- Integral diffusions
- Nonlinear (far from linear) equations
- Divergence – Nondivergence

Luis Caffarelli
The University of Texas at Austin

Review of regularity theory for second order equations

A diffusion process is one where the quantity under consideration, u, a density, a temperature, tries to revert to its (infinitesimal) surrounding average:

Example:

$$u_t = \Delta u \quad (\text{or } a_{ij}D_{ij}u, \text{ or } \operatorname{div}(a_{ij}D_{ij})$$

Indeed, we can see the Laplacian as an infinitesimal average:

$$\begin{aligned}\Delta u(x) &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^2} ⨍_{B_\varepsilon(x)} [u(y) - u(x)]\, dy \\ &= \frac{1}{\varepsilon^2}\left[⨍_{B_\varepsilon(x)} u - u(x) \right]\end{aligned}$$

or an elliptic equation

$$a_{ij}D_{ij}u(x) \text{ “=” } \int\limits_{\substack{\text{unit}\\ \text{sphere}}} u_{\sigma\sigma}(x)\omega(\sigma)\, d\sigma$$

as a weighted average of the second order directional derivatives

or a divergence equation as the flux across an infinitesimal sphere:

$$\operatorname{div} a_{ij}\nabla u = \lim_{\varepsilon \longrightarrow 0} \frac{1}{\varepsilon} \fint_{\partial B_\varepsilon} \nabla u \cdot A \cdot \nu \, d\sigma$$

In an “integral” diffusion process, u still tries to revert to its surrounding average, but it is not any more infinitesimal, in some sense, $u(x)$ is able to “see its surroundings”, i.e.:

$$u_t(x,t) = \int_{\mathbb{R}^n} [u(y,t) - u(x,t)] K(y,x)\, dy$$

for K a positive (non-negative) kernel or measure.

(If you ever have any doubt which is the “parabolic” sign, this is a way to remember.)

Minimal requirement is that

$$\int K(y,x)\min(|x-y|^2,1)\,dy < \infty$$

(Levi-Kintchine condition)

This guarantees that for a u that is C^2 near the origin and bounded at infinity the operator makes sense, at least for $K(x,y)$ symmetric in y around x:

$$K(x,\ x+y) = K(x,\ x-y)$$

We also note that if you expect the equation to force some regularity on u, the kernel

$$K(x,y)$$

must be rather singular in $x-y$. Indeed, that u convolved with a smooth function is nice does not say much about u.

The basic model family of examples are the kernels

$$K(x,y) = (1-s)|x-y|^{-(n+2s)}$$

for $0 < s < 1$.

The integral operator becomes then a convolution, with multiplyer

$$|\xi|^{2s}$$

and is referred to as the s (“fractional”) Laplacian.

Note that in this case, for the integral to converge, u has to provide some cancellation.

Before going to integral equations let us review second order equations:

- An important part of the classical theory concerns equations that are "small perturbations" of the Laplace (or constant coefficient) equations

Typical examples are the Shauder and Calderon-Zygmund theories, that assert that solutions to equations $a_{ij}(x)D_{ij}u$ with "smooth" coefficients behave "as well" as solutions to the Laplace equations. The two families of problems that are "really non linear" (cannot be treated as a perturbation of the linear theory) are respectively in the divergence and nondivergence case:

- Energy minimizers (calculus of variations)

  “u_0 minimizes (locally)

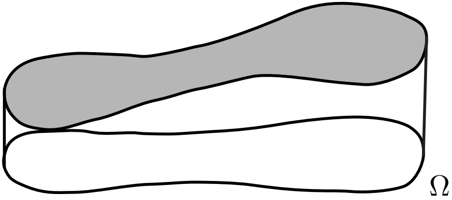
$$E(u) = \int F(\nabla u)\, dx$$

  with F smooth, strictly convex

Typical example:

Area minimization

$$A\,(\text{graph } u) \;=\; \int_\Omega \sqrt{1+|\nabla u|^2}\,dx$$



$u = f$
on $\partial\Omega$

Is strictly convex if $|\nabla u|$ is bounded.

- **Optimal control:**

Given a family of strictly positive (symmetric) matrices, for instance

- $\mathcal{A} = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 5 \end{bmatrix} \right\}$ or
- $\mathcal{A} = \{$“all matrices A such that $\frac{1}{2}I \leq A \leq 3I\}$ find a function u such that

$$F(D^2u) = \sup_{A\in\mathcal{A}} \sum a_{ij}D_{ij}u = 0$$

(also from geometry: Monge Ampere, symmetric functions of the Hessian, etc.)

Example b): "All matrices between $\frac{1}{2}I$ and $3I$" corresponds to the extremal Pucci operator. The solution u can be thought as an "upper envelop" of all possible solutions, v, to equations

$$a_{ij}(x)D_{ij}v = 0 \text{ for } \frac{1}{2}I \leq A \leq 3I$$

with prescribed Dirichlet data.

Indeed u is a supersolution for all those operators and at the same time is admissible since at every point it is a solution of one of the operators.

> "Find $u \to$ compute D^2u and its eigenvalues $\lambda_1 \ldots \lambda_n$, then we must have
>
> $$\frac{1}{2}\sum_{\lambda_j<0}\lambda_j + 3\sum_{\lambda_j>0}\lambda_j = 0$$

Regularity theory:

- **Divergence case:** (De Giorgi-Nash-Moser)

A minimizer u_0 to the variational integral

$$E(u_0) = \int F(\nabla u)$$

satisfies (the Euler Lagrange equation)

$$D_{x_i} F_{p_i}(\nabla u_0) = 0 \quad \text{(or: } F_{p_i p_j}(\nabla u_0) D_{ij} u_0 = 0$$

If we were to know that ∇u_0 is a Hölder function of x, Shauder estimates would kick in, and we would get:

“u_0 is as regular as F”

In principle, we only know that $\nabla u_0 \in L^2$, but, in turn the derivatives of u, satisfy: ($D_e u_0$ the directional derivative of u_0 in the direction e)

$$D_{x_i} F_{p_i p_j}(\nabla u_0) D_{x_j}(D_e u_0) = 0$$

DeGiorgi:

“We just forget that $F_{p_ip_j}(\nabla u_0)$ depends on u_0:

$$D_{x_i}\underbrace{F_{p_ip_j}}_{a_{ij}(x)}(\nabla u_0)D_j\underbrace{(D_e u)}_{w}=0\text{ ”}$$

We have $D_{x_i}a_{ij}(x)D_jw=0$) (where all we know of a_{ij} is that it is elliptic but we cannot assume any *regularity*)

“Nevertheless w is C^α”.

Then Shauder estimates kick in, and we have the desired regularity.

Given an equation invariant under $\mathbb{R}^{n+1}$ translations: Laplacian, minimal surface, fully nonlinear with a comparison principle, we should expect some sort of elliptic equation for incremental quotients and first derivatives

i.e.: If u, v are solutions in Ω and $u > v$ on $\partial\Omega$, the graph of u “cannot corss” the graph of v.

- Fully nonlinear: $F(D^2u) = 0$

We start as DeGiorgi: “Since the equation is translation invariant, and satisfy a “comparison principle” (two solutions, one above the other, cannot touch), first derivatives, $D_e u$, must satisfy an equation.”

(F is now a function of symmetric matrices and F_{ij} denotes the derivative of F with respect to the position i, j., not a second derivative.)

$$\underbrace{F_{ij}(D^2u)}_{a_{ij}(x)} D_{ij}(D_e u) = 0$$

(again, an elliptic equation, now in non divergence form, but no a priori regularity on the “coefficients”)

Krylov-Safanov: $\quad w = D_e u \quad$ is C^α (in fact a Harnack inequality holds)

Not enough yet for Schauder to kick in, since the coefficients $F_{ij}(D^2u)$ depend on D^2, not any more on ∇u.

Try with *second derivatives*: $D_{ee}u = v$

$$\underbrace{F_{ij}(D^2u)}_{a_{ij}} D_{ij} \underbrace{D_{ee}u}_{D_{ij}v} + \underbrace{D_{ije}u F_{ij,k\ell} D_{k\ell e}u}_{v^t \quad \mathcal{M} \quad v\geq 0} = 0$$

That is: Since F is convex, (any linear operator, $a_{ij}D_{ij}u$ is a linear function of D^2u) all pure second derivatives are supersolutions of the linearized operator, (i.e., like superharmonic functions, they all try to stay pointwise “above” their average) but, on the other hand D^2u lives in the Lipschitz surface $F(D^2u) = 0$.

Evans-Krylov: D^2u is C^α, Schauder estimates finally kick in for the first derivative equations, and u is as regular as F.

Non linear integral equations

Main issues and results

Integral operators: Two types: variational from continuum mechanics, non variational, from probability, optimal control.

Calculus of variations (“divergence” type integral equations)

We minimize an energy of the form:

$$E(u) = \iint \phi(u(x) - u(y))K(x,y)\,dx\,dy$$

Example: Fractional Laplacian

$$E(u): \quad (1-s)\iint (u(x)-u(y))^2 |x-y|^{-(n+2s)} \qquad (0 < s < 1)$$

Associated parabolic equation (Euler Lagrange)

$$u_t(x,t) = \int K(x,y)[u(y) - u(x)]\,dy$$

(Fractional Laplacian diffusion: $u_t = \Delta^s u$ for $K = |x-y|^{-n+2s}$)

Chen-Fife (Phase-transition-in solids)

Giacomish-Lebowitz (Phase Segregation)

Presutti (Phase transition)

G. Gilboa and S. Osher, Non local operators with applications to Image Processing

Some interesting “degenerate” variational equations

$$\iint [u(x) - u(y)]^p |x - y|^{-(n+ps)}\, dx\, dy$$

(“p” fractional Laplacian?)

$$\iint \chi_\Omega(x)\chi_{C\Omega}(y)|x - y|^{-(n+\sigma)}\, dx\, dy$$

($\sigma < 1$, fractional minimal surfaces, C-Roquejoffre-Savin)

$$\iint |u(x) - u(y)|\, |x - y|^{-(n+s)}\, dx\, dy$$

(“1” Laplacian, provides a foliation by “fractional” minimal surfaces)

What is the equivalent theory to De Giorgi-Nash-Moser?

We proceed as before:

A minimizer of $\iint [\phi(u(y) - u(x))] K(x-y)\, dx\, dy$

satisfies the Euler Lagrange equation:

$$\int \phi'(u(y) - u(x)) K(x-y)\, dy = \cdots$$

A first derivative $w(x) = D_e u$ then satisfies

$$\int \underbrace{K(x-y)}\, (w(y) - w(x))\, \underbrace{(\phi''(u(y) - u(x))}\, dy = \cdots$$

symmetric, bounded measurable kernel

DeGiorgi-Nash-Moser (parabolic):

Study regularity of “Bounded measurable coefficients equations” (satisfied by the incremental quotients of the nonlinear “divergence” equations):

$$u_t(x,t) = \int [u(y,t) - u(x,t)] K(x,y) dy$$

in the weak sense: given a test function φ

$$\int dt \int dx \, \varphi(x,t) u_t(x,t)$$
$$= \iiint [\varphi(x,t) - \varphi(y,t)] K(x,y) [u(x,t) - u(y,t)] \, dx \, dy \, dt$$

$K(x,y)$ symmetric.

Optimal control ("non-divergence")

Given a family of kernels $\mathbb{K} = \{K_\alpha(y)\}$ find u that solves, for x in a domain Ω.

$$0 = F(u) = \sup \text{“}L_\alpha u\text{”} = \sup_{K_\alpha \in \mathbb{K}} \int [u(x+y) + u(x-y) - 2u(x)] K_\alpha(y)\, dy$$

(For both "divergence" or "non-divergence" case the Dirichlet problem requires that, given a domain Ω, we prescribe data in *all* of $\mathbb{R}^n \setminus \Omega$, for the integrals to make sense. This corresponds to the fact that integral diffusions are generated by jump processes that may jump clear from $\partial\Omega$.)

Then the “bounded measurable coefficients” theory we need for the corresponding regularity for the nonlinear problems is the following:

Regularity theory for non-divergence “bounded measurable coefficients” (corresponding to Krylov-Safanov)

For each x, $\exists\, K_x(y)$ such that

$$L_x(u) = \int [u(x+y) + u(x-y) - 2u(x)] K_x(y) = 0$$

For our theory, we must “fix” the order, s, of the equation, i.e., the kernels must behave like a “fixed fractional Laplacian”, both in the “divergence” and “non divergence” case.

$$\lambda|x-y|^{-(n+2s)} \leq K(x,y), K_x(y) \leq \Lambda|x-y|^{-(n+2s)}$$

Divergence, bounded measurable (with C.H. Chan and A. Vasseur)

(see also work of Moritz Kassmann)

Theorem

Let $u(x,t)$ be a solution of the equation ("bounded measurable kernel", DeGiorgi type)

$$u_t(x,t) = \int [u(y,t) - u(x,t)]K(x,y)\,dy$$

Data in $L^2 \Rightarrow$ instantaneous $L^\infty \Rightarrow$ instantaneous C^α.

Corollary

Let v be a solution of $v_t = \int \phi'(v(x) - v(y))K(x-y)$ with ϕ strictly convex, smooth and symmetric. Then v is $C^{1,\alpha}$ in space.

Non-Divergence, bounded measurable (with L. Silvestre)

Theorem (Krylov-Safanov type)

Let $u(x)$ be a solution of the "bounded measurable coefficients" equation:

$$L_x(u) = \int [u(x+y) + u(x-y) - 2u(x)] K_x(y) = 0$$

in Ω. Then u satisfies a Harnack inequality and it is C^α

Theorem (Evans-Krylov type)

v satisfies an optimal control equation

$$F(v) = \sup_{K_\alpha \in \mathbb{K}} \int [v(x+y) + v(x-y) - 2v(x)] K_\alpha(y) = 0$$

then $v \in C^{2s+\alpha}$ (s the order of the equation) (Evans-Krylov type theorem).

As an application of the above techniques, we also study:

Integral Porous Media (with Soria and Vazquez)

- Conservation of mass:

$$\rho_t + \operatorname{div} \rho\vec{v} = 0 \qquad (\rho \text{ a density})$$

- Constitutive relation:

$$\vec{v} = -\nabla p = \qquad (p \text{ a ``pressure''})$$

- nonlocal pressure: (instead of $p \sim f(\rho)$ as in the porous media)

$$p = \int V(x-y)\rho(y)$$

For instance

$$V(x) = |x-y|^{-n+\varepsilon} \qquad (-\Delta^{-\varepsilon}\rho)$$

$$\rho_t = \operatorname{div} \rho\, \nabla(\Delta^{-\varepsilon} \rho) = \operatorname{div} \rho\, \nabla p$$

We then have:

- Initial data in L^1 (finite mass) becomes instantaneously bounded.
- Initial compact support (for ρ) $\Rightarrow$ compact support of ρ for all times.
- Decay in time to self similar solution.

[Arises in work of Giacomin (material sciences), Lions-Mas-Gallic and Ambrosio-Serfaty (semiconductivity).]

Theorems and proofs: In these lectures we will discuss the variational regularity theory.

Nonlinear integral variational problem

Let $\phi(t)$

- convex $\quad (0 < \lambda < \phi'' < \Lambda)$
- symmetric
- $\phi(0) = 0$

Variational integral

$$\iint \phi(u(x) - u(y))\, K(x-y)\, dx\, dy$$

(Note that $K(x-y)$ makes it translation invariant.)

Euler-Lagrange equation:

$$\int \phi'(u(y)-u(x))\,K(x-y)\,dy = \begin{cases} 0 \\ f(x,u) \\ u_t(x,\dots) \end{cases}$$

we write $(y = x+z)$:

$$\int \phi'(u(z+x)-u(x))\,K(z)\,dz = \cdots$$

If we use a derivative in the direction **e**, we get

$$\int \phi''[u(z+x)-u(x)][\underbrace{D_e u(z+x)}_{w(z+x)} - \underbrace{D_e u(x)}_{w(x)}]\,K(z)\,dz = \cdots$$

...and we change back to *y* and rearrange to

$$\int [w(y) - w(x)] \underbrace{[\, \underbrace{\phi''(u(y) - u(x))}_{\substack{\text{symmetric} \\ \text{function of} \\ x \text{ and } y}} \,]\, K(y - x)}_{\substack{\text{symmetric kernel} \\ \text{of } x \text{ and } y}} \, dy$$

(The actual computation is done for incremental quotients, as in the second order case.)

Theorem (Caffarelli-Chan-Vasseur)

Let $w(x,t)$ be a weak solution of

$$w_t(x,t) = \int [w(y,t) - w(x,t)]\, K(x,y)\, dy$$

with initial data in L^2.

Then w becomes instantaneously bounded and Hölder continuous.

The kernel K has to be symmetric and comparable to a fractional Laplacian:

$$(2-s)\lambda|x-y|^{-(n+s)} \le K(x,y) \le (2-s)\Lambda|x-y|^{-(n+s)}$$

The proof is based in the “non homogeneous” interplay of:

- **The Sobolev inequality**

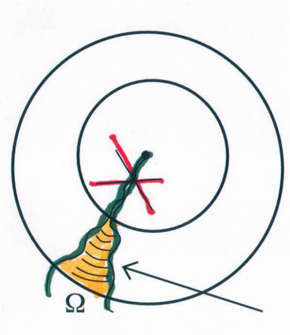
  (Derivatives of a function control the “size” of the function itself in a “locally higher” space).

- **Energy inequality**

  (Because of the equation, the function controls locally the (fractional) derivatives.)

Preliminary: A baby example of the DeGiorgi argument:

Let $\partial\Omega$ be a (generalized) minimal surface, and suppose that



"Volume $(\Omega \cap B_1)$ is very small",
then
"Ω cannot intersect $B_{1/2}$".

$\partial\Omega$, a set of minimal perimeter

Proof:

We cut Ω by a decreasing sequence of balls B^k of radius $r_k = \frac{1}{2} + 2^{-k}$ inductivity estimate the volume of $\Omega_k = \Omega \cap B^k$, and by the time we get to $B_{1/2}$ nothing will be left.

This is due to the non-linear interplay between *area* (“the size of the derivatives”) and *volume* (“the size of the function”):

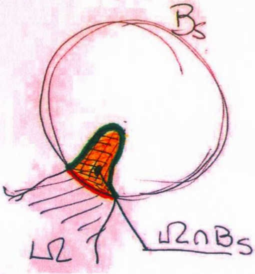
Area controls volume: (isoperimetric inequality)

$$\mathrm{Vol}(K) \le C[A(\partial K)]^{n/n-1}$$

In particular, in our case, if

$$K = \Omega \cap B_s$$

If we cut Ω with any ball, B_s,



$\text{Volume}\ (\Omega\cap B_s) \le C\ \text{Area}(\partial(B_s\cap\Omega))^{n/n-1}$

But, if $\partial\Omega$ is area minimizing

$\text{Area}((\partial\Omega)\cap B_s) \le \text{Area}(\partial B_s\cap\Omega)$

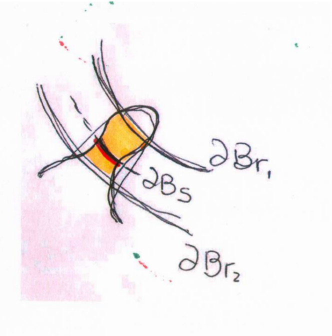
so $\text{Vol}(\Omega\cap B_s) \le C\ \text{Area}(\partial B_s\cap\Omega)^{n/n-1}$

$\text{Vol}\,(\Omega\cap B_s) \le$

$\le \text{Area}\,(\partial B_s \cap \Omega)^{n/n-1}$

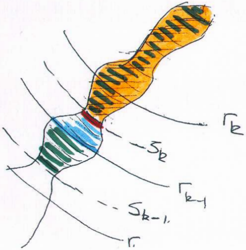
On the other hand,

Volume controls area: If $\frac{1}{2} \leq r_1 \leq r_2$, we cut Ω with $B_{r_2} \setminus B_{r_1}$, and choose the "slice ∂B_s" with smaller area,



$$
\begin{aligned}
&\text{Area}\,(\partial B_s \cap \Omega) \leq \\
&\leq (r_2 - r_1)^{-1} \int_{r_1}^{r_2} \text{Area}\,(\partial B_t \cap \Omega)\, dt \\
&\leq C \frac{\text{Vol}\,(\Omega \cap B_{r_2} \setminus B_{r_1})}{(r_2 - r_1)}
\end{aligned}
$$

We now choose $r_k = \frac{1}{2} + 2^{-k}$, and s_k between r_{2k-1} and r_{2k}, the "slice" with least area, and we write a recurrence relation for $A_k = \mathrm{Vol}(\Omega \cap B_{s_k})$



$$\underline{A_k = \mathrm{Vol}(\Omega \cap B_{s_k})} \leq C\, \underline{\mathrm{Area}(\partial B_{s_k} \cap \Omega)}^{n/n-1}$$

$$\leq C[2^k\, \underline{\mathrm{Vol}(\Omega \cap B_{r_k})]}^{n/n-1}$$

$$\leq C[2^k\, \underline{\mathrm{Vol}(\Omega \cap B_{s_{k-1}})]}^{n-1}$$

$$= C\underline{(2^k A_{k-1})^{n/n-1}}\,, \quad \text{i.e.}$$

$$A_k \leq C^k (A_{k-1})^{1+\varepsilon}\,.$$

But if A_0 is small enough, this recursive relation implies that A_k goes to zero, as k goes to infinity.

Back to our theorem.

Theorem (Caffarelli-Chan-Vasseur)

Let $w(x,t)$ be a weak solution of

$$w_t(x,t) = \int [w(y,t) - w(x,t)]\, K(x,y)\, dy$$

with initial data in L^2.

Then w becomes instantaneously bounded and Hölder continuous.

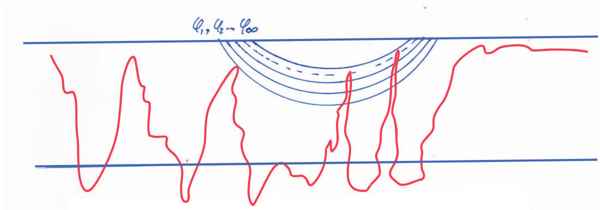
We will discuss the proof of the second part:

Bounded solutions are Hölder continuous

Main steps in the proof:

Given our bounded solution w, we are going to use test functions $w_\varphi = (w - \varphi)^+$ of compact support, (for a sequence of appropriate $\varphi_k \geq \varphi_{k-1}$.)

(This corresponds to cutting the minimal surface with spheres in the baby example.)



Let us see what type of “energy formula” we have....

We multiply the equation by w_φ and integrate in space and time:

$$\int_{T_1}^{T^2} \int w_\varphi(x,t) w_t(x,t)\, dx\, dt =$$

$$= \int_{T_1}^{T_2} dt \underbrace{\iint w_\varphi(x,t)[w(z,t) - w(x,t)]K(x,z)\, dx\, dz}_{\substack{\text{This expression can be symmetrized} \\ \text{(exchange } z \text{ and } x \text{ and substract) to}}}$$

$$- \iint \underbrace{(w_\varphi(x,t) - w_\varphi(z,t))K(x,z)(w(x,t) - w(z,t))}_{B(w_\varphi, w)}\, dx\, dz$$

The “energy formula” then reads

$$\int w_\varphi^2(x,\cdot)dx \,\Big|_{T_1}^{T_2} + \int dt\, B(w_\varphi, w) = 0$$

or (good terms on the left, bad on the right)

$$\int w_\varphi^2(x,T_2)\,dx + \int_{T_1}^{T_2} dt\, B(w_\varphi, w_\varphi) =$$

$$\int w_\varphi^2(x,T_1)dx + \int_{T_1}^{T_2} B(w_\varphi, w_\varphi - w)$$

We first discuss the equivalent of “area controls volume in a non homogeneous way”.

We prove “energy” controls an L^p norm, $p > 2$.

The bilinear term on the left, $B(w_\varphi, w_\varphi)$ is equivalent, by hypothesis to an H_s norm of w_φ, and from Sobolev inequality, controls, for each time, an L^p norm of w_φ, for some $p > 2$, depending on s and dimension.

In other words, the “energy” on the left controls

$$E(w_\varphi) \geq \sup_{T_1<t<T_2} \|w_\varphi\|^2_{L^2(\mathbb{R}^n)} + \int_{T_1}^{T_2} \|w_\varphi\|^2_{L^p(\mathbb{R}^n)} dt$$

$$\geq \left(\int_{T_1}^{T_2} \int_{\mathbb{R}^n} (w_\varphi)^q \right)^{2/q}$$

for some $2 < q < p$ from a standard interpolation.

In turn, the L^2 norm of w_φ in space and time (this is going to be our volume, A_k in the iteration) is controlled by

$$\int_{T_1}^{T_2}\int_{\mathbb{R}^n}(w_\varphi)^2 = \int_{T_1}^{T_2}\int_{\mathbb{R}^n}(w_\varphi)^2\chi_{w_\varphi>0}$$

$$\leq \text{(Hölder)}\left(\int_{T_1}^{T_2}\int (w_\varphi)^q\right)^{2/q}\cdot\cdot|\{w_\varphi>0\}|^{(\theta)}$$

(θ the corresponding inverse conjugate exponent to $q/2$)

We summarize:

$$A = \iint (w_\varphi)^2 \leq E\left(w_\varphi\right) \quad \left|\{w_\varphi > 0\}\right|^\varepsilon$$

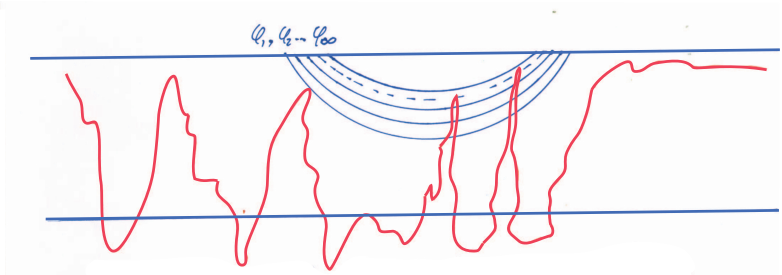
Our final objective is to show that

$$A_{k+1} \leq C^k (A_k)^{1+\theta} \qquad (\theta)$$

The sequence of cut offs and the *iterative relation*: (first attempt)

We now assume that w is defined and $w \leq 0$ for $-2 < t < 0$ and all of $\mathbb{R}^n$, and $w \leq -2$ "most of the time" for (x, t) in $\Gamma_2 = B_2 \times [-2, 0]$.
(a cylinder in space-time)

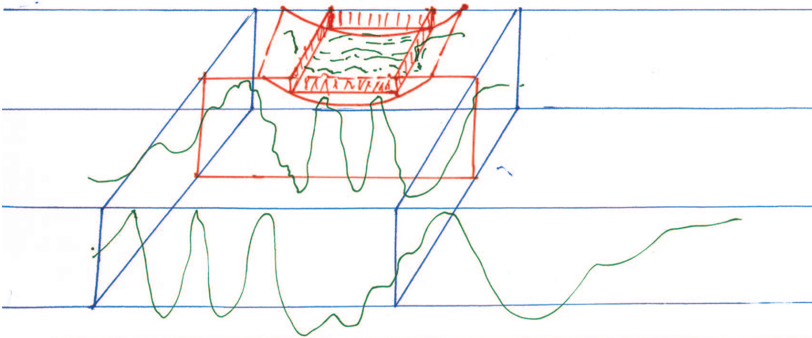
As in the minimal surface case, the idea is to make a sequence of cut offs, φ_k, with a dyadic gap $\varphi_k - \varphi_{k-1} \sim 2^{-k}$



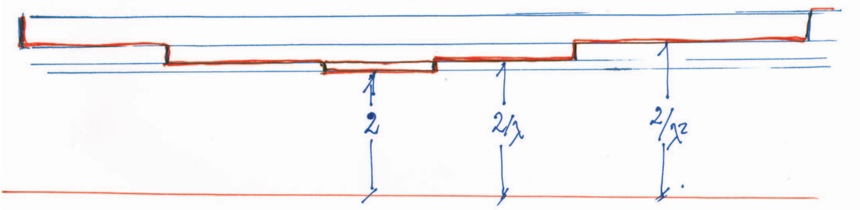
and play the Sobolev embedding (“area controls volume”) against the energy inequality (“volume controls area”) so that, by the time we “cut with φ_∞”, nothing is left.

If we achieve that, we attain a decay of the maximum of *w*, in a smaller domain, say

$$\Gamma_1 = B_1(0) \times [-1, 0]$$



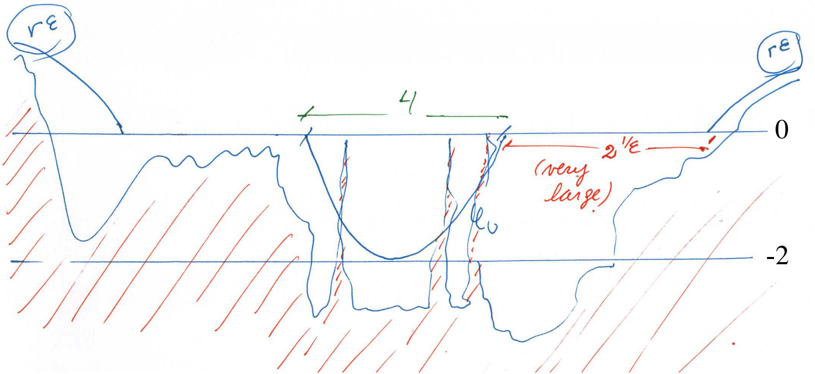
But if we manage to do this, and attempt to repeat the gain in oscillation by rescaling Γ_1 into Γ_2 and expand the graph of w to go back to the hypothesis "$w \leq 0$", we realize that our operator being global, we need to make a correction in our "normalized lemma" since as we renormalize w will not remain, globally, below 2, and will start to grow at infinity:



Let us now restate the “good” basic lemma:

- In $\mathbb{R}^n \times [-4, 0]$, we assume that the subsolution $w \leq \psi = \max(0, |x|^\varepsilon - 2)$, ε small, to be chosen
- In $\Gamma_4 = B_4 \times [-4, 0]$, “$w \leq \varphi_0$ most of the time” (i.e., $|\{w > \varphi_0\}| \leq \delta$) with φ_0: $\varphi_0 = \frac{1}{2}|x|^2 - 2$.

Note that a choice of a small growth rate, ε, at infinity may seem counter intuitive. But it is a necessary choice to ensure the convergence of the integrals, and an acceptable one that is attained by reducing the gain (the Hölder exponent) in the inductive process.



Define $\varphi_k = \min\left(\frac{1}{8}|x|^2 - (1+2^k), \psi\right)$, and $w_{\varphi_k} = w_k = (w - \varphi_k)^+$

Then

$$\text{“}w_\infty\text{”} = \text{“}(w-\varphi_\infty)^+\text{”} \equiv 0 \ \text{ in } \ \mathbb{R}^n \times [-1,0]\ ,$$

i.e.

$$w \le \min\left(\frac{1}{8}|x|^2 + 1, \psi\right) \ \text{ a.e.}$$

Proof: We will obtain an iterative relation for A_k:

$$A_k = \int_{T_k}^0 dt \int_{\mathbb{R}^n} [w_k(x,t)]^2\, dx$$

(with $T_k = -(1+2^{-k})$, so that $A_k \to 0$

More precisely, we will show that

$$A_k \le C^k (A_{k-1})^{\,1+\theta}$$

It is easy to prove that, due to the $1+\theta$ non-linearity, if A_0 is small enough, then $A_k \to 0$.

We already showed that

$$A_k = \int_{T_k}^{0} dt \int_{\mathbb{R}^n} [w_k(x,t)]^2 \, dx \le$$

$$\left[\iint dt\, dx \Big[w_k(x,t] \Big]^q \right]^{2/q} |\{w_k > 0\}|^\theta$$

And, from the Sobolev inequality,

$$A_k \leq E_{T_k,0}(w_k)|\{w_k > 0\}|^{\theta}$$

We now prove the "volume controls area" part, i.e. that both $E_{T_k,0}$ and $|\{w_k > 0\}|$ are bounded by $C^k A_{k-1}$.

a) Bound for $|\{w_k > 0\}|$: From the definition of φ_k, the fact that $w_k > 0$ implies that $w_{k-1} > 2^{-k}$, so

$$|\{w_k > 0\}| \leq |\{w_{k-1} > 2^{-k}\}| \leq 2^{2k} \iint dt\, dx (w_{k-1})^2$$

(Tchebichef)

b) About $E_{T_k,0}$, if we go back to the "good terms on the left, bad terms on the right" inequality, we have

$$E_{T_k,0} \leq \int w_k^2(x,T_k)dx + (-B(w_k,\varphi_k) - B(w_k,(w-\varphi_k)^-)\ )$$

$$= I + II + III$$

About I, we note that we may replace in our argument T_k by any initial time between T_{k-1} and T_k (that will only increase A_k)

So we may replace I by ("the slice with smallest area")

$$\inf_{T_{k-1}<T<T_k} \int w_k^2(x,T)dx \leq 2^k \int_{T_{k-1}}^{T_k} \int w_k^2(x,t)dx\,dt \leq 2^k A_{k-1}$$

About $B(w_k, \varphi_k)$, we write it as

$$\iint [w_k(x) - w_k(y)] \, [x_{w_k>0}(x) x_{w_k>0}(y)] \, K(x,y) \, [\varphi_k(x) - \varphi_k(y)]$$

By Hölder we get $II \leq$

(Cauchy)

$$\frac{1}{2}E + \frac{1}{2}\iint \underbrace{[\varphi_k(x) - \varphi_k(y)]^2 K(x,y)}_{*} [x_{w_k>0}(x) + x(y)]$$

From the Lipschitz regularity and slow growth of φ_k, $*$ is integrable in one of the variables, say y, for fixed x in the support of w_k.

Therefore, for this term we have the bound

$$\frac{1}{2}\iint \leq |\{w_k > 0\}|$$

and we proceed as in a).

Finally, the term

$$-B(w_\varphi, (w-\varphi)^-)$$

has the right sign, since the two functions have disjoint support, and that leaves us only with mixed terms

$$-\iint \underset{+}{w_\varphi(x)}\, \underset{+}{K(x,y)}\, [\underset{+}{-(w-\varphi)^-(y)}] \le 0$$

Summing up:

Among the terms on the right of the energy, the initial time slice is controlled by choosing the smaller value in T_{k-1}, T_k, the term $B(w_\varphi, \varphi)$ is split in a small multiple of $B(w_\varphi, w_\varphi)$ and $B^*(\varphi, \varphi)$, with B^* controlled by $|\{w_\varphi > 0\}|$, and $B(w_\varphi, (w - \varphi) \wedge 0)$ is a "good term" after all.

This completes the first lemma.

Second Step: From “$w \leq -2$ most of the time” to “$w \leq -2$ in some positive set”

In the first step, we proved that if $w \leq 0$ and in Γ_4, we knew that $|\{w \geq -2\}| \leq \delta$ for some critical value δ, then in the smaller cylinder Γ_1, w became strictly negative. But this is not a realistic assumption: Given a function w, in a cylinder Γ, if we choose a cut off level, say -1, all we can say is that

> “$w \leq -1$ at least half of the time in Γ,
> or $w \geq -1$ at least half of the time in Γ.”

Therefore, the proper result we need, to iterate the oscillation gain is:

Given any positive number μ, if $|\{w \leq -2\}| \geq \mu > 0$ in $\Gamma_4 = B_4 \times [-4, 0]$, then for some (very tiny) value $\lambda^*(\mu)$,

$$|\{w \geq -\lambda^*\} \cap \Gamma_1| \leq \delta$$

(i.e. the measure goes below the critical value δ, for the first part of the proof to apply and, as a consequence, we get the uniform gain

$$w|_{\Gamma_{1/4}} \leq -\frac{\lambda^*}{8}$$

(For us: $\mu = \frac{1}{2} |\Gamma_4|$ would do)

Geometry and parameters:

We will have four parameters in our proof:

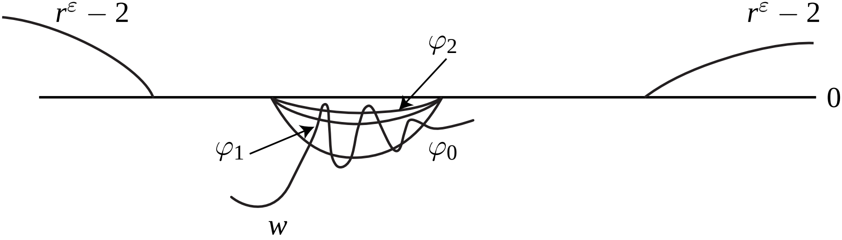
- μ given (small), δ critical mass from previous lemma
- λ will be “much smaller than μ and $\lambda \leq \mu^M, \delta^M$, a large power of μ and δ. The final λ^* will be in turn a power of λ.
- The constant ε, in the tail behavior will be chosen small at the end so that $2^{-s/\varepsilon} \leq (\lambda^*)^2$.

The normalized lemma will involve three consecutive cut offs:

$$\varphi_0 = \min\left(\psi, \frac{1}{8}(|x|^2 - 4)\right)$$

$$\varphi_1 = \min\left(\psi, \frac{\lambda}{8}(|x|^2 - 4)\right)$$

$$\varphi_2 = \min\left(\psi, \frac{\lambda^2}{8}(|x|^2 - 4)\right)$$



The lemma says that in going from the φ_0 cut off to the φ_2 cut off, i.e., from the set

$$\{w > \varphi_0\}$$

to

$$\{w > \varphi_2\}$$

"some mass" is lost, i.e., if $|\{w > \varphi_2\}|$ is not yet subcritical (i.e., $\leq \delta$) Then

$$|\{w < \varphi_2\}| \geq |\{w > \varphi_0\}| + \gamma(\delta, \mu)$$

This implies that after a finite number k of iterations (γ/δ iterations), $|\{w > \varphi_k\}|$ must become subcritical and the previous part applies.

Idea of the proof:

We start again with the energy inequality, but use better the “good” term

$$B(u_\varphi, (u-\varphi)^-) = \iint u_\varphi(x)K(x,y)(-(u-\varphi)^-(y))$$

that we just neglected before.

We have, for the φ_1 (intermediate) cut off:

$$\int (u_{\varphi_1})^2(x,t)\,dx\Big|_{T_1}^{T_2} + \int_{T_1}^{T_2} B(u_{\varphi_1}, u_{\varphi_1}) =$$

$$-\int_{T_1}^{T_2} B(u_\varphi, \varphi_1) - \int_{T_1}^{T_2} B(u_\varphi, (u-\varphi)^-)$$

The term

$$B(u_\varphi, \varphi_1) \leq \frac{1}{2}B(u_\varphi, u_\varphi) + 2\iint [\varphi_1(x) - \varphi_1(y)]K(x,y)[\varphi_1(x) - \varphi_k y][\chi_{B_1}(x)]$$

The first term $\frac{1}{2}B(u_\varphi, u_\varphi)$ is absorbed on the left and the second is smaller than $\lambda^2 + 2^{-\sigma/\varepsilon}$. We will choose ε at the end, when we know the number of steps we need so that all through our process $2^{-\sigma/\varepsilon} \leq \lambda^2$.

This leaves us with the inequality (with $(u-\varphi_1)^- = (u-\varphi_1)\wedge 0$)

$$\int w_\varphi^2 dx\Big|_{T_1}^{T_2} + \frac{1}{2}\int_{T_1}^{T_2} B(w_\varphi, w_\varphi)$$

$$+\Big[\int_{T_1}^{t_2}\int w_{\varphi_1}(x)K(x,y)[-(u-\varphi_1)^-(y)] \le C\,\lambda^2(T_1-T_1)$$

In particular, since the second and third terms are positive, we get that

a) $H(t) = \displaystyle\int w_\varphi^2(x,t)dx$

satisfies

$$H'(t) \le C\,\lambda^2$$

b) $\displaystyle\int_{T_1}^{T_2}\int w_{\varphi_1}(x)K(x,y)[-(u-\varphi_1^-(y))] \le C\,\lambda^2[T_2-T_1]$

An estimate on those time slices where the “good” extra term helps

From our hypothesis, in Γ_1 that

$$|\{w < \varphi_0\}| \geq \mu$$

the set of times Σ for which $|\{w(\ ,T)\varphi_0\}| \geq \mu/8$ has at least measure $\mu/2$.

We estimate now that except for a few of those time slices, $\int w_{\varphi_1}^2$ is very tiny:

If $\inf_{|x-y|\leq 1} K(x,y) = \alpha_0$ we have that

$$C\lambda^2 \geq \int_{-4}^{0} B(w_{\varphi_1}, (w-\varphi_1)^-)dt \geq C\,\alpha_0 \frac{\mu}{2} \iint_{t\in\Sigma} w_{\varphi_1}(x)\,dx\,dt$$

$$\geq C\,\alpha_0 \frac{\mu}{2\lambda} \iint_{t\in\Sigma} [w_{\varphi_1}(x)]^2\,dx\,dt$$

(since $w_{\varphi_1} \leq \lambda$)

In other words

$$\iint_{t\in\Sigma} [w_{\varphi_1}(x)]^2\,dx\,dt \leq \overline{C}\,\frac{\lambda^2}{\mu} \leq \overline{C}\,\lambda^{3-h}$$

for some small h (remember: $\lambda = \mu^M$).

In particular

$$\int w_{\varphi}^2(x)\,dx\,dt \leq \overline{C}\,\lambda^{3-\frac{1}{4}}$$

except for a very small subset F of t's, (say $\lambda^{1/8}$), still much smaller than $\mu \sim |\Sigma|$.

That is:

For most $t \in \Sigma$

$$\int w_{\varphi_1}^2(x)\, dx\, dt \leq \overline{C}\, \lambda^{3-\frac{1}{4}}$$

In search of an intermediate set, where w is between φ_0 and φ_2

Let us go now to w_{φ_2}.

Assume that for at least one time T_0,

$$|\{w_{\varphi_2} > 0\}| > \delta \,,$$

i.e., goes over critical for the first lemma and lets go backwards in time until we reach a slice of time $T_1 \in \Sigma$, where

$$\int w_{\varphi_1}^2(x)\,dx\,dt \leq C\,\lambda^{3-\frac{1}{4}}$$

At T_0, for the intermediate cut off, φ_1, we have

$$\int (w_{\varphi_1})^2 \geq \int (\varphi_1 - \varphi_2)^2 \chi_{\{w_{\varphi_2} > 0\}} \geq \lambda^2 \delta^4$$

(since $(\varphi_1 - \varphi_2) \geq \delta\lambda$ except for a ring of width δ)

while at T_1,

$$\int (w_{\varphi_1})^2 \leq C\,\lambda^{3-\frac{1}{2}}\ .$$

Thus, in going from T_0 backwards to T_1, $H(t) = \int (w_{\varphi_1})^2$ has crossed a range between two multiples of $\delta^4\lambda^2$, from say $\lambda^2 \frac{\delta^4}{8}$, to $\lambda^2 \frac{\delta^4}{16}$ (since $\lambda < \delta^M$).

Since $H'(t) \leq C\,\lambda^2$, in order to do so it needed in a range of times, of at least length $\sim \delta^4$.

We want to show that in this range, we pick up an intermediate set of D, of nontrivial measure, where $w_{\varphi_0} > 0$ and $w_{\varphi_2} = 0$, implying that the measure

$$\mathcal{A}_2 = |\{w_{\varphi_2} > 0\}|$$

effectively decreases some fixed amount from

$$\mathcal{A}_0 = |\{w_{\varphi_0} > 0\}|\ .$$

In these range of times, on one hand, given the gap between φ_1 and φ_2

$$|\{w_{\varphi_2} > 0\}| \leq C\,\delta$$

(if not $\int (w_{\varphi_1})^2 > \delta^4\lambda^2$) and on the other, those times for which

$$|\{w_{\varphi_0} < 0\}| \geq \mu$$

are in the exceptional subset F of very small size

$$\frac{\lambda^2\delta^4}{16} \geq \lambda^{3-\frac{1}{2}}\ , \quad \text{i.e.,}\ \frac{\delta^4}{16} \geq \lambda^{1/2}$$

Therefore, for these times in $[T_0, T_1]$ not in F, we have:

$$A(t) = |\{\varphi_0 \leq w \leq \varphi_2\}| \geq 1/2$$

That is

$$\mathcal{A}_0 = |\{w \geq \varphi_2\}| \leq |\{w \geq \varphi_0\}| - \int A(t)$$

$$\leq |\{w \geq \varphi_0\}| - \frac{1}{2}\,\frac{\delta^4}{16} = \mathcal{A}_0 - \frac{1}{2}\,\frac{\delta^4}{16}$$

Therefore, provided that the hypothesis $w^{-s/\varepsilon} \leq \lambda^2$ is satisfied in each renormalization,

$$w^{k+1} = \frac{1}{\lambda} w^k$$

we need $\dfrac{1}{\delta^4}$ steps to fall into the hypothesis of Lemma 1.

We must then choose ε, smaller than ε_0, so $\quad (\varepsilon_0 = \varepsilon_0(\lambda, \delta))$

$$2^{-s/\varepsilon} \leq \lambda^{2+\delta^2}$$

We note that the number of steps does not depend on ε_0, only in the fact that ε_0 complies with the inequality.
As we mentioned before, choosing a very small ε only means we prove Hölder continuity with a small exponent.

The oscillation theorem then follows.

Theorem

If w is bounded in $\mathbb{R}^n \times [-4, 4]$, then w is C^α in space and time in $\mathbb{R}^n \times [-2, 2]$.

Proof.

Given the sequence of cylinders $B_{2^{-k}} \times [-2^{-k}, 2^{-k}]$, we apply inductively the previous lemma to the rescalings of w or $(-w)$ depending if $w \leq \frac{1}{2}[\max(w) + \min(w)]$ at least half of the time, or $w \geq \frac{1}{2}[\max w + \min w]$ does. ☐