is minimzed for $j = n$. Let $x' = (x_1, \ldots, x_{n-1})$ and $Q' = I_1 \times \ldots \times I_{n-1}$ and let ϵ be half the distance from x' to $\partial Q'$. Then ϵ is comparable to the distance x to $\partial\partial Q$. We write for $i \neq n$

$$\int D_i\Gamma(y)1_Q(x-y)\,dy$$

$$= \int_{\mathbf{R}} \int_{R^{n-1}} D_i\Gamma(y)1_Q(x-y)\,dy'dy_n$$

Now the stroger use of the symmetry is that if i is not equal to n, then for fixed y_n the functions Γ and Γ_ϵ are symmetric under changing sign of the i-th component of x'. Thus we obtain

$$\int_{\mathbf{R}} \int_{R^{n-1}} D_i\Gamma(y)1_Q(x-y)\,dy'dy_n$$

$$\int_{\mathbf{R}} \int_{R^{n-1}} D_i\Gamma_\epsilon(y)1_Q(x-y)\,dy'dy_n$$

Note that ϵ may be much larger than the distance x_n to ∂I_n. Now if either i or j is different from n, since the partial derivatives commute, we may argue as in the proof of the previosu lemma to obtain the new inequalities of the present lemma. If $i = j = n$ we do not have the same argument, but we may deduce the estimates using that Δu is equal to 1_Q away from the boundary of Q and $\Delta\Gamma = \sum_{i=1}^n D_iD_i\Gamma$ and we already know that all but the terms $D_iD_i\Gamma$ for $i \neq n$ on the right hand side satisfy the desired estimates.

We now note that the above estimate can not further be much improved:

Lemma 3 *There are points arbitrarily close to any point in $\partial\partial Q$ such that*

$$\mathbf{max}_{i,j}|D_iD_ju(x)| \geq c(1 + |\log(\frac{\text{dist}(x, (\partial\partial Q)}{l(Q)})|)$$

Proof: By translation and scaling assume that $I_i = [0, 1)$ for all i and consider the points $x_i = x_j = 0$ in $\partial\partial Q$. Note that

$$D_i D_j \Gamma(y) = C \frac{y_i y_j}{|y|^{n+2}}$$

is positive in the set $y_i \leq 0$ and $y_j \leq 0$. Hence for $x_i \leq 0$ and $x_j \leq 0$ the integral

$$D_i D_j \Gamma * 1_Q(x) = \int D_i D_j \Gamma(y) 1_Q(x-y)\, dy$$

has a nonnegative integrand, and hence the estimates in the previous lemmata are sharp. □

Note that despite f being bounded, the second partials of $u = \Gamma * f$ are not bounded. By an application of the unform boudnedness principle, Banach Steinhaus, assuming continuity of f does not help.

Note first that if Q is a cube as in the previous lemma and f_ϵ is a continuous approximation to 1_Q, say positive and bounded by 1 that coincides with 1_Q outside an ϵ neighborhood of ∂Q, then $D_i D_j G * f_\epsilon)$ tends to $D_i D_j \Gamma * 1_Q$ at every point x with $x_i, x_j < 0$. Hence there is no C such that for any x with $x_i, x_j < 0$ we have

$$D_i D_j G * f_\epsilon(x) \leq C$$

for all $\epsilon > 0$.

Lemma 4 *There exists a continuous compactly supported function f and i, j such that $\Gamma * f$ is not twice continuosuly differentiable.*

Assume to get a contradiction that $\Gamma * f$ is twice continuously differentiable for every continuous f supported in $Q = [0, 1]^n$ satisfying $\|f\|_\infty \leq 1$. Note that this set of functions f is a complete metric space unde the distance function $d(f, g) = \sup_x |f(x) - g(x)|$. Then since $\Gamma * f$ decays outside $3Q$, by an estimate as in the above lemmata, and hence it does attain its maximum and thus is bounded. Thus the collection of linear functionals $\Lambda_y : f \to D_i D_j \Gamma * f(y)$ for every $y \in \mathbf{R}_n$ has the property that for each f we have $\Lambda_y f$ is bounded as y varies. By Banach Steinhaus, $\Lambda_y f$ is uniformly bounded, hence there is C such that

$$\|D_i D_j \Gamma * f\|_\infty \leq C$$

But this contradicts the observation discussed above. □

2 Dyadic cubes

Unless explicitly stated otherwise, all intervals in this section are of the form $[a, b)$ with $a < b$, thus they contain the left boundary point a but not the right boundary point b.

Definition 1 *The set of all intervals of the form $[2^k m, 2^k(m+1))$ is called the standard dyadic grid. An interval in the standard dyadic grid is called a standard dyadic interval.*

The length $l(I)$ of a standard dyadic interval I is an integer power of 2, namely 2^k. The number k is called the scale of the standard dyadic interval.

Lemma 5 *1. The class of standard dyadic intervals is invariant under the dilation $x \to 2x$ of the real line, more specifically,*

$$[a, b) \to [2a, 2b)$$

is a self-bijection of the set of standard dyadic intervals.

2. *The translation $x \to x + 1$ on the real line, i.e.*

$$[a, b) \to [a + 1, b + 1)$$

is a bijection of the set of standard dyadic intervals of scale at most 0.

3. *For every $k \in \mathbf{Z}$ and every $x \in \mathbf{R}$, x is contained in a unique standard dyadic interval of scale k.*

4. *For every standard dyadic interval I, the left and right halves defined by*

$$I_l = [a, \frac{a+b}{2}), \quad I_r = [\frac{a+b}{2}, b) \quad .$$

are again standard dyadic intervals and called the children of the (parent) interval I.

5. *If I and J are two standard dyadic intervals, then either $I \cap J = \emptyset$ or one of the two intervals is contained in the other.*

6. *Let $\mathbf{I}$ be some collection of standard dyadic intervals with scale bounded above,*

$$\sup_{I \in \mathbf{I}} l(I) < \infty \ ,$$

and let $\mathbf{I}_{\max}$ be the collection of maximal intervals in $\mathbf{I}$ with respect to set inclusion. Then the intervals in $\mathbf{I}_{\max}$ are pairwise disjoint and cover the union of all intervals in $\mathbf{I}$.

Proof:
1) This bijection amounts to the obvious bijection (translation)

$$(k, n) \to (k + 1, n)$$

in the parameter range $\mathbf{Z}^2$ of dyadic intervals.

2) This transformation amonts to the obvious bijection (nonlinear shearing)

$$(k, n) \to (k, n + 2^{-k})$$

of the parameter space $\mathbf{Z} \times \mathbf{Z}_{\leq 0}$ of dyadic intervals of scale at most 0.

3) By 1), it suffices to prove this for $k = 0$. Let n be the greatest integer smaller or equal to x (this exists by the archimedean principle and is unique by the total order of the integers) then $n \leq x < n + 1$ and n is the only integer satisfying these two inequalities.

4) By 1) and 2) it suffices to prove this for $I = [0, 1)$, where it follows from direct inspection.

5) Without loss of generality we may assume $|I| \geq |J|$. If I and J have the same scale, the claim follows from 3). Inductively, assume the claim is proven for $|I| = 2^k|J|$

and consider a pair of intervals with $|I| = 2^{k+1}|J|$. We need to show that if there exists $x \in J \cap I$. then $J \subset I$. Let $x \in J \cap I$. Let $\tilde{I}$ be child of I that contains x. By induction $J \subset \tilde{I}$ and hence $J \subset I$. This completes the proof.

6) Given any two different intervals in $\mathbf{I^{max}}$, since neither is contained in the other by maximality, they are disjoint by 5). For each point in the union of intervals in $\mathbf{I}$, there is a maximal dyadic interval containing this point (use the upper bound on the scale). Hence the point is in the union of elements in $\mathbf{I^{max}}$.

□

Definition 2 *A collection* $\mathbf{D}$ *of intervals is called a (not necessarily standard) dyadic grid if it satisfies the following properties*

1. *For each* $x \in \mathbf{R}$ *and* $k \in \mathbf{Z}$ *the point* x *is contained in some interval of length* 2^k *of the collection.*

2. *Any two intervals of the collection are disjoint or one is contained in the other.*

Given a dyadic grid, an interval of the grid is called a dyadic interval.

Note that one can sharpen the first statement in this definition. Namely, every point x is contained in exactly one interval of length 2^k. To see this, consider two intervals of the same length containing x. By the second property, one of the intervals is contained in the other, and since the intervals have the same length they are equal.

For each interval $[a, b)$ in a grid both children are in the grid as well. Namely, let I be the interval in the grid of half the length of $[a, b)$ that contains the point a and is guaranteed to exists by the first property. By the second property we have $I \subset [a, b)$. Hence I does not contain any points less than a and thus a is the left endpoint of I. Hence I is of the from $[a, (a + b)/2)$ and thus the left child of $[a, b)$. Now let J be the interval of the same length as I that contains $(a + b)/2$. Then J has to be disjoint from I by the second property, and hence $(a + b)/2$ is the left endpoint of J. Hence J is the right child of I.

Note that in a dyadic grid every interval is the child of some parent. Namely, let I be an interval in the grid and pick $x \in I$. Then x is contained in a unique interval of length $2l(I)$, Since I intersects at least one of the children of that interval, it must be equal to that child.

Lemma 6 *For every interval* $[a, b)$ *and every sequence* $d_n : \mathbf{N}_0 \to \{1, -1\}$ *there is a unique dyadic grid containing the interval* $[a, b)$ *and having the property that if* $[a_0, b_0) = [a, b)$ *and* $[a_{n+1}, b_{n+1})$ *denotes the parent of* $[a_n, b_n)$ *then* $[a_n, b_n)$ *is the left child of* $[a_{n+1}, b_{n+1})$ *if* $d_n = -1$ *and it is the right child of* $[a_{n+1}, b_{n+1})$ *if* $d_n = 1$*. The dyadic grid is a translate of the standard dyadic grid if and only if* $d_n = d_{n+1}$ *for all* n *larger than some* n_0*.*

Proof: Exercise □

Since each dyadic grid contains exactly one interval of length 1 containing 0, we may parameterize all dyadic grids by data $(a, d_0, d_1, \ldots)$ where $a \in (-1, 0]$ is the left endpoint of that interval and $d_n \in \{-1, 1\}$ is the sequence relative to the interval $[a, a + 1)$ as described in the above lemma.

Lemma 7 *For every point* x *there is a dyadic grid such that* x *is in the closed middle third of every dyadic interval containing* x*. For any two points* $x \neq y$ *there is a dyadic grid such that every dyadic interval of length larger than* $6|x - y|$ *contains both* x *and* y *in its middle half.*

Proof: For the first part, pick the interval $[x - 2/3, x + 1/3)$ and pick $d_n = (-1)^n$. Then the grid described by the previous lemma satisfies the desired properties. Namely, the interval of length 2^k in the grid containing x is of the form

$$[x - (2/3)2^k, x + (1/3)2^k)$$

for even k and of the form

$$[x - (1/3)2^k, x + (2/3)2^k)$$

for odd k, and one verifies immediately that for two consecutive k the corresponding intervals have a common endpoint and thus are related as parent and child.

For the second part, pick the point z the midpoint of x and y and consider the grid constructed in the first part with respect to z. □

Definition 3 *A dyadic grid in $\mathbf{R}^n$ consists of an n-tuple $(\mathbf{D}_1, \dots, \mathbf{D}_n)$ of dyadic grids of the real line. A dyadic cube with respect to this grid is a cube of the form $I_1 \times \ldots \times I_n$ with $I_j \in \mathbf{D}_j$ We define $l(Q) = l(I_i)$ where the right hand side is independent of i, and we define $|Q| = l(Q)^n$.*

Lemma 8 *Given a dyadic grid in $\mathbf{R}^n$, we have the following properties:*

1. *Every point of $\mathbf{R}^n$ is contained in exactly one cube of the grid of any given sidelength 2^k.*
2. *Given any two dyadic cubes, either they are disjoint or one is contained in the other.*
3. *Each dyadic cube is the union of 2^n pairwise disjoint dyadic cubes of half the sidelength. These are called the children of the cube. Moreover, each dyadic cube has a parent.*
4. *Given any collection $\mathbf{Q}$ of dyadic cubes with sidelength uniformly bounded above, then the set $\mathbf{Q_{max}}$ of maximal dyadic cubes in $\mathbf{Q}$ with respect to set inclusion consists of pairwise disjoint cubes and covers the same set as the original collection $\mathbf{Q}$ of cubes.*

2.1 Remarks on Lebesgue measure in $\mathbf{R}^n$ in terms of dyadic cubes

Fix a dyadic grid in $\mathbf{R}^n$

Exercise 1 *Given a dyadic cube Q_0 and some collection $\mathbf{Q}$ of dyadic cubes covering Q_0, i.e.*

$$Q_0 \subset \bigcup_{Q \in \mathbf{Q}} Q$$

then

$$|Q_0| \leq \sum_{Q \in \mathbf{Q}} |Q|$$

Hint: consider first the easier case that there exists a k such that $2^k|Q| \geq |Q_0|$ for every $Q \in \mathbf{Q}$. This case can be settled by induction on k. To pass to the general case, use a compactness argument.

Definition 4 *The Lebesgue outer measure of a set* $E \subset \mathbf{R}^n$ *is defined to be*

$$\mu(E) := \inf_{\mathbf{Q}} \sum_{Q \in \mathbf{Q}} |Q|$$

where the infimum is taken over oll coverings of E *by dyadic cubes, that is over all collections* $\mathbf{Q}$ *of dyadic cubes such that* $E \subset \bigcup_{Q \in \mathbf{Q}} Q$.

Note that Lebesgue outer measure is countably subadditive: If

$$E = \bigcup_{n=0}^{n} E_n$$

then

$$\mu(E) \leq \sum_{n=0}^{n} \mu(E_n)$$

This ultimately rests on the observation that given a covering for each E_n, then the union of these coverings is a covering of E.

Note further that for any k,

$$\mu(E) := \inf \sum_{Q \in \mathbf{Q}} |Q|$$

where the infimum is take over all collections $\mathbf{Q}$ such that $l(Q) \leq 2^k$ for all $Q \in \mathbf{Q}$.

Only the case of finite measure $\mu(E)$ is interesting, and by dilation by a power of 2 (possibly changing the grid) one may assume $\mu(E) \leq 1$: Then we can prove the above by induction: if $k = 0$ the statement is trivial and if the statement is true for k then replacing each cube in $\mathbf{Q}$ by its children proves the statement for $k + 1$.

Definition 5 *A set* E *is measurable in* $\mathbf{R}^n$ *if for every set* A *we have that*

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

Note that a dyadic cube Q is measurable. Namely, let A be a close to optimal covering of A we may assume all cubes in the covering have sidelength less than $l(Q)$. Then the covering falls into two parts, thos cubes contained in Q and those disjoint from Q. These are two coverings of $A \cap Q$ and $A \cap Q^c$.

Clearly the collection of measurable sets is closed under taking complements (by definition) and countable unions (easy exercise) and hence is a sigma algebra.

To see that open sets are measurable, we introduce the concept of Whitney decomposition, that is also of independent interest.

Definition 6 (Whitney decomposition) *Let* $\mathbf{D}$ *be a dyadic grid in* $\mathbf{R}^n$ *and let* Ω *be an open set in* $\mathbf{R}^n$. *The Whitney decomposition of* Ω *consists of the collection* $\mathbf{Q_{max}}$ *of all maximal dyadic cubes in* $\mathbf{Q}$, *where* $\mathbf{Q}$ *is the set of dyadic cubes* $Q \in \mathbf{D}$ *such that* $3Q \subset \Omega$.

Each Whitney cube is contained in Ω, hence the Whitney cubes cover a subset of Ω. They cover all of Ω, since each point in Ω has an ϵ ball about it in Ω, and thus any dyadic cube of sidelength 2^k with $2^k n < \epsilon$ containing x satisfies $3Q \subset \Omega$ and thus $\mathbf{Q}$ covers Ω and thus $\mathbf{Q_{max}}$ covers Ω. Moreover, $\mathbf{Q_{max}}$ is a disjoint cover of Ω.

In particular, open sets are measurable.

Whitney decompositions have further nice properties:

Lemma 9 *If Q, Q' are two adjacent cubes in a Whitney decomposition, then the ratio of sidelength of Q and Q' is at most 2.*

Proof: If Q and Q' are adjacent and $|Q'| < |Q|$, then $3Q' \subset 3Q$. If the ratio of sidelength of Q and Q' is more than 2, then there is a distinct cube Q'' adjacent to Q with $Q' \subset Q''$ and also $3Q'' \subset \Omega$. By maximality, Q' cannot be in the Whitney decomposition. □

Exercise 2 *If Q is a Whitney cube, then the distance from any point of Q to the complement of Ω is between $l(Q)$ and $\sqrt{n}l(Q)$.*

If one introduces Lebesgue measure via dyadic grids as above, it is a natural task to show that the Lebesgue measure does not depend on the choice of dyadic grid. To do so, one observes that for some fixed grid, every (non-dyadic) cube Q is a measurable set and has measure equal to $l(Q)^n$. From there one deduces that the definition of outer measure is independent of the grid.

2.2 Remarks on $L^\infty(\mathbf{R}^n)$, $L^1(\mathbf{R}^n)$, $M^1(\mathbf{R}^n)$ via dyadic cubes/martingales

Fix a dyadic grid in $\mathbf{R}^n$ and let $\mathbf{D}$ denote the set of all dyadic cubes in the grid.

We consider functions

$$f : \mathbf{D} \to \mathbf{R}$$

Note that the set $\mathbf{D}$ is countable, so the space of functions on $\mathbf{R}$ has the same cardinality as the space $\mathbf{R}$ itself, and stricly smaller cardinality than the space of functions from $\mathbf{R}^n$ to $\mathbf{R}$.

If $x \in \mathbf{R}^n$ is a point, we have for each $k \in \mathbf{Z}$ a unique cube $Q_{k,x}$ of sidelength 2^k containing x, and we may consider the limit

$$\lim_{k \to -\infty} f(Q_{k,x})$$

if it exists. If this limit exists for every x, then we obtain a function on $\mathbf{R}^n$. However by the above remarks on cardinality only relatively few functions can be generated in this way.

While we will not in general ask that this limit at x exists for every x, we do have the intuition that the functions $f : \mathbf{D} \to \mathbf{R}$ are trying to describe functions or generalized functions on $\mathbf{R}$. The value $f(Q)$ is then some approximation, or some average to the value of f inside the cube. It is thus natural to make a consistency assumption on f, namely that $f(Q)$ is the average of $f(Q')$ where Q' runs over the children of f.

Definition 7 (Martingale) *Given a grid $\mathbf{D}$, A function $f : \mathbf{D} \to \mathbf{R}$ is called a martingale if for every cube $Q \in \mathbf{D}$ we have*

$$f(Q)|Q| = \sum_{Q' \subset Q : 2l(Q') = l(Q)} f(Q')|Q'|$$

Note that the condition in this definition can also be written as

$$f(Q) = 2^{-n} \sum_{Q' \subset Q : 2l(Q') = l(Q)} f(Q') \ ,$$

which makes it plain that $f(Q)$ is the average of $f(Q')$ as Q' runs through the 2^n children of Q.

Exercise 3 (Truncation) *Let f be a martingale and let $\mathbf{Q}$ be some collection of pairwise disjoint cubes in $\mathbf{D}$. Then we can define the truncation $f_{\mathbf{Q}}$ of f as*

1. *If $Q \subset Q'$ for some $Q' \in \mathbf{Q}$, then this Q' is unique and*

$$f_{\mathbf{Q}}(Q) := f(Q')$$

2. *If $Q \not\subset Q'$ for any $Q' \in \mathbf{Q}$, then*

$$f_{\mathbf{Q}}(Q) := f(Q)$$

Prove that this truncation is a martingale.

Given a martingale f and a scale $k \in \mathbf{Z}$, we let f_k or $E_k f$ be the truncated martingale with respect to the collection

$$\mathbf{D}_k := \{Q \in \mathbf{D} : l(Q) = 2^k\}$$

.

Definition 8 *The space $L^\infty(\mathbf{D})$ is the space of all martingales f for which there exists $M > 0$ such that $\|f\|_\infty := \sup_{Q \in \mathbf{Q}} |f(Q)| < \infty$.*

Exercise 4 *The space $L^\infty(\mathbf{D})$ is a Banach space space (complete normed space).*

We will see below that there is a natural identification of the space $L^\infty(\mathbf{D})$ with the classical space $L^\infty(\mathbf{R})$ as martingales.

We also aim to obtain descriptions of the classical spaces $L^1(\mathbf{R}^n)$. The definitions are somewhat more subtle in this situation, we begin with:

Definition 9 *The space $M_d(\mathbf{D})$ is the space of all martingales f such that*

$$\|f\|_1 := \sup_{\mathbf{Q}} \sum_{Q \in \mathbf{Q}} |f_Q||Q| < \infty$$

where the supremum runs over all collections $\mathbf{Q}$ of pairwise disjoint dyadic cubes.

The following exercise shows that it suffices to test the collections of cubes of fixed scale

Exercise 5 *For any martingale f we have*

$$\|f\|_1 := \sup_k \sum_{Q: l(Q) = 2^k} |f(Q)||Q|$$

Note also that the quantity

$$\sum_{Q: l(Q) = 2^k} |f(Q)||Q|$$

is monotone in k, because

$$\sum_{Q: l(Q) = 2^k} |f(Q)||Q| = \sum_{Q: l(Q) = 2^k} | \sum_{Q' \subset Q, 2l(Q') = l(Q)} f(Q')|Q'||$$

$$\leq \sum_{Q: l(Q) = 2^k} \sum_{Q' \subset Q, 2l(Q') = l(Q)} |f(Q')||Q'| = \sum_{Q: l(Q) = 2^{k-1}} |f(Q')||Q'|$$

Lemma 10 *The space $M_d(\mathbf{D})$ is a complete normed linear space with the norm $\|f\|_1$.*

Proof: If f and g are martingales in $M_d(\mathbf{D})$ and c is a real number, then $f+g$ and cf are martingales in $M_d(\mathbf{D})$ and one observes $\|f+g\|_1 \le \|f\|_1 + \|g\|_1$ and $\|cf\|_1 = |c|\|f\|_1$. If $\|f\|_1 = 0$ then we have for every dyadic cube Q, viewed as a disjoint collection by itself, $|f(Q)||Q| \le \|f\|_1 = 0$ and hence $f(Q) = 0$. This makes $M_d(\mathbf{D})$ a normed space.

To see completeness, we proceed as follows. If $f^{(n)}$ is a Cauchy sequence of martingales, then for every Q the sequence $f^{(n)}(Q)$ is Cauchy and thus has a limit $f(Q)$. The data $f(Q)$ in itself defines a martingale, since finite averages commute with the limit.

To see that f is itself in $M_d(\mathbf{D})$, choose a subsequence of $f^{(n_m)}$ such that $\|f^{(n_{m+1})} - f^{(n_m)}\|_1 \le 2^{-m}$. Then for finite disjoint collections $\mathbf{Q}$ of cubes

$$\sum_{Q\in\mathbf{Q}} |f(Q)||Q| = \sum_{Q\in\mathbf{Q}} \lim_{M\to\infty} |f^{(0)}(Q) + \sum_{m=0}^{M} (f^{(n_{m+1})}(Q) - f^{(n_m)}(Q))|$$

$$\le \|f^{n_0}\|_1 + \sum_m \|f^{(n_{m+1})} - f^{(n_m)}\|_1 \le C < \infty$$

Similarly one sees that $f^{(n)}$ converges to f in norm. □

Note that for any finite Borel measure ν in $\mathbf{R}^n$ we may define a martingale $f \in M_d(\mathbf{D})$ by $f(Q) = \nu(Q)/|Q|$ Many elements, but not every element in $M_d(\mathbf{D})$ is of this form. For example cosndier the standard dyadic grid and the Diac mearure at the origin and let f be defined as above. Then the cube $[0,1)^n$ satisfies $f(Q) = 1$ while all other cubes Q' of sidelength 1 whose boundary contains 0 to the origin satisfy $f(Q') = 0$. Obviously this assignment is very particular to our chose of half open cubes. One cans ee that the martingale defined by

$$f(\prod_{i=1}^{n}[a_i, b_i)) = \nu(\prod_{i=1}^{n}(a_i, b_i]) / \prod_{i=1}^{n}|b_i - a_i|$$

for ν the Dirac measure at the origin is not of the form $\nu(Q)/|Q|$ for any Borel measure ν.

This suggests the following definition

Definition 10 *The space $M(\mathbf{D})$ is the space of all martingales $f \in M_d(\mathbf{D})$ such that for each $\epsilon > 0$ and dyadic cube Q there exists $\delta > 0$ such that for every collection $\mathbf{Q}$ of pairwise disjoint dyadic cubes contained in Q with $\partial(Q') \not\subset Q$ for all $Q' \in \mathbf{Q}$ with*

$$\sum_{Q\in\mathbf{Q}} |Q| \le \delta$$

we have

$$\sum_{Q\in\mathbf{Q}} |f(Q)||Q| \le \epsilon$$

The idea behind this definition is that no mass is put on the bad boundary, as represented by the bad cubes $Q' \subset Q$ with $\partial(Q') \not\subset Q$.

A related and more drastic definition is to avoid mass on any set of zero Lebesgue measure.

Definition 11 *The space $L^1(\mathbf{D})$ is the space of all martingales $f \in M(\mathbf{D})$ such that for every $\epsilon > 0$ there is a $\delta > 0$ such that for every collection $\mathbf{Q}$ of pairwise disjoint dyadic cubes with*

$$\sum_{Q \in \mathbf{Q}} |Q| \leq \delta$$

we have

$$\sum_{Q \in \mathbf{Q}} |f(Q)||Q| \leq \epsilon$$

In classical terms, $L^1(\mathbf{D})$ are the measures that are absolutely continuous with respect to Lebesgue measure.

Clearly $L^1(\mathbf{D})$ is a subspace of $M(\mathbf{D})$.

Lemma 11 *The spaces $M(\mathbf{D})$ and L^1 are closed subspaces of $M_d(\mathbf{D})$.*

Proof: We elaborate the proof in case $L^1(\mathbf{D})$, the case $M(\mathbf{D})$ being very similar.

Let $f^{(n)}$ be a sequence of martingales in $L^1(\mathbf{D})$ and assume that the sequence converegs to an element f in $M_d(\mathbf{R}^n)$. We need to prove that $f \in L^1(\mathbf{D})$.

For given ϵ we may choose an n so that

$$\|f^{(n)} - f\|_1 \leq \epsilon/2$$

We may then choose a δ so that for every disjoint collection $\mathbf{Q}$ of dyadic cubes with

$$\sum_{Q \in \mathbf{Q}} |Q| \leq \delta$$

we have

$$\sum_{Q \in \mathbf{Q}} |f^{(n)}(Q)||Q| \leq \epsilon/2$$

But then we also have

$$\sum_{Q \in \mathbf{Q}} |f(Q)||Q| \leq \sum_{Q \in \mathbf{Q}} |f(Q) - f^{(n)}||Q| + \sum_{Q \in \mathbf{Q}} |f^{(n)}(Q)||Q| \leq \epsilon/2 + \epsilon/2 = \epsilon$$

□

We need some auxiliary observations to simplify future calculations.

Exercise 6 *Let f be a martingale, let $\mathbf{Q}$ be a collection of pairwise disjoint cubes. Then the truncation $f_{\mathbf{Q}}$ satisfies*

$$\|f_{\mathbf{Q}}\|_\infty \leq \|f\|_\infty$$

$$\|f_{\mathbf{Q}}\|_1 \leq \|f\|_1$$

Moreover, if f is in $L^1(\mathbf{D})$ or $M(\mathbf{D})$, then so is $f_{\mathbf{Q}}$.

All of the above depended on a grid $\mathbf{D}$. In what follows we shall relate likewise spaces for different grids, and see that these spaces have a meaning independent of the grid.

Lemma 12 *If f is a martingale in $M(\mathbf{R}^n)$, then the set of all x for which*

$$\lim_{k\to\infty} f(Q_{k,x})$$

does not converge has Lebesgue measure zero. Here $Q_{k,x}$ is the cube of sidelength k which contains x.

When the limit as in this lemma converges, we call x a Lebesgue point of f. The lemma states that almost every point of $\mathbf{R}^n$ is a Lebesgue point of f.

Proof: First we argue that $f(Q_{k,x})$ is bounded above for almost every x. Pick $M > 0$ and consider the set of x such that $f(Q_{k,x})$ is not bounded above by M. This set is contained in the set of all cubes Q such that $f(Q) > M$. Let $\mathbf{Q}$ be the collection of maximal such cubes and let E be the set covered by $\mathbf{Q}$. Then

$$\mu(E) = \sum |Q| \leq M^{-1} \sum_{\mathbf{Q}} |f(Q)||Q| \leq M^{-1}\|f\|_1$$

Hence the measure of this set is bounded by $M^{-1}\|f\|_1$. By intersection the set of x such that $f(Q_{k,x})$ is not bounded above by any positive number has measure 0.

Likewise the set where $f(Q_{k,x}))$ is not bounded below by any number has measure zero, and hence this sequence is bounded fopr almost every x.

It then suffices to show that on the set where $f(Q_{k,x})$ is bounded above and below by M and $-M$ the sequence $f(Q_{k,x})$ converges almost everywhere. Note that a bounded sequence s is convergent if and only if for any pair of rational numbers $a < b$ there are at most finitely many crossings from a to b, i.e. there are only finite sequences $k_0, k_1, k_2, \dots$ with $s(i) \leq a$ for odd i and $s(i) \geq b$ for even i.

Let $-M \leq a < b \leq M$. There are two cases to consider, $a < 0$ or $b > 0$. W.l.o.g. we assume $b > 0$, the other case being similar.

Let $\mathbf{Q}_0$ be the set of maximal dyadic cubes with $f(Q) > b$ and we do not have $|f(Q')| > M$ for any Q' containing Q, and let E_0 be the union of these cubes. The set E_0 is finite since

$$\sum_{Q\in\mathbf{Q}_0} |Q| \leq b^{-1} \sum_{Q\in\mathbf{Q}_0} |f(Q)||Q| \leq b^{-1}\|f\|_1$$

Recursively define $\mathbf{Q}_{m+1}$ to be the set of maximal dyadic cubes Q inside E_m such that

1. if n is odd, then $f(Q) > b$ and we do not have $|f(Q')| > M$ for any Q' larger than M
2. if n is even then $f(Q) > b$ and we do not have $|f(Q')| > M$ for any Q' containing Q.

Let E_{m+1} be the union of the cubes in $\mathbf{Q}_{n+1}$. The sets E_m are clearly nested. Clearly infinite sequences of crossings between a and b can only occur in the intersection of all sets E_m or in the union of cubes satisfying $|f_Q| \geq M$. We thus have to show that the intersection of sets E_m has measure zero.

We show that the measure of these sets is bounded by a geometric decreasing sequence. For this it suffices to prove that for odd m and each $Q \in \mathbf{Q}_m$ we have

$$\sum_{Q'\in\mathbf{Q}_{m+1}:Q'\subset Q} |Q'| \leq (1-c)|Q|$$

for some $c > 0$ depending only on M, b and a. Cover the set

$$Q \setminus \bigcup_{Q' \in \mathbf{Q}_{m+1} : Q' \subset Q} Q'$$

by a collection Q' of cubes such that

$$\sum_{Q' \in \mathbf{Q}_{m+1} : Q' \subset Q} |Q'| + \sum_{Q' \in \mathbf{Q}'} |Q'| \leq |Q| \epsilon$$

This can be done by measurability of the union of $\mathbf{Q}_{m+1}$. Let F be the set covered by $\mathbf{Q}'$, we may assume that $\mathbf{Q}'$ is the set of maximal dyadic cubes contained in F.

Consider a parent Q'' of a cube $Q' \in \mathbf{Q}'$. The cube Q'' contains a cube in $\mathbf{Q}_{m+1}$ and thus satisfies $|f(Q'')| > -M$. The same holds for all those children of Q'' which are not in $\mathbf{Q}'$. Since by the martingale property

$$f(Q'')|Q''| = \sum_{Q' \text{ child of } Q''} f(Q')|Q'|$$

we have

$$\sum_{Q' \text{ child of } Q'', Q' \in \mathbf{Q}'} f(Q')|Q'| = f(Q'')|Q''| - \sum_{Q' \text{ child of } Q'', Q' \notin \mathbf{Q}'} f(Q')|Q'|$$

$$\geq -CM \sum_{Q' \text{ child of } Q'', Q' \in \mathbf{Q}'} |Q'|$$

where C may depend on the dimension n

We have by the martingale property:

$$0 = \sum_{Q' \subset Q, l(Q') = 2^k} (f(Q') - f(Q))|Q'|$$

Using the martingale property again to pass to larger cubes gives

$$0 = \sum_{Q' \subset Q : Q' \in \mathbf{Q}_{m+1}} (f(Q') - f(Q))|Q'| + \sum_{Q' \in \mathbf{Q}'} (f_{Q'} - f_Q)|Q'|$$

$$\geq (b - a) \sum_{Q' \subset Q : Q' \in \mathbf{Q}_{m+1}} |Q'| - CM \sum_{Q' \in \mathbf{Q}'} |Q'|$$

Hence

$$(b - a) \sum_{Q' \subset Q : Q' \in \mathbf{Q}_{m+1}} |Q'| \leq 2M \sum_{Q' \in \mathbf{Q}'} |Q'|$$

$$(b - a)|Q| \leq [2M + (b - a)] \sum_{Q' \in \mathbf{Q}'} |Q'|$$

$$\frac{b - a}{2M + (b - a)} |Q| \leq \sum_{Q' \in \mathbf{Q}'} |Q'|$$

And then

$$(1-\frac{b-a}{2M+(b-a)})|Q| \geq \sum_{Q'\subset Q:Q'\in\mathbf{Q}_{m+1}} |Q'|$$

This is what we planned to show. Hence $|E_m|$ shrinks to zero, hence the sequence $f(Q_{k,x})$ converges almost everywhere in the set where $f(Q_{k,x})$ is bounded by $|M|$ This proves the lemma. □

Lemma 13 *Let $f \in L^1(\mathbf{R})$ and let C be some constant. Then the following are equivalent:*

1. *$f(Q) \leq M$ for all Q*
2. *We have $\lim_{k\to\infty} f(Q_{k,x}) \leq M$ for every Lebesgue point x of f except possibly a set of measure zero.*

Proof: It is clear that 1) implies 2), we therefore turn attention to the converse. Assume we have $f(Q) > M+\epsilon$ for some Q. Since almost every point in Q is a Lebesgue point, and each Lebesgue point in Q is contained in a cube $Q' \subset Q$ with $f(Q') < M+\epsilon/2$, we find a finite collection $\mathbf{Q}$ of such cubes $Q' \subset Q$ which cover a set E of measure at least $(1-\delta)|Q|$. Let $\mathbf{Q'}$ be the collection of maximal dyadic cubes contained in $Q\setminus E$, by finiteness of $\mathbf{Q}$ this set is finite and covers $Q\setminus E$. We then have by the martingale property

$$\sum_{Q'\in\mathbf{Q}} f(Q')|Q'| + \sum_{Q'\in\mathbf{Q'}} f(Q')|Q'| = f(Q)|Q|$$

$$\begin{aligned}\sum_{Q'\in\mathbf{Q'}} f(Q')|Q'| &= \sum_{Q'\in\mathbf{Q'}} f(Q)|Q'| + \sum_{Q'\in\mathbf{Q}} (f(Q)-f(Q'))|Q'| \\ &\geq M\sum_{Q'\in\mathbf{Q'}} |Q'| + (\epsilon/2)\sum_{Q'\in\mathbf{Q}} |Q'|\end{aligned}$$

Now for δ small enough the first term on the right hand side is negligible compared to the second, hence

$$\sum_{Q'\in\mathbf{Q'}} f(Q')|Q'| \geq (\epsilon/4)\sum_{Q'\in\mathbf{Q}} |Q'| \geq (\epsilon/8)|Q|$$

However

$$\sum_{Q'\in\mathbf{Q'}} |Q'| \leq \delta|Q|$$

which is a contradiction to the L^1 property for δ small enough. □

Note that this lemma does not hold with $L^1(\mathbf{D})$ replaced by $M(\mathbf{D})$, and neither does the following corollary:

Lemma 14 *Assume we have two martingales f, g in $L^1(\mathbf{D})$, such that we have the limit $\lim_{k\to 0} f(Q_{k,x}) = \lim_{k\to 0} g(Q_{k,x})$ on almost every common Lebesgue point. Then $f = g$.*

Proof: Consider $f-g$, then by assumption $\lim_{k\to 0}(f-g)(Q_{k,x}) = 0$ on almost every common Lebesgue point, and hence by the previous Lemma with $M=0$ we have $(f-g)(Q) = 0$ for all Q. □.

Lemma 15 *Let E be a measurable set of finite measure. Then*

$$f(Q) := \mu(E \cap Q)|Q|^{-1}$$

defines a martingale f in $L^1(\mathbf{R}^n)$. We have

$$\|f\|_1 = \mu(E)$$

and on almost every Lebesgue point x the limit $\lim_{k\to-\infty} f(Q_{k,x})$ is equal to $1_E(x)$.

Proof: Clearly for any disjoint collection $\mathbf{Q}$ of dyadic cubes we have

$$\sum_{Q\in\mathbf{Q}} |f(Q)||Q|^{-1} = \sum_{Q\in\mathbf{Q}} \mu(E\cap Q) \le \mu(E)$$

and hence $\|f\|_1 \le \mu(E)$. Similarly,

$$\sum_{Q\in\mathbf{Q}} |f(Q)||Q|^{-1} = \sum_{Q\in\mathbf{Q}} \mu(E\cap Q) \le \sum_{Q\in\mathbf{Q}} |Q|$$

and hence $f \in L^1(\mathbf{D})$. Next we show that $\lim_{k\to-\infty} f(Q_{k,x}) = 1$ for almost every Lebesgue point x in E. Since $f(Q) \le 1$ for every Q, it suffices to show that for every $b < 1$ the set F of Lebesgue points in E where $\lim_{k\to-\infty} f(Q_{k,x}) < b$ has measure 0. Cover E by a collection $\mathbf{Q}$ of cubes Q such that $\sum_{Q\in\mathbf{Q}} |Q| \le |E| + \epsilon$ and let $E' = \bigcup_{Q\in\mathbf{Q}} Q$. Let $\mathbf{Q'}$ be some finite collection of cubes Q' contained in E' such that $f(Q) < b$. And let $\mathbf{Q''}$ be the maximal cubes contained in E' but disjoint from all cubes in $\mathbf{Q'}$. Then

$$\mu(E) + \epsilon \ge \sum_{Q\in\mathbf{Q}} |Q| = \sum_{Q'\in\mathbf{Q'}} |Q'| + \sum_{Q''\in\mathbf{Q''}} |Q''|$$

$$\ge \sum_{Q'\in\mathbf{Q'}} (1-b)|Q'| + \sum_{Q'\in\mathbf{Q'}} f(Q')|Q'| + \sum_{Q''\in\mathbf{Q''}} f(Q'')|Q''|$$

$$= \sum_{Q'\in\mathbf{Q'}} (1-b)|Q'| + \sum_{Q\in\mathbf{Q}} f(Q)|Q| = \sum_{Q'\in\mathbf{Q'}} (1-b)|Q'| + \mu(E)$$

Hence

$$\sum_{Q'\in\mathbf{Q'}} |Q'| \le \epsilon/(1-b)$$

Hence we also have

$$\sum_{Q'\subset E' : f(Q) < b'} |Q'| \le \epsilon/(1-b)$$

which shows that the set of Lebesgue points in E with limit at most b has measure at most $\epsilon/(1-b)$. Since ϵ was arbitrary, we have $\mu(F) = 0$.

Similarly we show that the limit on almost all Lebesgue points outside E is 0.

□

Exercise 7 *Let F denote the set of finite linear combinations*

$$\sum_{m=1}^{n} a_m 1_{E_m}$$

with measurable sets E_m and real numbers a_m. Define the martingale

$$f_Q = \sum_{m=1}^{n} a_m \mu(Q \cap E_m)|Q|^{-1}$$

Then the martingale converges almost everywhere to $\sum_{m=1}^{n} a_m 1_{E_m}$, in particular if two formal linear combinations coincide almost everywhere, then they yield the same martingale.

Lemma 16 *The finite linear combinations of characteristic functions of sets of finite measure are dense in $L^1(\mathbf{D})$.*

Proof: We use that the set of bounded functions is dense in $L^1(\mathbf{D})$ (exercise). Let $f \in L^1(\mathbf{D})$. Pick small ϵ and let E_m be the set of Lebesgue points of f with

$$\epsilon m < \lim_{k \to -\infty} f(Q_{k,x}) \leq \epsilon(m+1)$$

This set is measurable (exercise) Let

$$g_\epsilon = \sum_m \epsilon m 1_{E_m}$$

since f is bounded, all but finitely many summands in this formally infinite sum vanish. We then have $\lim_{\epsilon \to 0} \|f - g_\epsilon\|_1 = 0$ (exercise) This proves the lemma.

□

Note that the lemma identifies a dense subclass in $L^1(\mathbf{R}^n)$ that can be characterized as functions almost everywhere independently of the grid. This can be used to identify dense subspaces of $L^1(\mathbf{D})$ and $L^1(\mathbf{D}')$ for two different grids and show that these subspaces are canonically isomorphic. Hence $L^1(\mathbf{D})$ and $L^1(\mathbf{D}')$ are canonically isomorphic.

3 Haar basis

Fix a dyadic grid in $\mathbf{R}^n$. Let Q be a dyadic cube and let $\mathbf{Q}_Q$ denote the set of its children, there are 2^n cubes in $\mathbf{Q}_Q$.

Consider the space H_Q of all martingales f such that $f(Q') = 0$ if $Q \not\subset Q$ or $Q' = Q$ and $f(Q') = f(Q'')$ if $Q' \subset Q''$ and $Q'' \in \mathbf{Q}_Q$. It is clear that this martingale is determined by the 2^n values $f(Q'')$ with $Q'' \in \mathbf{Q}_Q$. Moreover, by the martingale property for the cube Q, we have a one dimensional constraint

$$\sum_{Q'' \in \mathbf{Q}_Q} f(Q'') = 0$$

and hence H_Q is at most $2^n - 1$ dimensional.

It is not hard to see that H_Q is exactly $2^n - 1$ dimensional. We construct a specific basis for this space. We parameterize the children

$$Q' = I'_1 \times \ldots \times I'_n$$

of

$$Q = I_1 \times I_n$$

by tuples $q = (q_1, \ldots, q_n)$ where $q_j = 0$ if I'_j is the left child of I_j and $q_J = 1$ if I'_j is the right child of I_j. Then for each further tuple $i = (i_1, \ldots, i_n)$ we define the Haar martingale $h_{Q,i} \subset H_Q$ by

$$h_{Q,i}(Q') = \prod_{j=1}^{n} (-1)^{i_j q_j}$$

provided the tuple h is not constant equal to 0.

We have the following identity

$$\sum_{Q' \in \mathbf{Q}_Q} h_{Q,i}(Q') h_{Q,i'}(Q') = 0$$

if $i \neq i'$ and

$$\sum_{Q' \in \mathbf{Q}_Q} h_{Q,i}(Q') h_{Q,i'}(Q') = 0$$

Exercise 8 *Let f be a martingale which is a finite linear combination of haar martingales*

$$f = \sum_{Q \in \mathbf{Q}} \sum_i f_{Q,i} h_{Q,i}$$

Then

$$f_{Q,i}|Q| = \sum_{Q' \in \mathbf{Q}_Q} f(Q') h_{Q,i}(Q')|Q'|$$

The corefficient $f_{Q,i}$ in the above lemma is called the Haar coefficient of f. It is defined for every martingale f by

$$f_{Q,i}|Q| = \sum_{Q' \in \mathbf{Q}_Q} f(Q') h_{Q,i}(Q')|Q'|$$

Note that for a Haar martingale $h_{Q,i}$ evfery point is a Lebesgue point. We may thus identify the Haar martimngale with a function that is defined everywhere, and called a Haar function.

4 Hölder continuity

Definition 12 *A function $f \in L^1(\mathbf{R})$ is called α-Hölder continuous if there is a constant C such that for every dyadic grid and avere Q in the grid we have*

$$\|f - f_{\{Q\}}\|_1 \leq C|Q| l(Q)^\alpha$$

Here $f_{\{}Q\}$ denotes the truncation of f with respect to the set $\{Q\}$.

Note that if we interpret f as function almost every where, then $\|f - f_{\{(Q)}\|_1$ is equal to

$$\int_Q |f(x) - (\frac{1}{|Q|} \int_Q f(y)\, dy)|\, dx$$

Lemma 17 *The following are equivalent for a function $f \in L^1(\mathbf{D})$*

1. *The function f is α- Hölder*
2. *There exists a C such that for every dyadic grid and every Haar coefficient*

$$|f_i(Q)| \leq Cl(Q)^\alpha$$

3. *Every point is a Lebesgue point of f and for any $x, y \in \mathbf{R}^n$ we have*

$$|f(x) - f(y)| \leq C|x - y|^\alpha$$

There exists a C such that for every dyadic grid and every dyadic cube

$$\|f - f_{\{(Q)\}}\|_\infty \leq Cl(Q)^\alpha$$

Proof: We have

$$f_{Q,i}|Q| = \sum_{Q' \in \mathbf{Q}_Q} f(Q')h_{Q,i}(Q')|Q'|$$

$$\leq \sum_{Q' \in \mathbf{Q}_Q} |f(Q')||Q'| = \sum_{Q' \in \mathbf{Q}_Q} |(f - f_{\{Q\}})(Q')||Q'| \leq \|f - f_\{(Q)\}\|_1$$

This proves that 1) implies 2) Next we have for $Q' \subset Q$

$$|f(Q') - f(Q)| = |\sum_{Q' \subset \neq Q'' \subset Q} f_{Q,i}h_{Q,i}(Q')|$$

$$\leq \sum_{Q' \subset \neq Q'' \subset Q} |f_{Q,i}| \leq \sum_{l(Q') < 2^k \leq l(Q)} C' 2^\alpha k \leq Cl(Q)^\alpha$$

This proves that $f(Q_{k,x})$ is a Cauchy sequence and hence every point is a Lebesgue point. Let x, y be two points and choose a grid such that there is a cube Q containing x and y with $l(Q) \leq 2|x - y|$. Then applying the above argument and takign a limit we obtain

$$|f(x) - f(y)| \leq |f(x) - f(Q)| + |f(Q) - f(y)| \leq Cl(Q)^\alpha \leq 2C|x - y|^\alpha$$

This shows that 2) implies 3). Since by 3) we have

$$|f(x) - f(y)| \leq Cl(Q)^\alpha$$

we also have

$$|f(x) - f(Q)| \leq Cl(Q)^\alpha$$

sicne $f(Q)$ is in teh convex hull of all $f(y)$ with $y \in Q$. Hence 3) implies 4). That 4) implies 1) follows by the triangle inequality. $\square$

Definition 13 *A function $f \in L^1(\mathbf{R}^n)$ is called an α- Hölder atom on the cube Q if*

$$f_{\{Q\}} = 0$$

and for every dyadic grid and every Q' in the dyadic grid we have

$$\|f - f_{\{Q'\}}\|_1 \leq |Q|(\frac{l(Q')}{l(Q)})^\alpha$$

Note the following observations (exercise): Hölder atoms are Hölder functions, they are supported on Q, the satisfy $\|f\|_1 \leq C_\alpha$ for some universal constant C_α depending only on α and they have vanishing integral $\int_{\mathbf{R}^n} f(x)\, dx = 0$.

In any grid containing a cube Q'' with $Q \subset Q''$ the function has an absolutely convergent expansion

$$f = \sum_{Q' \subset Q''} \sum_i f_{Q',i} h_{Q,i}$$

5 The Poisson equation for Hölder atoms

Let $\Delta u = f$ and assume f is an α-Hölder atom. the purpose of this section is to show that the entries of the Hessian of u are again (locally) Hölder continuous. As will turn out there is a particular $L^1(\mathbf{R})$ solution u (which by the maximum principle is unique), and we will establish that this solution is α-Hölder in the definition of the previous section.

Since f can be written as convergent series of linear combinations of Haar functions, the crucial step is to consider f a Haar function. Sicne each Haar function itself is a linear combination of characteristic functions of cubes, our previous calculations for cubes will get us a long ways.

We will analyse $D_i D_j u$ by calculating its Haar coefficients.

Lemma 18 *Let Q be some cube and $g : D_i D_j \Gamma * h_{Q,i}$. Then*

1. *For $x \notin 10Q$, we have*

$$g(x) \leq C(\frac{\operatorname{dist}(Q,x)}{l(Q)})^{-n-1}$$

2. $$\|g\|_1 \leq C|Q|$$

3. *If Q' is a cube, not necessarily from the same dyadic grid as Q, such that $l(Q) < l(Q')$ and*

$$\operatorname{dist}(Q, \partial Q'') \geq l(Q')/10$$

for every child Q'' of Q', then

$$g_{Q',i'} \leq C\frac{l(Q)}{l(Q')}(\frac{\operatorname{dist}(Q, \cup_{Q''}\partial Q'')}{l(Q')})^{-n-1}$$

where the union runs over the children Q'' of Q'.

Proof: Since each Haar function is a linear combination of functions of the form

$$1_{Q'} - 1_{Q''}$$

with two adjacent cchildren Q' and Q'' of Q, it suffies to prove the first two estimates with the function g replaced by $1_{Q'} - 1_{Q''}$. Write $Q'' = T_y(Q')$ where y is a multiple of length $l(Q)$ of a standard basis vector of $\mathbf{R}^n$, say the k-th basis vector. Then we have for $x \notin 10Q$

$$D_iD_j\Gamma * (1_{Q'} - 1_{Q''})(x)$$

$$= (D_iD_j\Gamma * 1_{Q'})(x) - T_y(D_iD_j\Gamma * 1_{Q'})(x)$$

$$= |y| \int_0^1 D_k(D_iD_j\Gamma * 1_{Q'})(x + ty)\, dt$$

By the estimates on the thrid partials of $\Gamma * 1_{Q'}$ we have (noting the distance from x to Q and from $+ty$ to Q are the same within a factor of 2)

$$|D_iD_j\Gamma * (1_{Q'} - 1_{Q''})(x)| \leq |y| \int_0^1 Cl(Q)^{-1}(\frac{\mathrm{dist}(Q,x)}{l(Q)})^{-n-1}\, dt$$

$$= C(\frac{\mathrm{dist}(Q,x)}{l(Q)})^{-n-1}$$

This proves the first estimate.

Since the pointwise estimate outside $10Q$ that is established in the frist inequality integrates to $C|Q|$, it suffices to estimate the L^1 norm of g restricted to $10Q$. But the L^1 norm on $10Q$ of $D_iD_j\Gamma * 1_{Q'}$ for any child Q' of Q is bounded by $C|Q|$ given the estimates estbalished before, hence the second estimate follows.

To see the third estimate, we note that

$$(D_iD_j\Gamma * h_{Q,i}, h_{Q',i'}) = -(h_{Q,i}, D_iD_j\Gamma * h_{Q',i'})$$

by antisymmetry of the kernel (and some technical arguments since the kernel is not integrable), it suffices to estimate

$$(h_{Q,i}, D_iD_j\Gamma * 1_{Q''})$$

for every child Q'' of Q'. Writing again $h_{Q,i}$ as linear combination of differences of characetristic functions of neighboring cubes as before, and unsing esrtimates for the derivative $D_kD_iD_j\Gamma * 1_{Q'}$ establishes the third estimate.

□

Now let f be an α-Hölder atom on a cube Q_0. We plan to estimate the Haar coefficient

$$(D_iD_j\Gamma * f)_{Q,i}$$

Choose a grid such that both Q and Q_0 are contaiend in the middle 90 percent of cubes of six times the sidelength, this is possible (exercise). let Q_1 be that cube containing Q_0.

Write f as an expansion

$$\sum_{Q' \subset Q_1} f_{Q',i} h_{Q',i}$$

We split this into three terms. The first term is

$$\sum_{Q' \subset Q_1 : Q' \subset 100Q} f_{Q',i} h_{Q',i}$$

We then have

$$|(D_i D_j \Gamma * \sum_{Q' \subset Q_1 : Q' \subset 100Q} f_{Q',i} h_{Q',i}, h_{Q,i})|$$

$$\leq \| D_i D_j \Gamma * \sum_{Q' \subset Q_1 : Q' \subset 100Q} f_{Q',i} h_{Q',i} \|_1$$

$$\leq \sum_{Q' \subset Q_1 : Q' \subset 100Q} |f_{Q',i}| \| D_i D_j \Gamma * h_{Q',i} \|_1$$

$$\leq \sum_{Q' \subset Q_1 : Q' \subset 100Q} |f_{Q',i}| |Q'|$$

$$\leq \sum_{k<10} \sum_{Q' \subset Q_1 : Q' \subset 100Q : 2^k l(Q') = l(Q)} C 2^{\alpha k} (\frac{l(Q)}{l(Q_0)})^\alpha |Q'|$$

$$\leq C \sum_{k<10} 2^{\alpha k} (\frac{l(Q)}{l(Q_0)})^\alpha |Q|$$

$$\leq C (\frac{l(Q)}{l(Q_0)})^\alpha |Q|$$

This is the desired estimate for the first term.

The next term is the sum over the set

$$Q' \subset Q_1 : l(Q') \leq l(Q), Q' \not\subset 100Q$$

We split this set further by a parameter $m \geq 0$ into sets

$$\mathbf{Q}_m = Q' \subset Q_1 : l(Q') \leq l(Q), Q' \subset 2^{m+1}Q, Q' \not\subset 2^m 100Q$$

It suffices to consider these sets separately, proving an estimate that decays geometrically in m. We have

$$|(D_i D_j \Gamma * \sum_{Q' \in \mathbf{Q}_m} f_{Q',i} h_{Q',i}, h_{Q,i})|$$

$$= |(\sum_{Q' \in \mathbf{Q}_m} f_{Q',i} h_{Q',i}, D_i D_j \Gamma * h_{Q,i})|$$

$$\leq \| \sum_{Q' \in \mathbf{Q}_m} f_{Q',i} h_{Q',i} \|_1 2^{-m(n+1)}$$

where we use the pointwise bound of

$$D_i D_j \Gamma * h_{Q,i}$$

on Q'. Now the L^1 norm is estimate by the estimates of the Hölder atom,

$$\leq \sum_{Q' \in \mathbf{Q}_m} |f_{Q',i}| |Q'| 2^{-m(n+1)}$$

$$\leq C \sum_{k<10} 2^{k\alpha} (\frac{l(Q)}{l(Q_0)})^\alpha \sum_{Q' \in \mathbf{Q}_m, 2^k l(Q') = l(Q)} |Q'| 2^{-m(n+1)}$$

$$\leq C \sum_{k<10} 2^{k\alpha} (\frac{l(Q)}{l(Q_0)})^\alpha 2^{nm} |Q| 2^{-m(n+1)}$$

$$\leq C 2^{-m} (\frac{l(Q)}{l(Q_0)})^\alpha |Q|$$

Finally we consider the sum

$$(\mathbf{D}_i D_j \Gamma * \sum_{Q' : l(Q') > 10 l(Q)} f_{Q',i'} h_{Q',i'}, h_{Q,i})$$

$$\leq \sum_{Q' : l(Q') > 10 l(Q)} C|Q| (\frac{l(Q')}{l(Q_0)})^\alpha \frac{l(Q)}{l(Q')} (\frac{\mathrm{dist}(Q, \cup_{Q''} \partial Q'')}{l(Q')})^{-n-1}$$

$$\leq \sum_{k>10} \sum_{Q' : l(Q') = 2^k l(Q)} C|Q| 2^{-k} 2^{\alpha k} (\frac{l(Q)}{l(Q_0)})^\alpha (1 + \frac{\mathrm{dist}(Q, Q')}{l(Q')})^{-n-1}$$

Here we have used that Q is not close to the boundary of any of the children of Q'. Using $\alpha < 1$ the last display is then bounded by

$$\leq \sum_{k>10} C|Q| 2^{-k+\alpha k} (\frac{l(Q)}{l(Q_0)})^\alpha \leq C|Q| (\frac{l(Q)}{l(Q_0)})^\alpha$$

The condition $\alpha < 1$ is used here but not crucial, a modification of the argument, using that f is an atom, will give the result without this assumption.