Trudinger-Moser inequality and beyond

Cyril Tintarev
Uppsala University

Weizmann Institute of Science, December 27 2012

Why in dimension 2 one can stop blowing bubbles and start making toy pyramids

Why in dimension 2 one can stop blowing bubbles and start making toy pyramids

- Trudinger-Moser (Yudovich, Peetre, Pohozhaev, Trudinger, Moser) inequality defines an optimal nonlinearity $\int e^{4\pi u^2}$ on a ball in Sobolev space for dimension 2, that exhibits more weak continuous behavior than its counterpart $\int |u|^{2^*}$ in higher dimensions.

Why in dimension 2 one can stop blowing bubbles and start making toy pyramids

- Trudinger-Moser (Yudovich, Peetre, Pohozhaev, Trudinger, Moser) inequality defines an optimal nonlinearity $\int e^{4\pi u^2}$ on a ball in Sobolev space for dimension 2, that exhibits more weak continuous behavior than its counterpart $\int |u|^{2^*}$ in higher dimensions.
- This suggests that the inequality can be refined. We produce several refinements, but argue that there an invariant local analog of $\int |u|^{2^*}$ does nor exist.

Why in dimension 2 one can stop blowing bubbles and start making toy pyramids

- Trudinger-Moser (Yudovich, Peetre, Pohozhaev, Trudinger, Moser) inequality defines an optimal nonlinearity $\int e^{4\pi u^2}$ on a ball in Sobolev space for dimension 2, that exhibits more weak continuous behavior than its counterpart $\int |u|^{2^*}$ in higher dimensions.

- This suggests that the inequality can be refined. We produce several refinements, but argue that there an invariant local analog of $\int |u|^{2^*}$ does nor exist.

- In the higher dimensions sequences approximating solutions to critical elliptic (Yamabe-type) problems may form concentration profiles in form of rescaled "standard bubbles". In dimension 2 the analogous sequences produce rescaled "toy pyramids".

Comparison of Sobolev and Trudinger-Moser inequalities

Sobolev ineq. $\mathcal{D}^{1,2}(\mathbb{R}^N)$, $N>2$	Trudinger-Moser, $H^1_0(\mathbb{D})$, $\mathbb{D}\subset\mathbb{R}^2$
$\sup_{\\|\nabla u\\|_2\le 1}\int_{\mathbb{R}^N}\|u\|^{2^*}\,\mathrm{d}x<\infty$	$\sup_{\\|\nabla u\\|_2\le 1}\int_{\mathbb{D}} e^{4\pi u^2}\,\mathrm{d}x<\infty$
$h\uparrow\infty, \sup\int_{\mathbb{R}^N} h(\|u\|)\|u\|^{2^*}\,\mathrm{d}x=\infty$	$\sup\int_{\mathbb{D}} h(\|u\|)e^{4\pi u^2}\,\mathrm{d}x=\infty$
$\mathcal{D}^{1,2}\hookrightarrow L^{2^*,2}$, Peetre '66 $\\|u\\|^2_{2^*,2}=\int\left\|\frac{u^*}{r}\right\|^2\mathrm{d}x$ (Hardy)	$H^1_0(\mathbb{D})\hookrightarrow L^{\infty,2,-1}$, Brezis-Wainger $\\|u\\|^2_{\infty,2,-1}=\int\left\|\frac{u^*}{r\log\frac{e}{r}}\right\|^2\mathrm{d}x$ (Leray)
$L^{2^*,2}\hookrightarrow L^{2^*}=L^{2^*,2^*}\hookrightarrow L^{2^*,\infty}$ $\int\|u\|^{2^*}$ no weak continuity at *any* u	$L^{\infty,2,-1}\hookrightarrow L^{\infty,\infty,-1/2}=\exp L^2$ $\int e^{4\pi u^2}$ "almost" weakly cont.

Weak continuity of the Moser functional

Lions' compactness result. Let $u_k \rightharpoonup u$ in $H_0^1(\mathbb{D})$ and $\|\nabla u_k\|_2 \leq 1$.
$J(u) = \int_{\mathbb{D}}(e^{4\pi u^2} - 1)\mathrm{d}x$.

- If $\|\nabla u_k\|_2 \leq \alpha < 1$, or $u \neq 0$, then $J(u_k) \to J(u)$.

Weak continuity of the Moser functional

Lions' compactness result. Let $u_k \rightharpoonup u$ in $H_0^1(\mathbb{D})$ and $\|\nabla u_k\|_2 \leq 1$.
$J(u) = \int_{\mathbb{D}}(e^{4\pi u^2} - 1)\mathrm{d}x$.

- If $\|\nabla u_k\|_2 \leq \alpha < 1$, or $u \neq 0$, then $J(u_k) \to J(u)$.
- If $u = 0$, $\|\nabla u_k\|_2 \to 1$, and the singular support of $w - \lim |\nabla u|^2 dx$ is anything but a single point, then $J(u_k) \to J(u)$.

Weak continuity of the Moser functional

Lions' compactness result. Let $u_k \rightharpoonup u$ in $H_0^1(\mathbb{D})$ and $\|\nabla u_k\|_2 \leq 1$.
$J(u) = \int_{\mathbb{D}}(e^{4\pi u^2} - 1)\mathrm{d}x$.

- If $\|\nabla u_k\|_2 \leq \alpha < 1$, or $u \neq 0$, then $J(u_k) \to J(u)$.
- If $u = 0$, $\|\nabla u_k\|_2 \to 1$, and the singular support of $w - \lim |\nabla u|^2 dx$ is anything but a single point, then $J(u_k) \to J(u)$.
- Adimurthi and CT (Annali SNS Pisa, to appear): $J(u_k) \to J(u)$ unless

$$\|\nabla(u_k - \mu_{t_k}(\cdot - y_k))\|_2 \to 0$$

for some $y_k \in \mathbb{D}$ and $t_k \to 0$, where

$$\mu_t(x) \stackrel{\text{def}}{=} (2\pi)^{-\frac{1}{2}}(\log\frac{1}{t})^{-\frac{1}{2}} \min\left\{\log\frac{1}{|x|}, \log\frac{1}{t}\right\}, \quad t \in (0,1),\ x \in \mathbb{D}.$$

(Moser function). The condition is *still not necessary*: convergence of $J(\lambda_k \mu_{t_k})$ with $\lambda_k \to 1$ depends on the sequence λ_k (Adimurthi & Prashanth)

Perfectly critical nonlinearity

- For $N > 2$ one can derive the critical nonlinearity from the Hardy inequality $\int_{\mathbb{R}^N} |\nabla u|^2 \geq \left(\frac{N-2}{2}\right)^2 \int_{\mathbb{R}^N} \frac{u^2}{r^2}$. Hardy functional lacks weak continuity, but only on sequences concentrating at zero.

Perfectly critical nonlinearity

- For $N > 2$ one can derive the critical nonlinearity from the Hardy inequality $\int_{\mathbb{R}^N} |\nabla u|^2 \geq \left(\frac{N-2}{2}\right)^2 \int_{\mathbb{R}^N} \frac{u^2}{r^2}$. Hardy functional lacks weak continuity, but only on sequences concentrating at zero.

.

- By using rearrangements, one forces the concentration to occur in the origin: $\int \frac{u^{*2}}{r^2}$. This is the standard definition of the $L^{2^*,2}$-quasinorm.

Perfectly critical nonlinearity

- For $N > 2$ one can derive the critical nonlinearity from the Hardy inequality $\int_{\mathbb{R}^N} |\nabla u|^2 \geq \left(\frac{N-2}{2}\right)^2 \int_{\mathbb{R}^N} \frac{u^2}{r^2}$. Hardy functional lacks weak continuity, but only on sequences concentrating at zero.

.

- By using rearrangements, one forces the concentration to occur in the origin: $\int \frac{u^{*2}}{r^2}$. This is the standard definition of the $L^{2^*,2}$-quasinorm.

- Radial estimate: $\sup_{r>0} u^*(r) r^{\frac{N-2}{2}} \leq C\|\nabla u\|_2$. The left hand side is the $L^{2^*,\infty}$- norm (weak L^{2^*}, Marcinkiewicz M_{2^*}). Hölder inequality defines quasinorms for a family of Lorentz spaces, $L^{2^*,q}$, $q > 2$. For radial functions all these quasinorms are invariant with respect to dilations $u \mapsto t^{\frac{N-2}{2}} u(t\cdot)$, thus no compactness.

Perfectly critical nonlinearity

- For $N > 2$ one can derive the critical nonlinearity from the Hardy inequality $\int_{\mathbb{R}^N} |\nabla u|^2 \geq \left(\frac{N-2}{2}\right)^2 \int_{\mathbb{R}^N} \frac{u^2}{r^2}$. Hardy functional lacks weak continuity, but only on sequences concentrating at zero.

.

- By using rearrangements, one forces the concentration to occur in the origin: $\int \frac{u^{*2}}{r^2}$. This is the standard definition of the $L^{2^*,2}$-quasinorm.
- Radial estimate: $\sup_{r>0} u^*(r) r^{\frac{N-2}{2}} \leq C\|\nabla u\|_2$. The left hand side is the $L^{2^*,\infty}$- norm (weak L^{2^*}, Marcinkiewicz M_{2^*}). Hölder inequality defines quasinorms for a family of Lorentz spaces, $L^{2^*,q}$, $q > 2$. For radial functions all these quasinorms are invariant with respect to dilations $u \mapsto t^{\frac{N-2}{2}} u(t\cdot)$, thus no compactness.
- Imbedding into $L^{2^*,2}$ is optimal in the class of RI spaces (Peetre 1966), but the quasinorm of $L^{2^*,2^*}$coincides with the L^{2^*}- norm.

Dilation-invariant nonlinearity

Same derivation in dimension 2 can be carried out with a twist and only at 90%. First of all there is no function space $\mathcal{D}^{1,2}(\mathbb{R}^2)$. Its role rather convincingly is taken over by $H_0^1(\mathbb{D})$ with the norm $\|\nabla u\|_2$.
On the unit disk we have analogs of translations and dilations acting on $H_0^1(\mathbb{D})$ and preserving $\|\nabla u\|_2$.

Dilation-invariant nonlinearity

Same derivation in dimension 2 can be carried out with a twist and only at 90%. First of all there is no function space $\mathcal{D}^{1,2}(\mathbb{R}^2)$. Its role rather convincingly is taken over by $H_0^1(\mathbb{D})$ with the norm $\|\nabla u\|_2$.
On the unit disk we have analogs of translations and dilations acting on $H_0^1(\mathbb{D})$ and preserving $\|\nabla u\|_2$.
Translations: $u\left(\frac{z-\zeta}{1-\bar{\zeta}z}\right)$, $\zeta \in \mathbb{D}$.

Dilation-invariant nonlinearity

Same derivation in dimension 2 can be carried out with a twist and only at 90%. First of all there is no function space $\mathcal{D}^{1,2}(\mathbb{R}^2)$. Its role rather convincingly is taken over by $H_0^1(\mathbb{D})$ with the norm $\|\nabla u\|_2$.
On the unit disk we have analogs of translations and dilations acting on $H_0^1(\mathbb{D})$ and preserving $\|\nabla u\|_2$.
Translations: $u\left(\frac{z-\zeta}{1-\bar{\zeta}z}\right)$, $\zeta \in \mathbb{D}$.
Dilations (a semigroup): $j^{-1/2}u(z^j)$, $j \in \mathbb{N}$, extend to radial functions to the group

$$h_s u(r) \stackrel{\text{def}}{=} s^{-\frac{1}{2}} u(r^s), \; s > 0.$$

The counterpart of Hardy inequality is the Leray inequality (dilation invariant)

$$\|\nabla u\|_2^2 \geq \frac{1}{4} \int_{\mathbb{D}} \frac{u^{*2}}{(r \log \frac{1}{r})^2} \, \mathrm{d}x \quad \text{Leray, 1933}$$

Replacing $1/r$ with e/r we have the quasinorm of the Lorentz-Zygmund space $L^{\infty,2;0}$.

The counterpart of Hardy inequality is the Leray inequality (dilation invariant)

$$\|\nabla u\|_2^2 \geq \frac{1}{4}\int_{\mathbb{D}} \frac{u^{*2}}{(r\log\frac{1}{r})^2}\,\mathrm{d}x \quad \text{Leray, 1933}$$

Replacing $1/r$ with e/r we have the quasinorm of the Lorentz-Zygmund space $L^{\infty,2;0}$.

Radial estimate: $\sup_{r>0} \frac{u^*(r)}{\sqrt{\log\frac{1}{r}}} \leq \frac{1}{\sqrt{2\pi}}\|\nabla u\|_2$. $L^{\infty,\infty;-1/2}$ coincides with the Orlicz space $\exp L^2$.

The counterpart of Hardy inequality is the Leray inequality (dilation invariant)

$$\|\nabla u\|_2^2 \geq \frac{1}{4}\int_{\mathbb{D}} \frac{u^{*2}}{(r\log\frac{1}{r})^2}\,\mathrm{d}x \quad \text{Leray, 1933}$$

Replacing $1/r$ with e/r we have the quasinorm of the Lorentz-Zygmund space $L^{\infty,2;0}$.
Radial estimate: $\sup_{r>0}\frac{u^*(r)}{\sqrt{\log\frac{1}{r}}} \leq \frac{1}{\sqrt{2\pi}}\|\nabla u\|_2$. $L^{\infty,\infty;-1/2}$ coincides with the Orlicz space $\exp L^2$.
The scale of Lorentz-Zygmund spaces $L^{\infty,p;1/p-1/2}$ is similar to the Lorentz scale and the $L^{\infty,\infty;-1/2}$-quasinorm is a clear analogue of the L^{2^*}-norm. As fiascos come, this is one of the nicest, a true Pyrrhic victory.

Why Möbius transformations?

- Poincaré disk model of the hyperbolic space $\mathbb{H}^2$: $d\mu = \frac{4}{(1-|x|^2)^2}dx$; $\|du\|^2_{L^2(\mathbb{H}^2)} = \|\nabla u\|_2^2$;

$$\sup_{u\in \dot{H}^1(\mathbb{H}^2), \|du\|_2\le 1} \int_{\mathbb{H}^2} (e^{4\pi u^2} - 1)d\mu < \infty.$$

"Hyperbolic refinement" of Trudinger-Moser: Adimurthi &CT, Sandeep and Mancini, 2010

Why Möbius transformations?

- Poincaré disk model of the hyperbolic space $\mathbb{H}^2$: $d\mu = \frac{4}{(1-|x|^2)^2}dx$; $\|du\|^2_{L^2(\mathbb{H}^2)} = \|\nabla u\|_2^2$;

$$\sup_{u\in \dot{H}^1(\mathbb{H}^2), \|du\|_2\le 1} \int_{\mathbb{H}^2} (e^{4\pi u^2} - 1)d\mu < \infty.$$

  "Hyperbolic refinement" of Trudinger-Moser: Adimurthi &CT, Sandeep and Mancini, 2010

- Möbius transformations η_ζ are isometries on $\mathbb{H}^2$.
  "Translation-invariant" Lorentz-Zygmund spaces have to use rearrangements relative to μ. This helps the quasinorms, but not the Trudinger-Moser functional!

Forcing invariance:

Creating an invariant functional in $H^1_{0,r}(\mathbb{D})$ by limit:
$\lim_{s\to 0} J(h_s u) - J(0) = 0$ and $\lim_{s\to\infty} J(h_s u) - J(0) = 2\pi \mathbf{1}_M(u)$ where $M = \{\mu_t\}_{t\in(0,1)}$.

Forcing invariance:

Creating an invariant functional in $H^1_{0,r}(\mathbb{D})$ by limit:
$\lim_{s\to 0} J(h_s u) - J(0) = 0$ and $\lim_{s\to\infty} J(h_s u) - J(0) = 2\pi\mathbf{1}_M(u)$ where $M = \{\mu_t\}_{t\in(0,1)}$.
If the functional $\int_{\mathbb{D}} F(r,u)d\mu$ is both Möbius-translation invariant and dilation invariant, then $F = 0$.

More refinements of Trudinger-Moser inequality

$$\sup_{u \in H_0^1(\mathbb{D}), \|\nabla u\|_2 \leq 1} \int_{\mathbb{D}} e^{4\pi(1+\lambda\|u\|_2^2)u^2} dx < \infty, \ \lambda < \lambda_1(\mathbb{D}) = 5.7...$$

More refinements of Trudinger-Moser inequality

$$\sup_{u\in H_0^1(\mathbb{D}),\|\nabla u\|_2\leq 1}\int_{\mathbb{D}} e^{4\pi(1+\lambda\|u\|_2^2)u^2}dx<\infty,\ \lambda<\lambda_1(\mathbb{D})=5.7...$$

- Under general conditions (*the singularity of V at the origin has to be* ***sub-Leray***) the inequality

$$\sup_{u\in H_0^1(\Omega)\|\nabla u\|^2-\int V(r)u^2\leq 1}\int_{\mathbb{D}} e^{4\pi u^2}<\infty$$

holds if and only if the quadratic form is subcritical in the sense of Agmon (positive without a virtual bound state). CT (preprint 2012)

- Expressing this with the Orlicz norm is not kind to the constant 4π.
- Adimurthi-Druet inequality follows from the case of $V(r)=\lambda$. Wang and Ye (2012) have $V=\frac{1}{(1-r^2)^2}$.

Trudinger-Moser on the plane

Bernhard Ruf (2005):

$$c(\lambda) = \sup_{u \in H^1(\Omega) \|\nabla u\|^2 + \lambda \int u^2 \leq 1} \int_{\mathbb{R}^2} e^{4\pi u^2} < \infty, .$$

Conjecture: $\int u^2$ can be replaced by a weaker term, as long as *some* coercivity is sustained. Perhaps not by $(\int_{\mathbb{D}} u)^2$?
In general, it's time to look for the counterpart of the Caffarelli-Cohn-Nirenberg inequalities.

- Critical sequences for the limit Sobolev functional in higher dimensions, $\|\nabla u\|^2 - \int |u|^{\frac{2N}{N-2}}$, develop elementary concentrations ("bubbles") $t^{\frac{N-2}{2}} w(t(\cdot - y))$. There is only one possible positive bubble, the "standard bubble"
- $w(x) = \frac{1}{(1+x^2)^{\frac{N-2}{2}}}$.
- Existence results require elimination of concentration. Once concentration is eliminated, the sequence converges. "How to blow bubbles" by Brezis and Coron, 1984.
- What happens in the case of $\|\nabla u\|^2 - \int e^{4\pi u^2}$?

- The elementary concentration: $t^{1/2}w(|x-y|^{1/t})$. (Druet, Struwe and others studied sequences of solutions that allowed Euclidean blowups).
- Convergence after elimination of bubbles: Adimurthi and CT, to appear in Annali SNS Pisa, radial case Adimurthi, do Ó and CT (2010)
- Instead of a standard bubble the profiles are infinitely many "toy pyramids" (David Costa and CT, preprint 2012).

How to build toy pyramids

- A radial function $\mu_{C_+,C_-} \in H_0^1(B)$, parametrized by closed disjoint sets $C_+, C_- \subset (0,1)$, is called a Moser-Carleson-Chang tower if $\mu_{C_+,C_-}(r) =$
$$\begin{cases} \sqrt{\frac{1}{2\pi}\log\frac{1}{r}}, & r \in C_+, \\ -\sqrt{\frac{1}{2\pi}\log\frac{1}{r}}, & r \in C_-, \\ A_n + B_n\log\frac{1}{r}, & r \in (a_n, b_n) \text{ (a connected complement of } (0,1)\setminus(C_+ \cup \end{cases}$$
- When the set C_+ consists of a single point and $C_- = \emptyset$, this is the original Moser function, and it uniquely minimizes$\|\nabla\mu_{C_+,C_-}\|_2$.
- The coefficients A_n, B_n are defined uniquely by the requirement of continuity.
- The function $\mu_{C_+,C_-}(r)$ has continuous derivative at every point of $(0,1)$ except the endpoints$\{a_n, b_n\}$.
- The number of times the function μ_{C_+,C_-} on $(0,1)$ changes sign does not exceed $\|\nabla\mu_{C_+,C_-}\|_2^2 - 1$.
- Restriction on C_-, C_+ : $\sum\frac{\sigma_n-1}{\sigma_n+1} < \infty$ where $\sigma_n = \sqrt{\frac{\log\frac{1}{a_n}}{\log\frac{1}{b_n}}}$.