Nonlinear Analysis of the Yamabe Problem

Zizheng Fang
Supervisor: Prof. Jason Lotay

October 17, 2025

Acknowledgements

I would like to thank my supervisor, Professor Jason Lotay, both for suggesting this fascinating topic and for patient guidance throughout.

This undergraduate summer research project was supported by the Mathematical Institute (University of Oxford) and by Exeter College, Oxford.

Contents

1 Introduction

1.1 The Yamabe Problem

The *Yamabe problem* asks whether, for every compact Riemannian manifold (M, g) of dimension $n \geq 3$, there exists a conformal metric $\tilde{g} = \rho g$ that has constant scalar curvature. A proof was announced in 1960 by Hidehiko Yamabe [Yam60], but an error was found in his proof, and what came to be known as the Yamabe problem was not solved in all its cases until 1984.

The solution has two essentially independent parts: an analytic condition that guarantees a scalar-flat conformal metric exists, and subtle geometric arguments to show this condition holds for all M not conformally equivalent to (S^n, g_{S^n}). Since S^n has constant scalar curvature, it follows that the Yamabe problem is solvable for all M.

The Yamabe Equation The *conformal Laplacian* is defined by $\square = -\Delta + aS$, where $a = \frac{n-2}{4(n-1)}$ and Δ is the Laplace-Beltrami operator. Writing $p = \frac{2n}{n-2}$ and $\varphi^{p-2} = \rho$, it can be computed that

$$\tilde{\square} u = \varphi^{1-p} \square (\varphi u). \tag{1.1}$$

In particular, $\tilde{S} = a^{-1} \tilde{\square}(1) = a^{-1} \varphi^{1-p} \square \varphi$, so $\tilde{S}$ is constant if and only if there exists some constant λ such that

$$\square \varphi = \lambda \varphi^{p-1}. \tag{1.2}$$

This is the *Yamabe equation.*

The Yamabe Functional It turns out the Yamabe equation is the Euler-Lagrange equation of the *Yamabe functional*, defined by

$$Q(\tilde{g}) = \frac{\int_M \tilde{S} \mathrm{d}V_{\tilde{g}}}{\left(\int_M \mathrm{d}V_{\tilde{g}} \right)^{2/p}}, \tag{1.3}$$

where $\tilde{g}$ varies over all conformal metrics.

It can be rewritten in terms of φ as

$$Q_g(\varphi) := aQ(\tilde{g}) = \frac{E(\varphi)}{\|\varphi\|_p^2},$$

where

$$E(\varphi) := \int_M \varphi \Box \varphi \mathrm{d}V_g = \int_M |\nabla \varphi|^2 + aS\varphi^2 \mathrm{d}V_g.$$

Given $\psi \in C^\infty(M)$, we may compute

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_g(\varphi + t\psi) = 2\|\varphi\|_p^{-2} \int_M (\Box\varphi - \|\varphi\|_p^{-p} E(\varphi)\varphi^{p-1})\psi \mathrm{d}V_g,$$

so φ is a critical point of the Yamabe functional Q_g if and only if the Yamabe equation (1.2) is satisfied. (The converse holds because $\Box\varphi = \lambda\varphi^{p-1}$ implies $E(\varphi) = \lambda\|\varphi\|_p^p$.)

The Yamabe Invariant By Hölder's inequality,

$$\left| \int_M aS\varphi^2 \mathrm{d}V_g \right| \leq C\|\varphi\|_p^2,$$

so Q_g is bounded below, which means Q is bounded below.

The *Yamabe invariant* is defined by

$$\begin{aligned} \lambda(M) &= \inf\{Q_g(\varphi) : \varphi > 0, \varphi \in C^\infty(M)\} \\ &= \inf\{aQ(\tilde{g}) : \tilde{g} \text{ conformal to } g\}. \end{aligned} \tag{1.4}$$

It depends only on the conformal class of (M, g).

Without loss of generality, we shall always assume M to be connected. Then the following results hold, and in combination they solve the Yamabe problem in the affirmative.

Theorem 1.1. *If $\lambda(M) < \lambda(S^n)$, then $\lambda(M) = Q_g(\varphi_0)$ for some $\varphi_0 > 0$, $\varphi_0 \in C^\infty(M)$.*

Theorem 1.2. *If M has dimension $n \geq 6$ and is not locally conformally flat, then $\lambda(M) < \lambda(S^n)$.*

Theorem 1.3. *If M either has dimension $n = 3, 4, 5$, or is locally conformally flat, then $\lambda(M) < \lambda(S^n)$ unless M is conformal to S^n.*

If the condition for Theorem 1.1 is satisfied, then φ_0 is a critical point of the Yamabe functional, so the conformal metric $\tilde{g} := \varphi_0^{p-2} g$ has constant scalar curvature. Theorems 1.2 and 1.3 guarantee that the condition $\lambda(M) < \lambda(S^n)$ is satisfied by all M not conformal to S^n.

The rest of this essay is organized as follows.

In Section 1.2, we explain the significance of the value $\lambda(S^n)$ and relate it to the best constant problem for the Sobolev inequality $\|u\|_{L^p(\mathbf{R}^n)} \le C\|\nabla u\|_{L^2(\mathbf{R}^n)}$.

Section 2 describes without proof solutions to Theorems 1.2 and 1.3. These involve explicit formulas of optimizers of the Sobolev inequality, which explains why it is important to find these optimizers in the first place.

Section 3 presents two proofs of the classification of optimizers of the Sobolev inequality.

Section 4 is concerned with the traditional approach to Theorem 1.1, going back to Yamabe's original (flawed) paper [Yam60], that derives a solution to the Yamabe equation as a limit of solutions to $\Box\varphi = \lambda\varphi^{s-1}$ as $s \uparrow p$. We also sketch a (somewhat unsatisfactory) argument along the same lines that an optimizer exists for the Sobolev inequality.

Section 5 describes methods to prove the sharp Sobolev inequality and Theorem 1.1 without having to pass through subcritical solutions.

1.2 The Role of the Sharp Sobolev Inequality

Let us recall that for any Riemannian manifold M, $D^{1,2}(M)$ denotes the space of L^p functions on M with L^2 gradient, endowed with the norm $\|u\|_{D^{1,2}(M)} = \|\nabla u\|_2$. If M has finite volume, then $D^{1,2}(M) = W^{1,2}(M)$ with an equivalent norm.

Theorem 1.4. *For all $n \ge 3$, there exists a constant Λ depending on n such that*

I. *For all $\varphi \in D^{1,2}(\mathbf{R}^n)$,*

$$\int_{\mathbf{R}^n} |\nabla\varphi|^2 \mathrm{d}x \ge \Lambda\Big(\int_{\mathbf{R}^n} \varphi^p \,\mathrm{d}x\Big)^{2/p} \tag{1.5}$$

with equality if and only if $\varphi = c(\mu^2 + |x-a|^2)^{-\frac{n-2}{2}}$ for some $\mu > 0$ and $a \in \mathbf{R}^n$.

II. *For all $\varphi \in W^{1,2}(S^n)$,*

$$\int_{S^n} |\nabla\varphi|^2 + \frac{n(n-2)}{4}\varphi^2 \mathrm{d}V_{g_{S^n}} \ge \Lambda\Big(\int_{S^n} \varphi^p \,\mathrm{d}V_{g_{S^n}}\Big)^{2/p} \tag{1.6}$$

with equality if and only if $\varphi(\omega) = c(1 - \xi\cdot\omega)^{-\frac{n-2}{2}}$ for some $\xi \in \mathbf{R}^{n+1}$, $|\xi| < 1$.

The sphere S^n has constant scalar curvature $S = n(n-1)$, so $aS \equiv \frac{n(n-2)}{4}$, and thus the left hand side of (1.6) equals $E(\varphi)$. Statement II says that $Q_{g_{S^n}}$ has sharp lower bound $\lambda(S^n) = \Lambda$, and $\varphi \equiv 1$ is a minimizer (which is to say, g_{S^n} minimizes Q in the conformal class of S^n). This explains the centrality of $\lambda(S^n)$ to the Yamabe problem.

Let $P = (0,\dots,0,1)$ be the north pole on $S^n \subset \mathbf{R}^{n+1}$. By identifying $S^n \setminus \{P\}$ with $\mathbf{R}^n$ via the stereographic projection $\pi : S^n \setminus \{P\} \longrightarrow \mathbf{R}^n$, given by

$$\pi(\omega^1, \ldots, \omega^{n+1}) := \frac{1}{1-\omega^{n+1}}(\omega^1, \ldots, \omega^n),$$

which is conformal:

$$(\pi^{-1})^* g_{S^n} = 4\big(|x|^2+1\big)^{-2} g_{\mathbf{R}^n},$$

we have the following

Lemma 1.5. *In Theorem 1.4, statements I and II are equivalent.*

Proof. Let $u^{p-2} = 4\big(|x|^2+1\big)^{-2}$. Then for $\varphi > 0$, $\varphi \in C^\infty(S^n)$, we have

$$(\pi^{-1})^*(\varphi^{p-2} g_{S^n}) = (u(\pi^{-1})^*\varphi)^{p-2} g_{\mathbf{R}^n},$$

and thus,

$$E(\varphi) = \int S \mathrm{d}V_g = E(u(\pi^{-1})^*\varphi),$$

$$\|\varphi\|^2_{L^p(S^n)} = \left(\int \mathrm{d}V_g\right)^{2/p} = \|u(\pi^{-1})^*\varphi\|^2_{L^p(\mathbf{R}^n)}.$$

Using the naturality of $\Box$ and its transformation formula (1.1) under conformal change of metric, it can be verified that the above hold for all $\varphi \in C^\infty(S^n)$.

Motivated by this, we shall write $\overline{\varphi} = u(\pi^{-1})^*\varphi$ whenever φ is a function on S^n. By approximating with $C^\infty(S^n)$ functions, it can be shown that $\varphi \leftrightarrow \overline{\varphi}$ is a continuous bijective map between $W^{1,2}(S^n)$ and $D^{1,2}(\mathbf{R}^n)$, satisfying $E(\varphi) = E(\overline{\varphi})$ and $\|\varphi\|_p^2 = \|\overline{\varphi}\|_p^2$.

Hence, $E(\varphi)/\|\varphi\|_p^2$ and $\|\nabla\overline{\varphi}\|_2^2/\|\overline{\varphi}\|_p^2$ have the same infimum $\Lambda > 0$, and their sets of minimizing functions are in bijective correspondance via $\varphi \leftrightarrow \overline{\varphi}$. $\Box$

2 Construction of Test Functions

Theorem 1.2. *If M has dimension $n \geq 6$ and is not locally conformally flat, then $\lambda(M) < \lambda(S^n)$.*

Theorem 1.3. *If M either has dimension $n = 3, 4, 5$, or is locally conformally flat, then $\lambda(M) < \lambda(S^n)$ unless M is conformal to S^n.*

In both cases a "test function" φ such that $Q_g(\varphi) < \lambda(S^n)$ can be constructed explicitly.

Recall that the Weyl tensor W is a conformal invariant that vanishes identically if and only if M is locally conformally flat. It helps describe the local conformal geometry at a point $P \in M$.

Theorem 2.1. *Let M be a Riemannian manifold and let $P \in M$. For each $N \geq 2$ there exists a conformal metric g on M, such that in the normal coordinates centered at P,*

$$\det(g_{ij}) = 1 + O(r^N).$$

If $N \geq 5$, then in these coordinates,

$$S = O(r^2),$$
$$\Delta S = \frac{1}{6}|W|^2.$$

Whenever conformal normal coordinates are used, N is assumed to be sufficiently large.

For Theorem 1.2, a local test function suffices. The construction is based on the family of optimizers

$$u_\varepsilon(x) := \left(\frac{\varepsilon}{|x|^2 + \varepsilon^2}\right)^{\frac{n-2}{2}}, \varepsilon > 0 \tag{2.1}$$

for the $\mathbf{R}^n$ Sobolev inequality (1.5). Let η_1 be a smooth radial cutoff function such that $0 \leq \eta_1 \leq 1$, $\eta_1 \equiv 1$ on B_1, and $\operatorname{supp} \eta_1 \subset B_2$; for $\rho > 0$, let $\eta_\rho(x) = \eta_1(\rho^{-1}x)$.

Theorem 2.2. *Suppose M has dimension $n \geq 6$ and is not locally conformally flat. Let P be a point such that $|W(P)| > 0$, let ρ be small enough that $\eta = \eta_\rho$ fits into a conformal normal neighbourhood of P, and let*

$$\varphi(x) = \begin{cases} \eta(x)u_\varepsilon(x), & \text{if } x \in B_{2\rho}, \\ 0, & \text{if } x \in M \setminus B_{2\rho}, \end{cases} \tag{2.2}$$

where $\varepsilon > 0$. Then $Q_g(\varphi) < \Lambda$ for sufficiently small ε.

The proof of Theorem 1.3 is more difficult, since it uses a global test function involving the Green function for $\Box$, and the analysis also relies on the positive mass theorem of general relativity.

Lemma 2.3. *Suppose $\lambda(M) > 0$. For $P \in M$, there exists a function $G_P > 0$, $G_P \in C^\infty(M \setminus \{P\})$, such that $\Box G_P = \delta_P$. Moreover, in normal coordinates at P,*

$$G_P(x) = \frac{1}{(n-2)\omega_{n-1}} r^{2-n}(1 + o(1)).$$

Proposition 2.4. *Suppose M has dimension $n = 3, 4, 5$ or is locally conformally flat, and suppose $\lambda(M) > 0$. Let $P \in M$, and let $G = (n-2)\omega_{n-1}G_P$. Then in a conformal normal neighbourhood,*

$$G(x) = r^{2-n} + A + \alpha(x),$$

where A is a constant and $\alpha(x) = O(r)$, $\nabla\alpha(x) = O(1)$.

The following result can be derived from the positive mass theorem.

Proposition 2.5. *In Proposition 2.4, $A \geq 0$, with equality if and only if M is conformal to S^n.*

Theorem 2.6. *Suppose M has dimension $n = 3, 4, 5$ or is locally conformally flat, and suppose $\lambda(M) > 0$. Consider a conformal normal neighbourhood at $P \in M$, and let $G(x) = r^{2-n} + A + \alpha(x)$ be as in Proposition 2.4. Let ρ be small enough, let $\eta = \eta_\rho$, and set*

$$\varphi(x) = \begin{cases} u_\varepsilon(x), & \text{if } x \in B_\rho, \\ \varepsilon_0(G(x) - \eta(x)\alpha(x)), & \text{if } x \in B_{2\rho} \setminus B_\rho, \text{ and} \\ \varepsilon_0 G(x), & \text{if } x \in M \setminus B_{2\rho}, \end{cases} \tag{2.3}$$

where $\varepsilon > 0$, and $\varepsilon_0 = u_\varepsilon(\rho)/(\rho^{2-n} + A)$. Then $Q_g(\varphi) < \Lambda$ for sufficiently small ε.

3 The Sobolev Inequality: Classification of Optimizers

We recall that there are two equivalent versions of the sharp Sobolev inequality:

Theorem 1.4. *For all $n \geq 3$, there exists a constant Λ depending on n such that*

I. *For all $\varphi \in D^{1,2}(\mathbf{R}^n)$,*

$$\int_{\mathbf{R}^n} |\nabla\varphi|^2 \mathrm{d}x \geq \Lambda \left(\int_{\mathbf{R}^n} \varphi^p \,\mathrm{d}x \right)^{2/p} \tag{1.5}$$

with equality if and only if $\varphi = c(\mu^2 + |x - a|^2)^{-\frac{n-2}{2}}$ for some $\mu > 0$ and $a \in \mathbf{R}^n$.

II. *For all $\varphi \in W^{1,2}(S^n)$,*

$$\int_{S^n} |\nabla\varphi|^2 + \frac{n(n-2)}{4}\varphi^2 \mathrm{d}V_{g_{S^n}} \geq \Lambda \left(\int_{S^n} \varphi^p \,\mathrm{d}V_{g_{S^n}} \right)^{2/p} \tag{1.6}$$

with equality if and only if $\varphi(\omega) = c(1 - \xi \cdot \omega)^{-\frac{n-2}{2}}$ for some $\xi \in \mathbf{R}^{n+1}$, $|\xi| < 1$.

The initial proof by Talenti [Tal76] uses rearrangement to reduce the $\mathbf{R}^n$ inequality to the case of radial functions, and applies a difficult one-variable integral inequality to finish the proof. Many different proofs have subsequently been found; most of them start by assuming the existence of a minimizer to $E(\varphi)/\|\varphi\|_p^2$, and then prove that it must be of the prescribed form.

An early result classifying such minimizers is the following.

Theorem 3.1 ([Oba71, Proposition 6.1])**.** *A conformal metric on S^n has constant scalar curvature if and only if it is isometric to (S^n, cg_{S^n}), $c > 0$.*

Obata's result implies that if $Q_{g_{S^n}}$ has a minimizer $\varphi > 0$, $\varphi \in C^\infty(S^n)$, then 1 is also a minimizer, since $\lambda(S^n) = Q(\varphi^{p-2}g_{S^n}) = Q(cg_{S^n}) = Q_{g_{S^n}}(1)$.

Lee and Parker [LP87] tried to find such a minimizer by renormalizing solutions to $\Box\varphi_s = \lambda\varphi_s^{s-1}$, $2 < s < p$, so that they would converge to a solution to $\Box\varphi = \lambda\varphi^{p-1}$. Unfortunately, in an attempt to deal with the difficult behaviour of $\Box_{S^n}$ under the dilation $(\pi^{-1}\delta_\alpha\pi)^*$ (where $\delta_\alpha(x) = \alpha^{-1}x$), they made a mistake by invoking Fatou's lemma with the inequality sign pointing in the wrong direction.

By renormalizing on $\mathbf{R}^n$, and with the help of Gidas-Ni-Nirenberg's symmetry result [GNN79], Schoen and Yau [SY94] showed the existence of a minimizer $\varphi > 0, \varphi \in C^\infty(\mathbf{R}^n)$. Obata's theorem does not immediately apply, since it is a strictly weaker result: with the notation of Lemma 1.5, $\varphi > 0, \varphi \in C^\infty(S^n)$ imply $\overline{\varphi} > 0, \overline{\varphi} \in C^\infty(\mathbf{R}^n)$, but not the other way round. A suitable adaptation of Obata's proof was developed by Chang, Gursky and Yang [CGY03]. We present their proof in Section 3.1.

In Section 3.2, we give an alternative proof by Rupert Frank and Elliott Lieb [FL12], which proceeds by examining the second variation of the Yamabe functional—making no use at all of the regularity of φ. A key tool is a "center of mass condition" that any nonnegative minimizer can be made to satisfy with a suitable transformation.

3.1 Obata's Method

We reproduce the proof of Obata's theorem given in [LP87, Proposition 3.1].

Theorem 3.1 ([Oba71, Proposition 6.1]). *A conformal metric on S^n has constant scalar curvature if and only if it is isometric to (S^n, cg_{S^n}), $c > 0$.*

Proof. Let $\overline{g} = g_{S^n}$, and let $g = v^{-2}\overline{g}$ be a conformal metric with constant scalar curvature. By the Killing-Hopf theorem, it suffices to prove that g has constant curvature.

Let $B_{jk} = R_{jk} - n^{-1}Sg_{jk}$ be the traceless Ricci tensor. Since $\overline{g}$ has constant curvature, we compute

$$0 = \overline{B}_{jk} = B_{jk} + (n-2)v[(v^{-1})_{jk} - n^{-1}\Delta(v^{-1})g_{jk}], \tag{3.1}$$

where the covariant derivatives and the index liftings are performed with respect to g. Now since S is constant, using the contracted Bianchi identity

$$R^i{}_{l,i} = \frac{1}{2}S_{,l}$$

we deduce that

$$B^i{}_{l,i} = R^i{}_{l,i} - n^{-1}(S^{,i}g_{il} + Sg^i{}_{l,i}) = 0. \tag{3.2}$$

Combined with the fact that B is traceless and symmetric, we compute

$$\begin{aligned}\int_{S^n} B_{jk}B^{jk}v^{-1}\mathrm{d}V_g &= -(n-2)\int_{S^n} B^{jk}[(v^{-1})_{jk} - n^{-1}\Delta(v^{-1})g_{jk}]\mathrm{d}V_g \\ &= -(n-2)\int_{S^n} B^{jk}(v^{-1})_{jk}\mathrm{d}V_g \\ &= (n-2)\int_{S^n} B^{jk}{}_{,k}(v^{-1})_j\mathrm{d}V_g \\ &= 0.\end{aligned}$$

Hence $|B| \equiv 0$ on S^n, and we are done. □

Theorem 3.2 ([CGY03, Theorem 1.1])**.** *A conformal metric on* $\mathbf{R}^n$, $n \geq 3$ *with finite volume has constant scalar curvature if and only if it has the form* $g = v^{-2}g_{\mathbf{R}^n}$, $v = c(\mu^2 + |x-a|^2)$ *for some* $c, \mu > 0$, $a \in \mathbf{R}^n$.

Step 1: It suffices to show that g has constant curvature. The invocation of Killing-Hopf in Theorem 3.1 was a red herring. We recall from (3.1) that

$$B_{jk} = -(n-2)v[(v^{-1})_{jk} - n^{-1}\Delta(v^{-1})g_{jk}],$$

with covariant derivatives and index liftings performed with respect to g. In Euclidean coordinates this simplifies to

$$B_{jk} = (n-2)v^{-1}[\partial_{jk}v - n^{-1}\Delta_0 v\delta_{jk}],$$

where we write Δ_0 for the Euclidean Laplacian.

Suppose $B \equiv 0$. Then $\partial_{jk}v = 0$ for $j \neq k$, so $\partial_k v = f_k(x^k)$ for each k, and so

$$v = F_1(x^1) + \cdots + F_n(x^n).$$

Now $F_1''(x^1) = \cdots = F_n''(x^n)$, so there exists some constant c such that $F_i''(x^i) \equiv 2c$ for all i, and hence we have

$$v = c|x|^2 + r_i x^i + s.$$

Since $v > 0$ and $\mathrm{Vol}(g) = \int_{\mathbf{R}^n} v^{-n} < \infty$, we conclude that $v = c(\mu^2 + |x-a|^2)$ for some $c, \mu > 0$ and $a \in \mathbf{R}^n$. □

Step 2: Reduction to a "tail estimate". Let the test functions η_ρ be as in Section 2. To prove that $B \equiv 0$, it suffices to show

$$\int_{\mathbf{R}^n} v^{-1}B_{jk}B^{jk}\eta_\rho^2\mathrm{d}V_g = 0$$

for any $\rho > 0$ however large. Successively using (3.1), the tracelessness of B, and (3.2), we have

$$\begin{aligned}\int_{\mathbf{R}^n} B_{jk}B^{jk}v^{-1}\eta_\rho^2 \mathrm{d}V_g &= -(n-2)\int_{\mathbf{R}^n} B^{jk}[(v^{-1})_{jk} - n^{-1}\Delta(v^{-1})g_{jk}]\eta_\rho^2\mathrm{d}V_g \\ &= -(n-2)\int_{\mathbf{R}^n} B^{jk}(v^{-1})_{jk}\eta_\rho^2\mathrm{d}V_g \\ &= (n-2)\int_{\mathbf{R}^n} B^{jk}(v^{-1})_j(\eta_\rho^2)_k\mathrm{d}V_g.\end{aligned}$$

Using Cauchy's inequality we thus have

$$\begin{aligned}\int_{\mathbf{R}^n} |B|^2v^{-1}\eta_\rho^2\mathrm{d}V_g &\leq (n-2)\int_{\mathbf{R}^n} |B||\nabla(v^{-1})||\nabla(\eta_\rho^2)|\mathrm{d}V_g \\ &= 2(n-2)\int_{\mathbf{R}^n} |B||\nabla v||\nabla\eta_\rho|v^{-2}\eta_\rho\mathrm{d}V_g \\ &\leq 2(n-2)\left(\int_{B_{2\rho}\setminus B_\rho} |B|^2v^{-1}\eta_\rho^2\mathrm{d}V_g\right)^{\frac{1}{2}}\left(\int_{\mathbf{R}^n}|\nabla v|^2|\nabla\eta_\rho|^2v^{-3}\mathrm{d}V_g\right)^{\frac{1}{2}}.\end{aligned}$$

We claim that it suffices to show that $\int_{\mathbf{R}^n}|\nabla v|^2|\nabla\eta_\rho|^2v^{-3}\mathrm{d}V_g$ is bounded above as $\rho \to \infty$. If we assume that this condition is obtained, then

$$\int_{\mathbf{R}^n} |B|^2v^{-1}\eta_\rho^2\mathrm{d}V_g \leq 4(n-2)^2\int_{\mathbf{R}^n}|\nabla v|^2|\nabla\eta_\rho|^2v^{-3}\mathrm{d}V_g$$

is also bounded, so

$$\int_{\mathbf{R}^n} |B|^2v^{-1}\mathrm{d}V_g < \infty.$$

This means

$$\int_{B_{2\rho}\setminus B_\rho} |B|^2v^{-1}\eta_\rho^2\mathrm{d}V_g \to 0 \text{ as } \rho \to \infty,$$

and hence

$$\int_{\mathbf{R}^n} |B|^2v^{-1}\eta_\rho^2\mathrm{d}V_g \to 0 \text{ as } \rho \to \infty,$$

and this finishes the proof. □

Changing to the Euclidean metric, and writing ∇_0 for the Euclidean gradient, we have

$$\begin{aligned}\int_{\mathbf{R}^n}|\nabla v|^2|\nabla\eta_\rho|^2v^{-3}\mathrm{d}V_g &= \int_{\mathbf{R}^n}|\nabla_0 v|^2|\nabla_0\eta_\rho|^2v^{-n+1}\mathrm{d}x \\ &\leq C\rho^{-2}\int_{B_{2\rho}\setminus B_\rho}|\nabla_0 v|^2v^{-n+1}\mathrm{d}x,\end{aligned}$$

so it suffices to show $\int_{B_{2\rho}\setminus B_\rho}|\nabla_0 v|^2v^{-n+1}\mathrm{d}x = O(\rho^2)$, and this (technically equivalent) "tail estimate" is the one we shall refer to subsequently.

We also require two lemmas before coming back to the proof of Theorem 3.2.

Lemma 3.3. *If a conformal metric on* $\mathbf{R}^n$*,* $n \geq 3$ *has scalar curvature* $S \leq 0$*, then it does not have finite volume.*

Proof. Let $g = \varphi^{p-2} g_{\mathbf{R}^n}$. We recall that $S = -a^{-1}\varphi^{1-p}\Delta_0\varphi$, so $\Delta_0\varphi \geq 0$, φ is subharmonic. By the mean value property we have

$$\int_{B_R} \varphi \mathrm{d}x \geq |B_R| f(0)$$

for all $R > 0$. Hence, by Hölder's inequality,

$$\|\varphi\|_{L^p(B_R)} \geq \frac{\|\varphi\|_{L^1(B_R)}}{\|1\|_{L^q(B_R)}} \geq \frac{|B_R| f(0)}{\|1\|_{L^q(B_R)}} = \|f(0)\|_{L^p(B_R)},$$

so $\mathrm{Vol}(\varphi) = \|\varphi\|_p = \infty$. □

Lemma 3.4. *If a conformal metric on* $\mathbf{R}^n$*,* $n \geq 3$ *has finite volume and scalar curvature* $S \geq 0$*,* $\sup S < \infty$*, then the metric has the form* $g = v^{-2} g_{\mathbf{R}^n}$*,* $v(x) = O(|x|^2)$ *as* $|x| \to \infty$*.*

Proof. Let $g = \varphi^{p-2} g_{\mathbf{R}^n}$. Then $S = -a^{-1}\varphi^{1-p}\Delta_0\varphi$. Let us consider the conformal automorphism defined on $\mathbf{R}^n \setminus \{0\}$ by $\kappa(x) = x/|x|^2$. We compute $\kappa^* g = (|x|^{2-n}\kappa^*\varphi)^{p-2} g_{\mathbf{R}^n}$, so letting

$$\hat{\varphi}(x) = |x|^{2-n}\varphi\left(\frac{x}{|x|^2}\right),$$

(this is the *Kelvin transform* of φ), we have

$$S\left(\frac{x}{|x|^2}\right) = -a^{-1}\hat{\varphi}^{1-p}\Delta_0\hat{\varphi}. \tag{3.3}$$

We are interested in the behaviour of $\hat{\varphi}$ near the origin, because $v(x) \leq C|x|^2$ for large $|x|$ if and only if $\hat{\varphi}(x) \geq C^{(2-n)/2}$ for small $|x|$. We claim that $\hat{\varphi}$ can be extended to a superharmonic function on $\mathbf{R}^n$. This would imply the desired result: assuming for contradiction that there is a sequence of points $x_n \to 0$ such that $\hat{\varphi}(x_n) \to 0$, then

$$\int_{B_1} \hat{\varphi}(x)\mathrm{d}x = \lim_{n\to\infty} \int_{B_{1-|x_n|}(x_n)} \hat{\varphi}(x)\mathrm{d}x \leq \lim_{n\to\infty} |B_{1-|x_n|}|\hat{\varphi}(x_n) = 0,$$

which contradicts the fact that $\hat{\varphi} > 0$ on $\mathbf{R}^n \setminus \{0\}$.

By (3.3), we have

$$(\Delta_0 + a\hat{\varphi}^{p-2} S(x/|x|^2))\hat{\varphi} = 0 \text{ on } \mathbf{R}^n \setminus \{0\}.$$

By the finite volume assumption, $\hat{\varphi} \in L^p(\mathbf{R}^n)$, $p > n/(n-2)$, and because we assume S to be bounded, we have $a\hat{\varphi}^{p-2}S(x/|x|^2) \in L^{n/2}(\mathbf{R}^n)$, so the conditions of Theorem A.1 are satisfied, hence the same equation holds weakly over all of $\mathbf{R}^n$. In particular, $\Delta_0\hat{\varphi} \leq 0$ weakly over $\mathbf{R}^n$, which implies that $\hat{\varphi}$ is equal a.e. to a superharmonic function $\bar{\hat{\varphi}}$. Since $\hat{\varphi}$ is superharmonic on $\mathbf{R}^n \setminus \{0\}$, we have $\hat{\varphi} = \bar{\hat{\varphi}}$ on $\mathbf{R}^n \setminus \{0\}$, and this finishes the proof. □

Step 3: Reduction of the "tail estimate" to a growth estimate. We recall that the desired estimate is

$$\int_{B_{2\rho}\setminus B_\rho} |\nabla_0 v|^2 v^{-n+1} \mathrm{d}x = O(\rho^2) \text{ as } \rho \to \infty.$$

Let $\tilde{\eta}_\rho$ be a nonnegative smooth function identically 1 on $B_{2\rho} \setminus B_\rho$, supported on $B_{5\rho/2} \setminus B_{\rho/2}$, and such that $|\nabla\tilde{\eta}_\rho| \leq C\rho^{-1}$. Moreover, we assume ρ to be sufficiently large for Lemma 3.4 to apply. Using the formula

$$S = 2(n-1)v\Delta_0 v - n(n-1)|\nabla_0 v|^2,$$

multiplying both sides by $v^{-n+1+\alpha}\tilde{\eta}_\rho^4$ and integrating by parts we get

$$\begin{aligned}
\int_{\mathbf{R}^n} S v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x &= 2(n-1)\int_{\mathbf{R}^n} v^{-n+2+\alpha}\Delta_0 v \tilde{\eta}_\rho^4 \mathrm{d}x - n(n-1)\int_{\mathbf{R}^n} |\nabla_0 v|^2 v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x \\
&= -2(n-1)\int_{\mathbf{R}^n} \nabla_0 v \nabla_0(v^{-n+2+\alpha}\tilde{\eta}_\rho^4)\mathrm{d}x - n(n-1)\int_{\mathbf{R}^n} |\nabla_0 v|^2 v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x \\
&= (n-4-\alpha)(n-1)\int_{\mathbf{R}^n} |\nabla_0 v|^2 v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x - 2(n-1)\int_{\mathbf{R}^n} \nabla_0 v \nabla_0(\tilde{\eta}_\rho^4) v^{-n+2+\alpha}\mathrm{d}x.
\end{aligned}$$

Choosing $\alpha \neq n-4$, we have

$$\int_{\mathbf{R}^n} |\nabla_0 v|^2 v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x = C_1\int_{\mathbf{R}^n} S v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x + C_2\int_{\mathbf{R}^n} \nabla_0 v \nabla_0(\tilde{\eta}_\rho^4) v^{-n+2+\alpha}\mathrm{d}x. \tag{3.4}$$

Let us assume that $0 \leq 1+\alpha \leq 3+\alpha \leq n$. Then by Hölder's inequality,

$$\int_{\mathbf{R}^n} v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x \leq \left\|v^{-n+1+\alpha}\right\|_{n/(n-1-\alpha)}\left\|\tilde{\eta}_\rho^4\right\|_{n/(1+\alpha)} \leq C\rho^{1+\alpha}.$$

Similarly, with $3+\alpha$ in the place of $1+\alpha$, we have

$$\begin{aligned}
\left|\int_{\mathbf{R}^n} \nabla_0 v \nabla_0(\tilde{\eta}_\rho^4) v^{-n+2+\alpha}\mathrm{d}x\right| &\leq 4\rho^{-1}\int_{\mathbf{R}^n} |\nabla_0 v| v^{-n+2+\alpha}\tilde{\eta}_\rho^3 \mathrm{d}x \\
&\leq C\rho^{-1}\left(\int_{\mathbf{R}^n} |\nabla_0 v|^2 v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x\right)^{1/2}\left(\int_{\mathbf{R}^n} v^{-n+3+\alpha}\tilde{\eta}_\rho^2 \mathrm{d}x\right)^{1/2} \\
&\leq C\rho^{(1+\alpha)/2}\left(\int_{\mathbf{R}^n} |\nabla_0 v|^2 v^{-n+1+\alpha}\tilde{\eta}_\rho^4 \mathrm{d}x\right)^{1/2}.
\end{aligned}$$

The left hand side of (3.4) (let's call it F) satisfies $F \leq C\rho^{1+\alpha} + C(\rho^{1+\alpha}F)^{1/2}$, and so $F \leq C\rho^{1+\alpha}$.

For $n = 3$, we simply choose $\alpha = 0$ to get

$$\int_{B_{2\rho}\setminus B_\rho} |\nabla_0 v|^2 v^{-n+1} \mathrm{d}x \leq F \leq C\rho \leq C\rho^2.$$

For $n \geq 4$ we choose $\alpha = -1$. Then $F \leq C$, and we apply Lemma 3.4 to get

$$\int_{B_{2\rho}\setminus B_\rho} |\nabla_0 v|^2 v^{-n+1} \mathrm{d}x \leq C\rho^2 \int_{B_{2\rho}\setminus B_\rho} |\nabla_0 v|^2 v^{-n} \mathrm{d}x \leq C\rho^2 F \leq C\rho^2.$$

This finishes the proof. □

3.2 Symmetry Tricks

Theorem 3.5 ([FL12, Section 3]). *Suppose $\varphi \geq 0$, $\varphi \in W^{1,2}(S^n) \setminus \{0\}$ minimizes $Q_{g_{S^n}}$ and satisfies the "center of mass condition"*

$$\int_{S^n} \omega^j \varphi^p \mathrm{d}V_{g_{S^n}} = 0,\ j = 1, \ldots, n+1, \tag{3.5}$$

where $(\omega^1, \ldots, \omega^{n+1}) = x \in S^n \subset \mathbf{R}^{n+1}$. Then φ is a constant.

Proof. Since the first variation of $Q_{g_{S^n}}$ is zero at φ, the second variation is given by

$$0 \leq \frac{\mathrm{d}^2}{\mathrm{d}t^2} Q_{g_{S^n}}(\varphi + tw) = \frac{\frac{\mathrm{d}^2}{\mathrm{d}t^2} E(\varphi + tw)\|\varphi\|_p^2 - E(\varphi)\frac{\mathrm{d}^2}{\mathrm{d}t^2}\|\varphi + tw\|_p^2}{\|\varphi\|_p^4}. \tag{3.6}$$

for $w \in W^{1,2}(S^n)$. Now

$$\|\varphi + tw\|_p^2 = \left(\|\varphi\|_p^p + t\int_{S^n} p\varphi^{p-1} w \mathrm{d}V_{g_{S^n}} + \frac{t^2}{2}\int_{S^n} p(p-1)\varphi^{p-2} w^2 \mathrm{d}V_{g_{S^n}} + o(t^2)\right)^{2/p},$$

so if w satisfies $\int_{S^n} \varphi^{p-1} w \mathrm{d}V_{g_{S^n}} = 0$, we have

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\|\varphi + tw\|_p^2 = \frac{2}{p}\|\varphi\|_p^{2-p}\int_{S^n} p(p-1)\varphi^{p-2} w^2 \mathrm{d}V_{g_{S^n}},$$

and (3.6) becomes (after multiplying both sides by $\frac{1}{2}\|\varphi\|_p^2$)

$$0 \leq E(w) - (p-1)E(\varphi)\|\varphi\|^{-p}\int_{S^n} \varphi^{p-2} w^2 \mathrm{d}V_{g_{S^n}}. \tag{3.7}$$

Because φ satisfies the center-of-mass condition (3.5), we may choose $w = \omega^j\varphi$ in (3.7); summing over $j = 1, \ldots, n+1$ we get

$$0 \le \sum_{j=1}^{n+1} E(\omega_j \varphi) - (p-1)E(\varphi).$$

An elementary calculation ([FL12, Lemma 3.2]) gives

$$\sum_{j=1}^{n+1} E(\omega_j \varphi) = E(\varphi) + n \int_{S^n} \varphi^2 \mathrm{d}V_{g_{S^n}},$$

so we end up with

$$0 \le n \int_{S^n} \varphi^2 \mathrm{d}V_{g_{S^n}} - (p-2)E(\varphi) = -(p-2) \int_{S^n} |\nabla \varphi|^2 \mathrm{d}V_{g_{S^n}}.$$

So φ must be a constant. $\square$

It remains only to fulfill the center-of-mass condition (3.5). For $\xi \in S^n$, let $\pi_\xi : S^n \to \mathbf{R}^n$ be the stereographic projection that identifies ξ with $0 \in \mathbf{R}^n$, and for $\delta > 0$, let $\gamma_{\delta,\xi}(\omega) = \pi_\xi^{-1} \delta \pi_\xi(\omega)$ for $\omega \in S^n$. Since both $x \mapsto \delta x$ and π_ξ are conformal, so $\gamma = \gamma_{\delta,\xi}$ is conformal.

Suppose $(\gamma^{-1})^* g_{S^n} = \psi^{p-2} g_{S^n}$. For $\varphi \in W^{1,2}(S^n) \setminus \{0\}$, let $\tilde{\varphi} = \psi (\gamma^{-1})^* \varphi$. Then as in Lemma 1.5, we have $E(\varphi) = E(\tilde{\varphi})$, $\|\varphi\|_p^2 = \|\tilde{\varphi}\|_p^2$. In particular, if φ is a minimizer of $Q_{g_{S^n}}$, then so is $\tilde{\varphi}$. Moreover, with regard to the center-of-mass condition, we have

$$\begin{aligned}
\int_{S^n} \tilde{\omega} \tilde{\varphi}(\tilde{\omega})^p \mathrm{d}V_{g_{S^n}}(\tilde{\omega}) &= \int_{S^n} \gamma(\omega) \psi(\tilde{\omega})^p \varphi(\omega)^p \mathrm{d}V_{g_{S^n}}(\tilde{\omega}) \\
&= \int_{S^n} \gamma(\omega) \varphi(\omega)^p \mathrm{d}V_{(\gamma^{-1})^* g_{S^n}}(\tilde{\omega}) \\
&= \int_{S^n} \gamma(\omega) \varphi(\omega)^p \mathrm{d}V_{g_{S^n}}(\omega).
\end{aligned}$$

The following result can be deduced from Brouwer's fixed point theorem.

Lemma 3.6 ([FL12, Lemma B.1])**.** *Suppose $f \in L^1(S^n)$ satisfies $\int_{S^n} f \mathrm{d}V_{g_{S^n}} \neq 0$. Then there exists $0 < \delta \le 1, \xi \in S^n$ such that the $\mathbf{R}^{n+1}$-valued integral*

$$\int_{S^n} \gamma_{\delta,\xi}(\omega) f(\omega) \mathrm{d}V_{g_{S^n}}(\omega) = 0.$$

Theorem 3.5 combined with Lemma 3.6 imply that if $\varphi \ge 0$, $\varphi \in W^{1,2}(S^n) \setminus \{0\}$ is a minimizer of $Q_{g_{S^n}}$ (equivalently, $\overline{\varphi} \ge 0$, $\overline{\varphi} \in D^{1,2}(\mathbf{R}^n) \setminus \{0\}$ is a minimizer of $\|\nabla \overline{\varphi}\|_2^2 / \|\overline{\varphi}\|_p^2$), then there exist $0 < \delta \le 1$ and $\xi \in S^n$ such that $\tilde{\varphi} = \psi_{\delta,\xi}(\gamma_{\delta,\xi}^{-1})^* \varphi$ is a constant. Hence all nonnegative minimizers are isometric to (S^n, cg_{S^n}), and in fact have the form

$$\varphi = c\gamma^*_{\delta,\xi}(1/\psi_{\delta,\xi}) = c\left(\frac{2\delta}{(\delta^2+1)+(\delta^2-1)\xi\cdot\omega}\right)^{(n-2)/2}, c > 0. \tag{3.8}$$

We claim that this implies Theorem 1.4, provided there exists some minimizer $\varphi \in W^{1,2}(S^n)\setminus\{0\}$. Since $\nabla\varphi_\pm = \pm 1_{\pm\varphi>0}\nabla\varphi$, $E(\varphi) = E(\varphi_+) + E(\varphi_-)$; and since $\|\varphi\|_p^2 \le \|\varphi_+\|_p^2 + \|\varphi_-\|_p^2$, we have $E(\varphi_\pm) = \Lambda\|\varphi_\pm\|_p^2$ where $\Lambda = E(\varphi)/\|\varphi\|_p^2$. Applying (3.8) to $\varphi_\pm$, we see that $\varphi = \pm$(3.8).

4 Convergence of Subcritical Solutions

This section follows [SY94]. We recall that

$$\lambda(M) = \inf\{Q_g(\varphi) : \varphi > 0, \varphi \in C^\infty(M)\}.$$

This turns out to be equivalent to

$$\lambda(M) = \inf\{Q_g(\varphi) : \varphi \in W^{1,2}(M) \setminus \{0\}\}.$$

To see the equivalence, we note that Q_g is well-defined and continuous on $W^{1,2}(M) \setminus \{0\}$. For all $\varphi \in W^{1,2}(M)$, at least one of $\varphi_\pm$ (wlog, let's say φ_+) is nonzero with $Q_g(\varphi_+) \le Q_g(\varphi)$. And φ_+ can be approximated in $W^{1,2}(M)$ by smooth, positive functions: we can split it up using a partition of unity and use convolution to mollify the pieces on local charts.

4.1 A Proof of Theorem 1.1

Theorem 1.1. *If $\lambda(M) < \lambda(S^n)$, then $\lambda(M) = Q_g(\varphi_0)$ for some $\varphi_0 > 0$, $\varphi_0 \in C^\infty(M)$.*

For $2 < s < p$, the perturbed functional $Q_{g,s}$ is defined by

$$Q_{g,s}(\varphi) = \frac{E(\varphi)}{\|\varphi\|_s^2}. \tag{4.1}$$

The properties of Q_g discussed in Section 1.1 are all shared by $Q_{g,s}$. By Hölder's inequality, $Q_{g,s}$ is bounded below, so we may define

$$\lambda_s = \inf\{Q_{g,s}(\varphi) : \varphi \in W^{1,2}(M) \setminus \{0\}\}.$$

And a function $\varphi \in W^{1,2}(M) \setminus \{0\}$ is a critical point if and only if it is a (weak) solution to the subcritical equation

$$\Box\varphi = \lambda\varphi^{s-1}. \tag{4.2}$$

The point of considering $Q_{g,s}$ where $s < p$ is that the embedding $W^{1,2}(M) \hookrightarrow L^s(M)$ is compact;

whereas p is exactly the critical exponent at which compactness does not hold. The proof below demonstrates why this matters.

Proposition 4.1. *For $2 < s < p$, there exists $u \geq 0$, $u \in W^{1,2}(M)$, such that $Q_{g,s}(u) = \lambda_s$.*

Proof. Let $u_i \geq 0$, $u_i \in W^{1,2}(M)$ be a minimizing sequence for $Q_{g,s}$, normalized so that $\|u_i\|_s = 1$. Then u_i is bounded in $W^{1,2}(M)$:

$$\begin{aligned} \|u_i\|_2 &\leq C\|u_i\|_s, \text{ and} \\ \|\nabla u_i\|_2^2 &= E(u_i) - \int_M aSu_i^2 \mathrm{d}V_g \leq E(u_i) + C\|u_i\|_s^2. \end{aligned}$$

Hence, we may pass to a subsequence that converges weakly in $W^{1,2}(M)$ *and strongly in $L^s(M)$*. Let u be the limit. Then

$$\begin{aligned} \|u\|_s &= \lim_{i\to\infty} \|u_i\|_s = 1, \\ \int_M aSu^2 \mathrm{d}V_g &= \lim_{i\to\infty} \int_M aSu_i^2 \mathrm{d}V_g, \text{ and} \\ \int_M |\nabla u|^2 \mathrm{d}V_g &\leq \lim_{i\to\infty} |\nabla u_i|^2 \mathrm{d}V_g \text{ since } \nabla u \text{ is the weak limit,} \end{aligned}$$

so taken together we have

$$Q_{g,s}(u) \leq \lim_{i\to\infty} Q_{g,s}(u_i) = \lambda_s.$$

Since λ_s is the infimum of $Q_{g,s}$ we must have $Q_{g,s}(u) = \lambda_s$. □

Using an elliptic regularity “bootstrapping" argument, it can be shown that the minimizer given by Proposition 4.1 is a smooth and strictly positive solution to the subcritical equation (4.2).

Proposition 4.2. *For $2 < s < p$, there exists a function $\varphi_s > 0$, $\varphi_s \in C^\infty(M)$, $\|\varphi_s\|_s = 1$ satisfying $\Box\varphi_s = \lambda_s\varphi_s^{s-1}$.*

Proof. Let φ_s be the weak solution u obtained in Proposition 4.1. Since $\varphi_s \in L^p$, we have $\lambda_s\varphi_s^{s-1} \in L^{p/(s-1)}$, so by L^p estimates $\varphi_s \in W^{2,p/(s-1)}$, and by Sobolev embedding $\varphi_s \in L^{q_1}$ for all $1/q_1 \geq (s-1)/p - 2/n$. By the same token, if $\varphi_s \in L^q$, then $\varphi_s \in L^{q'}$ for all $1/q' \geq (s-1)/q - 2/n$.

The map $f(x) = (s-1)x - 2/n$ iterates to $f^{(n)}(x) = (s-1)^n(x - 2/(ns-2n)) + 2/(ns-2n)$, so since $s - 1 > 1$ and $1/p - 2/(ns-2n) < 0$, we have $f^{(n)}(1/p) < 0$ for sufficiently large n, and thus $\varphi_s \in L^q$ for all $q < \infty$. Applying L^p estimates one more time we have $\varphi_s \in W^{2,q}$ for all $q < \infty$, and using Sobolev embedding with $q > n$, we have $\varphi_s \in C^{1,\alpha}$ for some $\alpha > 0$. Then $\lambda_s\varphi_s^{s-1}$ is Hölder continuous, so by Schauder estimates $\varphi_s \in C^{2,\alpha}$.

Let $c \geq 0$ be so large that $-\Delta\varphi_s + c\varphi_s \geq 0$. Then by the maximum principle, if $\varphi_s(x) = 0$ at some point x, then φ_s must be constantly zero, which cannot be true; hence, $\varphi_s > 0$.

Thus for all k, if $\varphi_s \in C^{k,\alpha}$, then $\lambda_s \varphi_s^{s-1} \in C^{k,\alpha}$, and so by Schauder estimates $\varphi_s \in C^{k+2,\alpha}$. Applying Schauder estimates iteratively we get $\varphi_s \in C^\infty$. □

It turns out that φ_s converge to a solution for the critical equation (1.2), *provided that* $\lambda(M) < \lambda(S^n)$. Theorem 1.1 follows immediately.

Proposition 4.3. *Suppose $\lambda(M) < \lambda(S^n)$. Let φ_s be as in Proposition 4.2. Then φ_s converge in C^0 along a subsequence to a function φ satisfying $\varphi > 0$, $\varphi \in C^\infty(M)$, with $Q_g(\varphi) = \lambda(M)$.*

First, we need an elementary lemma.

Lemma 4.4. $\limsup_{s\to p} \lambda_s \leq \lambda(M)$*, and if* $\lambda(M) \geq 0$*, then we have* $\lim_{s\to p} \lambda_s = \lambda(M)$.

Proof. For all $u \in W^{1,2}(M) \setminus \{0\}$ we have $\limsup_{s\to p} \lambda_s \leq \lim_{s\to p} Q_{g,s}(u) = Q_g(u)$, so letting $Q_g(u) \to \lambda(M)$ we get the desired bound. Now suppose $\lambda(M) \geq 0$. Then for all $u \in W^{1,2}(M) \setminus \{0\}$, since $E(u) \geq 0$, we have $Q_{g,s}(u) = E(u)/\|u\|_s \geq \|1\|_{ps/(p-s)} Q_g(u)$. Letting $Q_{g,s}(u) \to \lambda(M)$, we have $\lambda_s \geq \|1\|_{ps/(p-s)}\lambda(M)$, so $\limsup_{s\to p} \lambda_s \geq \lambda(M)$, and combined with the previous bound we get $\lim_{s\to p} \lambda_s = \lambda(M)$. □

(In the case $\lambda(M) < 0$, choosing $E(u) < 0$, we have that $Q_{g,s}(u)/\|1\|_s$ is nondecreasing in s, so $\lim_{s\to p} \lambda_s$ exists and is $\leq \lambda(M)$. But we shall soon see that it always equals $\lambda(M)$.)

Back to Proposition 4.3. It suffices to show that φ_s are bounded in C^0 (at least for all s sufficiently close to p), for then by L^p and Schauder estimates they are bounded in $C^{k,\alpha}$ for all k, and by the Arzela-Ascoli theorem a subsequence φ_{s_i} converges in C^k for all k. We may also assume that λ_{s_i} converge to $\limsup_{s\to p} \lambda_s$. Then $\varphi = \lim \varphi_{s_i} \in C^\infty$ satisfies $\square\varphi = (\limsup_{s\to p} \lambda_s)\varphi^{p-1}$, so $Q_g(\varphi) = \limsup_{s\to p} \lambda_s \leq \lambda(M)$. Since $\lambda(M)$ is the infimum of Q_g, equality is achieved, and φ is a smooth nonnegative minimizer of Q_g. As in the proof of Proposition 4.2, by the maximum principle, we have $\varphi > 0$.

The proof of Proposition 4.3 presented in [LP87] derives the C^0 bound by first showing that for $r > p$ sufficiently close to p, we have $\sup_{s_0 \leq s < p} \|\varphi_s\|_r < \infty$ for some $s_0 < p$. Then the same iterative argument in Proposition 4.2, with s replaced by p and $1/p$ replaced by $1/r$, can be applied (further increasing s_0 if necessary) to give a $C^{1,\alpha}$ bound.

Another proof, given in [SY94], directly addresses the question of uniform boundedness.

Proof of Proposition 4.3. Let us suppose for contradiction that φ_s are not bounded in C^0 as $s \to p$. Then there exists a sequence $s_i \to p$, $z_i \in M$, such that $m_i = \sup \varphi_{s_i} = \varphi_{s_i}(z_i) \to \infty$. Since M is compact, we may assume $z_i \to z_0 \in M$.

We shall work in a normal neighbourhood around z_0, having radius ρ. In Euclidean coordinates, φ_{s_i} satisfies the equation

$$-\frac{1}{\sqrt{G}}\partial_j(\sqrt{G}g^{jk}\partial_k\varphi_{s_i}) + aS\varphi_{s_i} - \lambda_{s_i}\varphi_{s_i}^{s_i-1} = 0.$$

Let x_i be the coordinates of z_i, and let $u_i(x) = m_i^{-1}\varphi_{s_i}(m_i^{1-s_i/2}x+x_i)$ for $|x| < m_i^{-1+s_i/2}(\rho-|x_i|) \to \infty$. Then $0 \le u_i \le 1$ satisfies

$$\begin{aligned} 0 &= m^{1-s_i}\left[-\frac{m_i^{s_i-1}}{\sqrt{\tilde{G}}}\partial_j(\sqrt{\tilde{G}}\tilde{g}^{jk}\partial_k u_i) + a\tilde{S}m_i u_i - \lambda_{s_i}m_i^{s_i-1}u_i^{s_i-1}\right] \\ &= -\frac{1}{\sqrt{\tilde{G}}}\partial_j(\sqrt{\tilde{G}}\tilde{g}^{jk}\partial_k u_i) + a\tilde{S}m_i^{2-s_i}u_i - \lambda_{s_i}u_i^{s_i-1}, \end{aligned}$$

where tilde signifies pullback by $x \mapsto m_i^{1-s_i/2}x + x_i$. Expanding the equation into nondivergence form $-a^{jk}\partial_{jk}u_i + b^k\partial_k u_i + cu_i = f$, we have

$$\begin{cases} a^{jk} = \tilde{g}^{jk} & \to \delta^{jk} \text{ in } C^1 \text{ on all bounded domains in } \mathbf{R}^n \\ b^k = -\partial_j\tilde{g}^{jk} - (\partial_j \log \tilde{G})\tilde{g}^{jk} & \to 0 \text{ in } C^0 \text{ on all bounded domains in } \mathbf{R}^n \\ c = a\tilde{S}m_i^{2-s_i} & \to 0 \text{ in } C^0 \text{ on all bounded domains in } \mathbf{R}^n, \end{cases}$$

so since u_i and $f = \lambda_{s_i}u_i^{s_i-1}$ are uniformly bounded, by L^p and Schauder estimates, $\|u_i\|_{C^{2,\alpha}(B_R)}$ is bounded for all $R > 0$. By Arzela-Ascoli, we may assume that u_i is C^2-convergent on all B_R. Then the limit u satisfies (with Δ denoting the Euclidean Laplacian)

$$-\Delta u = \lambda(M)u^{p-1},$$

with $0 \le u \le 1$, $u(0) = 1$. By the maximum principle, $u > 0$.

To derive a contradiction, it suffices to show that $\|u\|_p \le 1$ and $\|\nabla u\|_2 < \infty$. Then by the dominated convergence theorem we have

$$\int_{\mathbf{R}^n} |\nabla u||\nabla(u - \eta_\rho u)|\mathrm{d}x \to 0 \text{ as } \rho \to \infty,$$

and so letting $\rho \to \infty$ on both sides of

$$\int_{\mathbf{R}^n} \nabla u \nabla(\eta_\rho u)\mathrm{d}x = \int_{\mathbf{R}^n} (-\Delta u)(\eta_\rho u)\mathrm{d}x = \lambda(M)\int_{\mathbf{R}^n} u^p\eta_\rho \mathrm{d}x,$$

we get

$$\int_{\mathbf{R}^n} |\nabla u|^2\mathrm{d}x = \lambda(M)\int_{\mathbf{R}^n} u^p\mathrm{d}x \le \lambda(M)\left(\int_{\mathbf{R}^n} u^p\mathrm{d}x\right)^{2/p} < \Lambda\left(\int_{\mathbf{R}^n} u^p\mathrm{d}x\right)^{2/p},$$

which contradicts Theorem 1.4.

To show $\|u\|_p \leq 1$, we apply Fatou's lemma to the estimate

$$\int_{|x|<m_i^{-1+s_i/2}(\rho-|x_i|)} u_i^{s_i}\sqrt{\tilde{G}}\mathrm{d}x = m_i^{(n/2-1)s_i-n}\int_{B_{\rho-|x_i|}(x_i)} \varphi_{s_i}^{s_i}\sqrt{G}\mathrm{d}x < \|\varphi_{s_i}\|_{s_i}^{s_i} = 1.$$

The exact same argument also yields $\|\nabla u\|_2 < \infty$. □

4.2 Finding an Optimizer for the Sobolev Inequality

For any bounded Lipschitz domain Ω in $\mathbf{R}^n$, let $\lambda(\Omega) = \inf\{\|\nabla u\|_2^2/\|u\|_p^2 : u \in D^{1,2}(\Omega)\}$. Since the quotient $\|\nabla u\|_2^2/\|u\|_p^2$ is invariant under translation and scaling, so translation and scaling of the domain Ω leaves $\lambda(\Omega)$ is invariant. On the other hand, $\lambda(n\Omega) \downarrow \lambda(\mathbf{R}^n)$ as $n \to \infty$, since, assuming wlog that $0 \in \Omega$, the domains $n\Omega$ cover $\mathbf{R}^n$. Hence, $\lambda(\Omega) = \lambda(\mathbf{R}^n)$ for any Ω. We write Λ for their common value.

Let $\lambda_s(\Omega) = \inf\{\|\nabla u\|_2^2/\|u\|_s^2 : u \in D^{1,2}(\Omega)\}$ for $2 < s < p$. Under the assumption that Ω is bounded, exactly as for Lemma 4.4 we have $\lim_{s\to p}\lambda_s(\Omega) = \lambda(\Omega) = \Lambda$. With a view to obtaining an optimizer for the $\mathbf{R}^n$ Sobolev inequality, we shall consider the domain $B = B_1$.

Exactly as in the proofs of Proposition 4.1 and Proposition 4.2, using minimizing sequences and a bootstrapping argument we can show that for any $2 < s < p$ there exists $\varphi_s > 0$, $\varphi_s \in C^\infty(B)$, $\|\varphi_s\|_s = 1$, satisfying

$$\begin{cases} -\Delta\varphi_s = \lambda_s(B)\varphi_s^{s-1} & \text{in } B \\ \varphi_s = 0 & \text{on } \partial B. \end{cases}$$

But unlike the proof of Proposition 4.3, we *don't* want φ_s to be C^0-bounded. Indeed, let us suppose that there exists a sequence $s_i \to p$ such that $m_i = \sup\varphi_{s_i} = \varphi_{s_i}(x_i) \to \infty$. By a result of Gidas-Ni-Nirenberg [GNN79, Theorem 1], or alternatively by a rearrangement inequality, all φ_s are radial and monotonically decreasing from the origin; in particular, $x_i = 0$. Setting $u_i(x) = m_i^{-1}\varphi_{s_i}(m_i^{1-s_i/2}x + x_i)$, we have (as in the proof of Proposition 4.3) that u_i is C^2-convergent on all B_R, up to subsequences. The limit $u \in C^2(\mathbf{R}^n)$ satisfies

$$-\Delta u = \Lambda u^{p-1},$$

and is nonnegative, radial, and monotonically decreasing from the origin. (At this point one may as well look at the equation as a one-dimensional problem. But without assuming a radial symmetry result, a priori the domains of u_i need not tend to $\mathbf{R}^n$, as the sequence x_i could slink toward the boundary.)

By Fatou's lemma, $\|u\|_p \leq 1$, and on the other hand

$$\int_{\mathbf{R}^n} u^p \mathrm{d}x = \int_{\mathbf{R}^n} |\nabla u|^2 \mathrm{d}x \geq \left(\int_{\mathbf{R}^n} u^p \mathrm{d}x\right)^{2/p},$$

so $\|u\|_p = 1$, and hence u is an optimizer of the Sobolev inequality on $\mathbf{R}^n$.

We shall now prove the non-boundedness result. Suppose for contradiction that φ_s are C^0-bounded for $s_0 \leq s < p$. By regularity theory, φ_s are bounded in $C^{2,\alpha}(B)$ for some α, so there is a subsequence φ_{s_i} that converges in $C^2(B)$; and the limit u satisfies

$$\begin{cases} -\Delta u = \Lambda u^{p-1} & \text{in } B \\ u = 0 & \text{on } \partial B. \end{cases}$$

However, this contradicts the fact that $\Lambda(\Omega)$ is never achieved for a bounded domain Ω. Indeed, extending u by zero to B_2, we see that u achieves the infimum $\Lambda(B_2) = \Lambda$. Being a minimizer of the Yamabe quotient on B_2, u satisfies $-\Delta u = \Lambda u^{p-1}$ in B_2, and by regularity theory $u \in C^2(B_2)$. Since u is zero on $B_2 \setminus B_1$, by the maximum principle we have a contradiction.

5 Direct Methods

5.1 The Brezis-Lieb Lemma

In 1983, Brezis and Lieb [BL83] discovered an elementary integration-theory result with applications in the calculus of variations. Following [BL83, Section 4, Example (B)] and [Heb14, Theorem 2.9], we use the Brezis-Lieb lemma to give a direct proof of Theorem 1.1. (Proofs of the Brezis-Lieb lemma itself are provided in Appendix B.)

Theorem B.1 (Brezis-Lieb). *Suppose f_n is a bounded sequence in L^p, $1 < p < \infty$, such that $f_n \to f$ a.e. Then $f_n \rightharpoonup f$ weakly in L^p, and*

$$\lim_{n\to\infty} \int \left||f_n|^p - |f_n - f|^p - |f|^p\right| = 0. \tag{B.1}$$

In particular,

$$\int |f_n|^p = \int |f|^p + \int |f_n - f|^p + o(1). \tag{B.2}$$

Theorem 1.1. *If $\lambda(M) < \lambda(S^n)$, then $\lambda(M) = Q_g(\varphi_0)$ for some $\varphi_0 > 0$, $\varphi_0 \in C^\infty(M)$.*

First we observe that it is enough to show that there exists a minimizer $\varphi \in W^{1,2}(M) \setminus \{0\}$ of Q_g. For then either of $\varphi_\pm$ is a minimizer, and so we may assume that $\varphi \geq 0$. Now since φ is a weak solution of $\Box\varphi = \lambda\varphi^{p-1}$, by regularity theory ([Str08, Lemma B.3]) we have $\varphi \in C^\infty(M)$, and by the maximum principle $\varphi > 0$.

We shall also need an optimal Sobolev inequality for manifolds. Stronger results exist but the version given in [LP87, Theorem 2.3] suffices for our purpose: applying the $\mathbf{R}^n$ version of Theorem 1.4 to a sufficiently fine partition of unity, it can be shown that for any $\varepsilon > 0$ there exists a constant $C_\varepsilon(M)$ such that

$$\|\varphi\|_p^2 \leq (1+\varepsilon)\Lambda^{-1}\|\nabla\varphi\|_2^2 + C_\varepsilon(M)\|\varphi\|_2^2 \text{ for all } \varphi \in W^{1,2}(M).$$

Proof of Theorem 1.1 based on the Brezis-Lieb lemma. Let $u_i \in W^{1,2}(M) \setminus \{0\}$ be a minimizing sequence for Q_g, with $\|u_i\|_p = 1$. Then as in Proposition 4.1, we have that u_i is bounded in $W^{1,2}(M)$. We assume that $u_i \rightharpoonup u$ in $W^{1,2}(M)$ for some u.

But we can't just assume $u_i \to u$ in $L^p(M)$, so we instead suppose that $u_i \to u$ in $L^2(M)$. We furthermore assume that $u_i \to u$ a.e., so as to be able to apply the Brezis-Lieb lemma.

The goal is to show that $\|u\|_p = 1$, for as in Proposition 4.1 we have

$$E(u) \leq \liminf_{i\to\infty} E(u_i) = \lambda(M),$$

which implies that $Q(u) = \lambda(M)$.

Let $\varepsilon > 0$ be small enough that

$$\frac{\Lambda}{1+\varepsilon} > \lambda(M).$$

By weak $W^{1,2}$-convergence and strong L^2-convergence, using the Sobolev inequality on manifolds we have

$$\begin{aligned}\frac{\Lambda}{1+\varepsilon}\|u_i - u\|_p^2 &\leq \|\nabla(u_i - u)\|_2^2 + C\|u_i - u\|_2^2 \\ &= \|\nabla u_i\|_2^2 - \|\nabla u\|_2^2 + o(1) \\ &= E(u_i) - E(u) + o(1) \\ &\leq \lambda(M)\Big(\|u_i\|_p^2 - \|u\|_p^2\Big) + o(1),\end{aligned}$$

where the last step is due to $E(u_i) = \lambda(M)\|u_i\|_p^2 + o(1)$, $E(u) \geq \lambda(M)\|u\|_p^2$.

On the other hand, by pointwise convergence, the Brezis-Lieb lemma gives

$$\|u_i\|_p^p = \|u\|_p^p + \|u_i - u\|_p^p + o(1),$$

and so

$$\|u_i\|_p^2 \leq \|u\|_p^2 + \|u_i - u\|_p^2 + o(1).$$

Hence,

$$\frac{\Lambda}{1+\varepsilon}\Big(\|u_i\|_p^2 - \|u\|_p^2\Big) \le \frac{\Lambda}{1+\varepsilon}\|u_i - u\|_p^2 + o(1) \le \lambda(M)\Big(\|u_i\|_p^2 - \|u\|_p^2\Big) + o(1).$$

Since $\|u_i\|_p = 1$, by Fatou's lemma $\|u\|_p \le 1$, so $\|u_i\|_p^2 - \|u\|_p^2 = 1 - \|u\|_p^2 \ge 0$, and thus by the above estimate $\|u\|_p = 1$. □

(From the relation $\|\nabla(u_i - u)\|_2^2 = \lambda(M) - E(u) + o(1)$, and the fact that $E(u) = \lambda(M)$, it follows that $u_i \to u$ strongly in $W^{1,2}(M)$.)

5.2 The Concentration-Compactness Principle

The first rearrangement-free proof of the sharp Sobolev inequality was due to [Lio85], as an example application of his "concentration-compactness principle". This is not one theorem, but "a general principle ...that *all minimizing sequences are relatively compact if and only if the sub-additivity inequalities are strict*" ([Lio84]).

Theorem 5.1 ([Lio85, Theorem I.1])**.** *Suppose $u_m \in D^{1,2}(\mathbf{R}^n)$, $\|u_m\|_p = 1$, $\|\nabla u_m\|_2 \to \Lambda^{1/2}$. Then up to translation and dilation, u_m converges strongly in $D^{1,2}(\mathbf{R}^n)$.*

The first concentration-compactness results that we shall need is the following.

Theorem C.1 (First Concentration-Compactness Lemma [Lio84, Lemma I.1])**.** *Any sequence μ_m of probability measures on $\mathbf{R}^n$ has a subsequence satisfying one of the following three conditions:*

> **Compactness.** *There exists a sequence $x_m \in \mathbf{R}^n$ such that for any $\varepsilon > 0$, there exists $R > 0$ satisfying*
>
> $$\int_{B_R(x_m)} \mathrm{d}\mu_m > 1 - \varepsilon \text{ for all } m.$$
>
> **Vanishing.** *For all $R > 0$,*
>
> $$\sup_{x\in\mathbf{R}} \int_{B_R(x)} \mathrm{d}\mu_m \to 0 \text{ as } m \to \infty.$$
>
> **Dichotomy.** *There exist $0 < \lambda < 1$, $x_m \in \mathbf{R}^n$, $R_m \to \infty$, such that*
>
> $$\begin{cases} \displaystyle\int_{B_{R_m}(x_m)} \mathrm{d}\mu_m = \lambda + o(1), \\ \displaystyle\int_{B_{2R_m}(x_m)\setminus B_{R_m}(x_m)} \mathrm{d}\mu_m = o(1), \\ \displaystyle\int_{\mathbf{R}^n\setminus B_{2R_m}(x_m)} \mathrm{d}\mu_m = 1 - \lambda + o(1). \end{cases}$$

Now we may begin the proof of Theorem 5.1.

STEP 1: OBTAINING TIGHTNESS. There exist $\tilde{x}_m \in \mathbf{R}^n$, $\tilde{R}_m > 0$ such that

$$\int_{B_{\tilde{R}_m}(x_m)} |u_m|^p \mathrm{d}x = \sup_{x \in \mathbf{R}^n} \int_{B_{\tilde{R}_m}(x)} |u_m|^p \mathrm{d}x = \frac{1}{2},$$

so the rescaled sequence $v_m(x) = \tilde{R}_m^{n/p} u_m(\tilde{R}_m x + \tilde{x}_m)$ satisfies

$$\int_{B_1(0)} |v_m|^p \mathrm{d}x = \sup_{x \in \mathbf{R}^n} \int_{B_1(x)} |v_m|^p \mathrm{d}x = \frac{1}{2}. \tag{5.1}$$

Moreover, we have $\|v_m\|_p = \|u_m\|_p = 1$, $\|v_m\|_2 = \|u_m\|_2 \to \Lambda$.

Applying the first concentration-compactness lemma (Theorem C.1) to $|v_m|^p \,\mathrm{d}x$, we see that "vanishing" is precluded by the rescaling. The proof is complete if we can show (1) "dichotomy" is also impossible, and (2) assuming "compactness", v_m is strongly convergent in $D^{1,2}(\mathbf{R}^n)$. We shall now prove (1), and deal with (2) only after introducing the *second concentration-compactness lemma.*

We recall that in the dichotomy scenario, there exist $0 < \lambda < 1$, $x_m \in \mathbf{R}^n$, $R_m \to \infty$, such that

$$\begin{cases} \displaystyle\int_{B_{R_m}(x_m)} \mathrm{d}\mu_m = \lambda + o(1), \\ \displaystyle\int_{B_{2R_m}(x_m) \setminus B_{R_m}(x_m)} \mathrm{d}\mu_m = o(1), \\ \displaystyle\int_{\mathbf{R}^n \setminus B_{2R_m}(x_m)} \mathrm{d}\mu_m = 1 - \lambda + o(1). \end{cases}$$

We begin by noticing that

$$\|\nabla v_m\|_2^2 \geq \|(\nabla v_m)\eta_{R_m}\|_2^2 + \|(\nabla v_m)(1 - \eta_{R_m})\|_2^2.$$

By the triangle inequality

$$\|(\nabla v_m)\eta_{R_m}\|_2 \geq \|\nabla(v_m \eta_{R_m})\|_2 - \|v_m \nabla \eta_{R_m}\|_2,$$

and by Hölder's inequality

$$\|v_m \nabla \eta_{R_m}\|_2 \leq C R_m^{-1} \|v_m\|_{L^2(B_{2R_m} \setminus B_{R_m})} \leq C \|v_m\|_{L^p(B_{2R_m} \setminus B_{R_m})} = o(1),$$

which means that we have

$$\|(\nabla v_m)\eta_{R_m}\|_2 \geq \|\nabla(v_m \eta_{R_m})\|_2 + o(1).$$

Since $\|(\nabla v_m)\eta_{R_m}\|_2 \le \|\nabla v_m\|_2$ is bounded, $\|\nabla(v_m\eta_{R_m})\|_2$ is also bounded, so we deduce that

$$\|(\nabla v_m)\eta_{R_m}\|_2^2 \ge \|\nabla(v_m\eta_{R_m})\|_2^2 + o(1).$$

Similarly $\|(\nabla v_m)(1-\eta_{R_m})\|_2^2 \ge \|\nabla(v_m(1-\eta_{R_m}))\|_2^2 + o(1)$, so

$$\|\nabla v_m\|_2^2 \ge \|(\nabla v_m\eta_{R_m})\|_2^2 + \|(\nabla v_m(1-\eta_{R_m}))\|_2^2 + o(1),$$

and by Sobolev's inequality,

$$\begin{aligned}\|\nabla v_m\|_2^2 &\ge \Lambda(\|v_m\eta_{R_m}\|_p^2 + \|v_m(1-\eta_{R_m})\|_p^2) + o(1)\\ &= \Lambda(\lambda^{2/p} + (1-\lambda)^{2/p}) + o(1).\end{aligned}$$

But since $\lambda^{2/p} + (1-\lambda)^{2/p} > 1$ this contradicts the assumption that $\|\nabla v_m\|_2 \to \Lambda$. □

Before the second concentration-compactness lemma can be stated, it must be explained what is meant by weak convergence of measures. The definition is as follows: $\mu_m \rightharpoonup \mu$ if

$$\int_{\mathbf{R}^n} \xi \mathrm{d}\mu_m \to \int_{\mathbf{R}^n} \xi \mathrm{d}\mu \text{ for all } \xi \in C_c(\mathbf{R}^n).$$

A related concept, used in probability, uses continuous bounded functions instead of continuous functions with compact support; it is sometimes called tight convergence because it is equivalent to weak convergence plus tightness. Convergence of measures is useful because any sequence of measures μ_m satisfying

$$\sup_{m\in\mathbf{N}} \int_K \mathrm{d}\mu_m < \infty \text{ for all compact } K \subset \mathbf{R}^n$$

has a weakly convergent subsequence. The difficulty in proving this fact mainly lies in showing that the limiting linear functional is in fact a measure, and always involve an application of the Riesz-Markov representation theorem, or some other such result (e.g. that any distribution function $F : \mathbf{R} \to [0,1]$ corresponds to a probability measure on $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$). More information can be found in [EG25, Section 1.9].

Theorem D.1 (Second Concentration-Compactness Lemma [Lio85, Lemma I.1]). *Suppose u_m is a bounded sequence in $D^{1,2}(\mathbf{R}^n)$, such that*

$$\begin{aligned} u_m &\rightharpoonup u \textit{ weakly in } D^{1,2}(\mathbf{R}^n),\\ |\nabla u_m|^2\mathrm{d}x &\rightharpoonup \mu \textit{ weakly in the sense of measures,}\\ |u_m|^p\mathrm{d}x &\rightharpoonup \nu \textit{ weakly in the sense of measures.}\end{aligned}$$

Then the limits satisfy

$$\mu \geq |\nabla u|^2 \mathrm{d}x + \Lambda \sum_{j \in J} \nu_j^{2/p} \delta_{x_j},$$
$$\nu = |u|^p \mathrm{d}x + \sum_{j \in J} \nu_j \delta_{x_j},$$

where Λ is the best constant for the Sobolev inequality (1.5).

STEP 2: USING TIGHTNESS. We have shown in the first part of the proof that $|v_m|^p \mathrm{d}x$ belongs (up to subsequences) to the "compactness" case of the first concentration-compactness lemma: that is, there exists $x_m \in \mathbf{R}^n$ such that for all $\varepsilon > 0$, there exists $R > 0$ satisfying

$$\int_{B_R(x_m)} |v_m|^p \mathrm{d}x > 1 - \varepsilon \text{ for all } m.$$

By (5.1), for all $\varepsilon < 1/2$, we must have $B_R(x_m) \cap B_1(0) \neq \emptyset$, so $B_R(x_m) \subset B_{2R+1}(0)$, and so the above hold for the sequence $x_m = 0$. (In other words, $|v_m|^p \mathrm{d}x$ is a tight sequence of probability measures.)

Suppose $v_m \rightharpoonup v$ weakly in $D^{1,2}(\mathbf{R}^n)$, $|\nabla v_m|^2 \mathrm{d}x \rightharpoonup \mu$, and $|v_m|^p \mathrm{d}x \rightharpoonup \nu$. Then

$$\int_{\mathbf{R}^n} \mathrm{d}\mu \leq \liminf_{m \to \infty} \int_{\mathbf{R}^n} |\nabla v_m|^2 \mathrm{d}x = \Lambda,$$

while tightness of $|v_m|^p \mathrm{d}x$ implies that

$$\int_{\mathbf{R}^n} \mathrm{d}\nu = 1.$$

By the second concentration-compactness lemma,

$$\mu \geq |\nabla v|^2 \mathrm{d}x + \Lambda \sum_{j \in J} \nu_j^{2/p} \delta_{x_j},$$
$$\nu = |v|^p \mathrm{d}x + \sum_{j \in J} \nu_j \delta_{x_j}.$$

Since $2/p < 1$, we have

$$\begin{aligned} \Lambda \geq \int_{\mathbf{R}^n} \mathrm{d}\mu &\geq \|\nabla v\|_2^2 + \Lambda \sum_{j \in J} \nu_j^{2/p} \\ &\geq \Lambda \|v\|_p^2 + \Lambda \sum_{j \in J} \nu_j^{2/p} \qquad (5.2) \\ &\geq \Lambda (\|v\|_p^p + \sum_{j \in J} \nu_j)^{2/p} = \Lambda \int_{\mathbf{R}^n} \mathrm{d}\nu = \Lambda, \qquad (5.3) \end{aligned}$$

with equality at (5.3) only if at most one of $\|v\|_p^p$ and ν_j, $j \in J$ is nonzero. By (5.1), $\nu_j \leq 1/2$ for all $j \in J$. Therefore $\|v\|_p = 1$. At this point we may already conclude from equality at (5.2) that v is an optimizer for the Sobolev inequality. Strong convergence in $D^{1,2}(\mathbf{R}^n)$ follows from weak convergence and $\|\nabla v_m\|_2 \to \|\nabla v\|_2$. □

The fact that $1/2$ is the supremum in (5.1) is needed to rule out the situation $\nu = \delta_x$, $x \in \partial B_1(0)$. This assumption is where dilation comes in.

What would happen if we tried to make a minimizing sequence converge without using dilation? Let $u_m = C(|x - me_1|^2 + 1)^{(2-n)/2}$, where $e_1 = (1, 0, \ldots, 0) \in \mathbf{R}^n$, C is the appropriate normalizing constant, and let us consider the rescaling $v_m = \tilde{R}_m^{n/p} u_m(\tilde{R}_m x)$, $\int_{B_1(0)} |v_m|^p \mathrm{d}x = 1/2$. Then the rescaling factor satisfies $m < \tilde{R}_m \leq m + \tilde{R}_0$, so the supremum of v_m is achieved at $\tilde{R}_m^{-1} m e_1 \to e_1$. Thus $v_m \rightharpoonup \delta_{e_1}$, as expected.

We end this essay by noting that the concentration-compactness principle can be used in a similar way to prove Theorem 1.1, the main difference being that the first concentration-compactness lemma is trivial on a compact manifold. Details can be found in [Neu18].

A Removal of Singularities

Theorem A.1. *Suppose u is a weak solution of $(\Delta + h)u = 0$ on $\mathbf{R}^n \setminus \{0\}$, where $h \in L^{n/2}(\mathbf{R}^n)$, and suppose $u \in L^q(\mathbf{R}^n)$ for some $q > n/(n-2)$. Then $(\Delta + h)u = 0$ weakly on all of $\mathbf{R}^n$.*

Proof. Let $\varphi \in C_c^\infty(\mathbf{R}^n)$. For all $\varepsilon > 0$, let the test function η_ε be defined as in Section 2. Then since $(1 - \eta_\varepsilon)\varphi \in C_c^\infty(\mathbf{R}^n \setminus \{0\})$, we have

$$\int_{\mathbf{R}^n} (u\Delta\varphi + hu\varphi)\mathrm{d}x = \int_{\mathbf{R}^n} (u\Delta(\eta_\varepsilon\varphi) + hu\eta_\varepsilon\varphi)\mathrm{d}x. \tag{A.1}$$

We claim that the right hand side of (A.1) tends to 0 as $\varepsilon \to 0$. Since the left hand side is constant it must be zero, and the desired result follows.

By Hölder's inequality, hu is locally integrable, so

$$\int_{\mathbf{R}^n} hu\eta_\varepsilon\varphi \mathrm{d}x \to 0 \text{ as } \varepsilon \to 0.$$

Since $|\eta_\varepsilon| \leq 1$, $|\nabla\eta_\varepsilon| \leq C\varepsilon^{-1}$, and $|\Delta\eta_\varepsilon| \leq C\varepsilon^{-2}$, we have

$$|\Delta(\eta_\varepsilon\varphi)| = |\Delta\eta_\varepsilon\varphi + 2\nabla\eta_\varepsilon\nabla\varphi + \eta_\varepsilon\Delta\varphi| \leq C\varepsilon^{-2},$$

and thus by Hölder's inequality

$$\left|\int_{\mathbf{R}^n} u\Delta(\eta_\varepsilon\varphi)\mathrm{d}x\right| \le C\varepsilon^{-2}\int_{B_{2\varepsilon}} |u|\mathrm{d}x$$
$$\le C\varepsilon^{-2}\|u\|_q \operatorname{Vol}(B_{2\varepsilon})^{(q-1)/q}$$
$$= C\varepsilon^{-2}\varepsilon^{n(q-1)/q}.$$

The assumption that $q > n/(n-2)$ implies $n(q-1)/q > 2$, and the proof is finished. □

B The Brezis-Lieb Lemma

Theorem B.1 (Brezis-Lieb). *Suppose f_n is a bounded sequence in L^p, $1 < p < \infty$, such that $f_n \to f$ a.e. Then $f_n \rightharpoonup f$ weakly in L^p, and*

$$\lim_{n\to\infty}\int \big||f_n|^p - |f_n - f|^p - |f|^p\big| = 0. \tag{B.1}$$

In particular,

$$\int |f_n|^p = \int |f|^p + \int |f_n - f|^p + o(1). \tag{B.2}$$

For $0 < p \le 1$, we have the inequality $\big||f_n|^p - |f_n - f|^p\big| \le |f|^p$, so by the dominated convergence theorem (B.1) still holds; in addition, the condition $\sup\|f_n\|_p < \infty$ can be weakened to $\|f\|_p < \infty$. The weak convergence statement is no longer true: a counterexample is the sequence $(1,0,0,\ldots), (0,1,0,\ldots), (0,0,1,\ldots), \ldots$ in ℓ^p, since $\Lambda(x_1, x_2, \ldots) = \sum_{k=1}^\infty x_k$ is a continuous linear functional on ℓ^p.

In the rest of this section we assume $p > 1$ unless otherwise specified. But first, we prove the weak convergence part of the theorem.

Proof of weak convergence. It suffices to show that, under the additional assumption that $f_n \rightharpoonup g$ weakly in L^p, we have $g = f$. Indeed, this implies every subsequence has a subsubsequence converging weakly to f, and hence $f_n \rightharpoonup f$ weakly.

For $n \in \mathbf{N}$, let C_n be the set of finite linear combinations of $\{f_n, f_{n+1}, f_{n+2}, \ldots\}$. By Mazur's theorem, the closure of C_n in L^p is weakly closed, so $g \in \overline{C}_n$, and hence g is the a.e. limit of a sequence of finite linear combinations of $\{f_n, f_{n+1}, f_{n+2}, \ldots\}$. In particular, for almost all x, if $|f_m(x) - f(x)| < \varepsilon$ for all $m \ge n$, then $|g(x) - f(x)| \le \varepsilon$. Letting $n \to \infty$ and then $\varepsilon \to 0$ we see that $g = f$ a.e. □

B.1 Proof by Vitali's Convergence Theorem

We first recall Vitali's theorem. A collection $\{f_j : j \in \mathcal{J}\}$ of integrable functions on a measure space $(\Omega, \mathcal{F}, \mu)$ are said to be *equi-integrable* if the following conditions are all satisfied:

(1) The functions f_j satisfy

$$\sup_{j \in \mathcal{J}} \int_\Omega |f_j| < \infty.$$

(2) For all $\varepsilon > 0$ there exists $A_\varepsilon \in \mathcal{F}$ with $\mu(A_\varepsilon) < \infty$ such that

$$\sup_{j \in \mathcal{J}} \int_{\Omega \setminus A_\varepsilon} |f_j| < \varepsilon.$$

(3) For all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\sup_{j \in \mathcal{J}} \int_A |f_j| < \varepsilon \text{ for all } \mu(A) < \delta.$$

Theorem (Vitali's Convergence Theorem). *If a sequence of equi-integrable functions f_n converge a.e. to f, then $f_n \to f$ in L^1.*

The following proof of (B.2) is taken from [Str08, Equation (4.4)]. Unfortunately it does not adapt easily to (B.1).

Proof 1. For all $n \in \mathbf{N}$, we have

$$\begin{aligned}
\|f_n\|_p^p - \|f_n - f\|_p^p &= -\int_\Omega \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} |f_n - \theta f|^p \mathrm{d}\theta \mathrm{d}\mu \\
&= p \int_\Omega \int_0^1 \operatorname{Re}\left[\overline{f}(f_n - \theta f)\right] |f_n - \theta f|^{p-2} \mathrm{d}\theta \mathrm{d}\mu \\
&= p \int_0^1 \int_\Omega \operatorname{Re}\left[\overline{f}(f_n - \theta f)\right] |f_n - \theta f|^{p-2} \mathrm{d}\mu \mathrm{d}\theta.
\end{aligned}$$

For $0 \le \theta \le 1$, the family $f|f_n - \theta f|^{p-1}$, $n \in \mathbf{N}$ is equi-integrable. Indeed, since $f_n - \theta f$ is bounded in L^p, all three conditions for equi-integrability follow from Hölder's inequality and the fact that they hold for $|f|^p$. So by Vitali's theorem and the dominated convergence theorem,

$$\lim_{n \to \infty} \left(\|f_n\|_p^p - \|f_n - f\|_p^p\right) = p \int_0^1 \int_\Omega (1 - \theta)^{p-1} |f|^p \mathrm{d}\mu \mathrm{d}\theta = \|f\|_p^p.$$

This finishes the proof. ∎

Now we set out to prove Vitali's theorem.

Lemma B.2. *Suppose $\{f_j : j \in \mathcal{J}\}$ satisfy condition (2). Then they are equi-integrable if and only if*

$$\sup_{j\in\mathcal{J}} \int_{|f_j|>c} |f_j| \to 0 \text{ as } c \to \infty.$$

Proof. If $\{f_j : j \in \mathcal{J}\}$ are equi-integrable, then by condition (1),

$$\sup_{j\in\mathcal{J}} \mu(|f_j| > c) \leq \frac{1}{c} \sup_{j\in\mathcal{J}} \int_{\Omega} |f_j| \to 0 \text{ as } c \to \infty,$$

and the desired result follows from condition (3). Conversely, choosing c large enough so that

$$M := \sup_{j\in\mathcal{J}} \int_{|f_j|>c} |f_j| < \infty,$$

we have

$$\sup_{j\in\mathcal{J}} \int_{\Omega} |f_j| \leq 1 + c\mu(A_1) + M < \infty,$$

so condition (1) is satisfied. For all $\varepsilon > 0$, let c be large enough such that

$$\sup_{j\in\mathcal{J}} \int_{|f_j|>c} |f_j| < \frac{\varepsilon}{2}.$$

Then

$$\int_A |f_j| \leq c\mu(A) + \frac{\varepsilon}{2} < \varepsilon \text{ for all } \mu(A) < \frac{\varepsilon}{2c}.$$

This proves condition (3). □

Proof of Vitali's convergence theorem. Using the characterization of Lemma B.2, it can be shown that the sequence $g_n = |f_n - f|$ is equi-integrable. For all $\varepsilon > 0$, since $f_n \to f$ a.e., we have $\mu(g_n \geq \varepsilon) \to 0$ as $n \to \infty$, so

$$\limsup_{n\to\infty} \int_{g_n\geq\varepsilon} g_n \leq \limsup_{n\to\infty} \left(c\mu(g_n \geq \varepsilon) + \int_{g_n\geq c} |g_n| \right) \leq \sup_{n\in\mathbb{N}} \int_{g_n\geq c} |g_n|.$$

Since g_n is equi-integrable, the right hand side can be taken arbitrarily small, which means

$$\int_{g_n\geq\varepsilon} g_n \to 0 \text{ as } n \to \infty.$$

On the other hand, let $\eta > 0$ be fixed; then for all $n \in \mathbb{N}$ we have

$$\int_{g_n<\varepsilon} g_n \leq \varepsilon\mu(A_\eta) + \eta.$$

Since both η and ε are arbitrary, this finishes the proof. □

B.2 Proof by Weak Convergence

The weak convergence result can be used to give a very elegant proof of (B.1). We follow [Bre11, Exercise 4.17].

Proof 2. Since $|f_n - f|$ is bounded in L^p and converges a.e. to 0, $|f_n - f| \rightharpoonup 0$ weakly in L^p, and similarly $|f_n - f|^{p-1} \rightharpoonup 0$ weakly in $L^{p/(p-1)}$. Hence,

$$\int |f_n - f||f|^{p-1} + |f_n - f|^{p-1}|f| \to 0 \text{ as } n \to \infty.$$

Using Lemma B.3 with $a = f_n - f$ and $b = f$, we have

$$\int ||f_n|^p - |f_n - f|^p - |f|^p| \to 0 \text{ as } n \to \infty,$$

and (B.1) follows immediately. □

Lemma B.3. *For all $p > 1$, there exists a constant C_p such that*

$$||a+b|^p - |a|^p - |b|^p| \leq C_p(|a||b|^{p-1} + |a|^{p-1}|b|) \textit{ for all } a, b \in \mathbb{C}.$$

Proof. Wlog suppose $|a| \leq |b|$. By homogeneity we may assume $b = 1$, and the equation becomes

$$||a+1|^p - |a|^p - 1| \leq C_p(|a| + |a|^{p-1}), \tag{B.3}$$

with $|a| \leq 1$. Since the left hand side is bounded above, it suffices to prove (B.3) for all $|a| \leq 1/12345$. Under this additional assumption, we have $||a+1|^p - 1| \leq C||a+1| - 1| \leq C|a|$. We also have $|a|^p \leq |a|$, so $||a+1|^p - |a|^p - 1| \leq C|a|$. This finishes the proof. □

B.3 Proof by Egorov's Theorem

A more direct proof of (B.1), not involving weak convergence, was given in [Lie83].

Proof 3. By Lemma B.3, it suffices to show that

$$\int |f_n - f||f|^{p-1} + |f_n - f|^{p-1}|f| \to 0.$$

By Hölder's inequality,

$$\int |f_n - f||f|^{p-1} + |f_n - f|^{p-1}|f| \leq 2\Big(\|(f_n - f)1_E\|_p \|f\|_p + \|f_n - f\|_p \|f 1_{E^c}\|_p\Big)$$

for any measurable set E.

Now $f_n \to f$ a.e. with respect to the finite measure $|f|^p \mathrm{d}\mu$, so by Egorov's theorem, for all $\varepsilon > 0$ there is a measurable set E with $\|f 1_{E^c}\|_p^p < \varepsilon$ such that $f_n \to f$ uniformly on E. Thus

$$\limsup_{n\to\infty} \left(\|(f_n - f)1_E\|_p \|f\|_p + \|f_n - f\|_p \|f 1_{E^c}\|_p \right) \leq C\varepsilon^{1/p},$$

and we are done. □

B.4 The Original Proof

The original paper [BL83] gave a proof uniting the cases $p > 1$ and $0 < p \leq 1$. However, their phrasing seems to suggest that the proof from [Lie83] had been discovered first.

Proof 4. Using Lemma B.4 with $a = f$ and $b = f_n - f$ gives

$$||f_n|^p - |f_n - f|^p - |f|^p| \leq ||f_n|^p - |f_n - f|^p| + |f|^p \leq \varepsilon |f_n - f|^p + (1 + C_{p,\varepsilon})|f|^p,$$

so the quantity

$$W_{\varepsilon,n} = (||f_n|^p - |f_n - f|^p - |f|^p| - \varepsilon |f_n - f|^p)_+$$

is dominated by $(1 + C_{p,\varepsilon})|f|^p$, and so by the dominated convergence theorem

$$\int W_{\varepsilon,n} \to 0 \text{ as } n \to \infty.$$

Now,

$$\int ||f_n|^p - |f_n - f|^p - |f|^p| \leq \int W_{\varepsilon,n} + \varepsilon \int |f_n - f|^p,$$

and since $\|f_n - f\|_p^p \leq (\|f_n\|_p + \|f\|_p)^p$ is bounded, the limsup of the right hand side can be made arbitrarily small. This finishes the proof. □

Lemma B.4. *For any $p > 0$, $\varepsilon > 0$, there exists a constant $C_{p,\varepsilon}$ such that*

$$||a + b|^p - |b|^p| \leq \varepsilon |b|^p + C_{p,\varepsilon} |a|^p \text{ for all } a, b \in \mathbb{C}.$$

Proof. For $0 < p \leq 1$,

$$|a + b|^p - |b|^p \leq (|a| + |b|)^p - |b|^p \leq |a|^p,$$

and replacing b by $-(a + b)$ we have

$$|b|^p - |a + b|^p \leq |a|^p,$$

which yields the desired inequality.

Suppose $p > 1$. The function $t \mapsto t^p$ is convex on $[0, \infty)$, so

$$|a + b|^p \leq (|a| + |b|)^p = \left((1 - \lambda)\frac{|a|}{1 - \lambda} + \lambda\frac{|b|}{\lambda}\right)^p \leq (1 - \lambda)^{-p+1}|a|^p + \lambda^{-p+1}|b|^p$$

for any $0 < \lambda < 1$. Setting $\lambda = (1 + \varepsilon_0)^{1/(-p+1)}$ we get

$$|a + b|^p - |b|^p \leq C|a|^p + \varepsilon_0|b|^p,$$

and replacing b by $-(a + b)$ we have

$$|b|^p - |a + b|^p \leq C|a|^p + \varepsilon_0|a + b|^p \leq (1 + \varepsilon_0)C|a|^p + \varepsilon_0(1 + \varepsilon_0)|b|^p.$$

Since $\varepsilon_0(1 + \varepsilon_0)$ can be made arbitrarily small, this finishes the proof. □

C The First Concentration-Compactness Lemma

Given a probability measure μ on $\mathbf{R}^n$, its *concentration function* is defined by

$$Q(r) = \sup_{x \in \mathbf{R}^n} \int_{B_r(x)} \mathrm{d}\mu.$$

It is a nondecreasing, nonnegative function on $[0, \infty)$ with $\lim_{r\to\infty} Q(r) = 1$. (If $\mu \ll \mathrm{d}x$, then Q is continuous, and in particular $\lim_{r\to 0} Q(r) = 0$.)

Let μ_m be a sequence of probability measures on $\mathbf{R}^n$, and let Q_m be their concentration functions. Then up to subsequences, $Q_m \to Q$ at all points of continuity of Q. Let $\lambda = \lim_{n\to\infty} Q(n)$. Then $0 \leq \lambda \leq 1$.

Case I: $\lambda = 1$. Then for all $\varepsilon > 0$, there exists $R > 0$ such that $Q_m(R) > 1 - \varepsilon$ for all m, so there exists a sequence $x_m \in \mathbf{R}^n$ such that

$$\int_{B_R(x_m)} \mathrm{d}\mu_m > 1 - \varepsilon \text{ for all } m.$$

Actually, x_m can be chosen independently of ε. Let x_m be chosen for $\varepsilon_0 = \frac{1}{2}$, so that

$$\int_{B_{R_0}(x_m)} \mathrm{d}\mu_m > \frac{1}{2} \text{ for all } m \text{ and some fixed } R_0 > 0.$$

Now for any $0 < \varepsilon < \frac{1}{2}$, there exists $R_1 = R_1(\varepsilon) > 0$ and a sequence y_m such that

$$\int_{B_{R_1}(y_m)} \mathrm{d}\mu_m > 1 - \varepsilon \text{ for all } m,$$

and since $\frac{1}{2} + 1 - \varepsilon > 1$, it follows that $B_{R_0}(x_m) \cap B_{R_1}(y_m) \neq \emptyset$, so $|x_m - y_m| < R_0 + R_1$, and hence $B_{R_1}(y_m) \subset B_{R_0+2R_1}(x_m)$. We set $R = R_0 + 2R_1(\varepsilon)$ to finish the proof.

Case II: $\lambda = 0$. Then $\lim_{m\to\infty} Q_m(R) = 0$ for all R.

Case III: $0 < \lambda < 1$. Let $R_m \to \infty$ be a sequence such that $Q(R_m) > \lambda - 1/m$. Using a diagonal argument, we can pass to a subsequence of μ_m such that

$$\lambda - \frac{1}{m} < Q_m(R_m) \leq Q_m(2R_m) < \lambda + \frac{1}{m} \text{ for all } m,$$

and so there exists a sequence x_m such that

$$\lambda - \frac{1}{m} < \int_{B_{R_m}(x_m)} \mathrm{d}\mu_m \leq \int_{B_{2R_m}(x_m)} \mathrm{d}\mu_m < \lambda + \frac{1}{m} \text{ for all } m,$$

which means

$$\int_{B_{R_m}(x_m)} \mathrm{d}\mu_m = \int_{B_{2R_m}(x_m)} \mathrm{d}\mu_m = \lambda + o(1).$$

To summarize, we have the following trichotomy.

Theorem C.1 (First Concentration-Compactness Lemma [Lio84, Lemma I.1])**.** *Any sequence μ_m of probability measures on $\mathbf{R}^n$ has a subsequence satisfying one of the following three conditions:*

> **Compactness.** *There exists a sequence $x_m \in \mathbf{R}^n$ such that for any $\varepsilon > 0$, there exists $R > 0$ satisfying*
>
> $$\int_{B_R(x_m)} \mathrm{d}\mu_m > 1 - \varepsilon \textit{ for all } m.$$
>
> **Vanishing.** *For all $R > 0$,*
>
> $$\sup_{x\in\mathrm{R}} \int_{B_R(x)} \mathrm{d}\mu_m \to 0 \textit{ as } m \to \infty.$$
>
> **Dichotomy.** *There exist $0 < \lambda < 1$, $x_m \in \mathbf{R}^n$, $R_m \to \infty$, such that*
>
> $$\begin{cases} \displaystyle\int_{B_{R_m}(x_m)} \mathrm{d}\mu_m = \lambda + o(1), \\ \displaystyle\int_{B_{2R_m}(x_m)\setminus B_{R_m}(x_m)} \mathrm{d}\mu_m = o(1), \\ \displaystyle\int_{\mathrm{R}^n\setminus B_{2R_m}(x_m)} \mathrm{d}\mu_m = 1 - \lambda + o(1). \end{cases}$$

D The Second Concentration-Compactness Lemma

Theorem D.1 (Second Concentration-Compactness Lemma [Lio85, Lemma I.1]). *Suppose u_m is a bounded sequence in $D^{1,2}(\mathbf{R}^n)$, such that*

$$
\begin{aligned}
u_m &\rightharpoonup u \text{ weakly in } D^{1,2}(\mathbf{R}^n),\\
|\nabla u_m|^2 \mathrm{d}x &\rightharpoonup \mu \text{ weakly in the sense of measures,}\\
|u_m|^p \mathrm{d}x &\rightharpoonup \nu \text{ weakly in the sense of measures.}
\end{aligned}
$$

Then the limits satisfy

$$
\begin{aligned}
\mu &\geq |\nabla u|^2 \mathrm{d}x + \Lambda \sum_{j\in J} \nu_j^{2/p} \delta_{x_j},\\
\nu &= |u|^p \mathrm{d}x + \sum_{j\in J} \nu_j \delta_{x_j},
\end{aligned}
$$

where Λ is the best constant for the Sobolev inequality (1.5).

Proof. Suppose $|\nabla(u_m - u)|^2 \mathrm{d}x \rightharpoonup \lambda$ and $|u_m - u|^p \mathrm{d}x \rightharpoonup \omega$. Then for any $\xi \in C_c^\infty(\mathbf{R}^n)$, $\xi \geq 0$,

$$
\begin{aligned}
\int_{\mathbf{R}^n} \xi^p \mathrm{d}\omega &= \int_{\mathbf{R}^n} \xi^p |u_m - u|^p \mathrm{d}x + o(1)\\
&\leq \liminf_{m\to\infty} \Lambda^{-p/2} \|\nabla(\xi(u_m - u))\|_2^{p/2}.
\end{aligned}
$$

By the triangle inequality,

$$
\|\nabla(\xi(u_m - u))\|_2 \leq \|\xi\nabla(u_m - u)\|_2 + \|(\nabla\xi)(u_m - u)\|_2.
$$

There is a subsequence of u_m that converges to u strongly in $L^2(\operatorname{supp}\xi)$, so

$$
\liminf_{m\to\infty} \|\nabla(\xi(u_m - u))\|_2 \leq \liminf_{m\to\infty} \|\xi\nabla(u_m - u)\|_2,
$$

and thus we have

$$
\begin{aligned}
\int_{\mathbf{R}^n} \xi^p \mathrm{d}\omega &\leq \Lambda^{-p/2} \liminf_{m\to\infty} \|\xi\nabla(u_m - u)\|_2^{p/2}\\
&= \Lambda^{-p/2} \left(\int_{\mathbf{R}^n} \xi^2 \mathrm{d}\lambda \right)^{p/2}.
\end{aligned}
$$

This can be rewritten as

$$
\int_{\mathbf{R}^n} \xi^2 \mathrm{d}\lambda \geq \Lambda \left(\int_{\mathbf{R}^n} \xi^p \mathrm{d}\omega \right)^{2/p}.
$$

For any open set U, letting $\xi \uparrow 1_U$ we have

$$\int_U \mathrm{d}\lambda \geq \Lambda\left(\int_U \mathrm{d}\omega\right)^{2/p}.$$

All locally finite Borel measures on $\mathbf{R}^n$ are regular, so the above holds with any measurable set in place of U. In particular, $\omega \ll \lambda$. By the Lebesgue-Besicovitch differentiation theorem, $\omega = f\mathrm{d}\lambda$ where

$$f(x) = \lim_{r\to 0} \frac{\int_{B_r(x)} \mathrm{d}\omega}{\int_{B_r(x)} \mathrm{d}\lambda} \text{ exists for } \lambda\text{-almost all } x.$$

So for any x not an atom of λ, provided that the above limit exists, we have $f(x) = 0$. Let us write

$$\omega = \omega_0 + \sum_{j\in J} \nu_j \delta_{x_j},$$

where ω_0 is atomless and nonnegative. Letting $\xi \downarrow 1_x$, we have

$$\lambda \geq \Lambda \nu_j^{2/p} \delta_{x_j} \text{ for all } j \in J.$$

In particular, any atom of ω is an atom of λ. Now $\omega_0 = f_0\mathrm{d}\lambda$ where $f_0 \leq f$ λ-a.e., so $f_0 = 0$ (λ-a.e.) outside J, and by construction $f_0 = 0$ in J as well. Hence, $\omega_0 = 0$,

$$\omega = \sum_{j\in J} \nu_j \delta_{x_j}.$$

(We have now shown that the theorem holds when $u = 0$.)

For all $\xi \in C_c(\mathbf{R}^n)$, by weak convergence, we have

$$\int_{\mathbf{R}^n} \xi(\mathrm{d}\mu - \mathrm{d}\lambda) = \lim_{m\to\infty} \int_{\mathbf{R}^n} \xi(|\nabla u_m|^2 - |\nabla(u_m - u)|^2)\mathrm{d}x = \int_{\mathbf{R}^n} \xi|\nabla u|^2 \mathrm{d}x,$$

so

$$\mu = |\nabla u|^2 \mathrm{d}x + \lambda \geq |\nabla u|^2 \mathrm{d}x + \Lambda \sum_{j\in J} \nu_j^{2/p} \delta_{x_j}.$$

By the strong Brezis-Lieb lemma (B.1),

$$\int_{\mathbf{R}^n} \xi(\mathrm{d}\nu - \mathrm{d}\omega) = \lim_{m\to\infty} \int_{\mathbf{R}^n} \xi(|u_m|^p - |u_m - u|^p)\mathrm{d}x = \int_{\mathbf{R}^n} \xi|u|^p \mathrm{d}x,$$

so we have

$$\nu = |u|^p \mathrm{d}x + \omega = |u|^p \mathrm{d}x + \sum_{j\in J} \nu_j \delta_{x_j},$$

and the proof is finished. □

References

[BL83] Haïm Brezis and Elliott Lieb. "A relation between pointwise convergence of functions and convergence of functionals". In: *Proc. Amer. Math. Soc.* 88.3 (1983), pp. 486–490. DOI: 10.2307/2044999.

[Bre11] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011, pp. xiv+599. DOI: 10.1007/978-0-387-70914-7.

[CGY03] Sun-Yung A. Chang, Matthew J. Gursky, and Paul C. Yang. "Entire solutions of a fully nonlinear equation". In: *Lectures on partial differential equations.* Vol. 2. New Stud. Adv. Math. Int. Press, Somerville, MA, 2003, pp. 43–60. URL: https://web.math.princeton.edu/~chang/cgy-nir-012102.pdf.

[EG25] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Second Edition. Textbooks in Mathematics. CRC Press, Boca Raton, FL, 2025, pp. xi+327. ISBN: 978-1-032-94644-3. DOI: 10.1201/9781003583004.

[FL12] Rupert L. Frank and Elliott H. Lieb. "A new, rearrangement-free proof of the sharp Hardy-Littlewood-Sobolev inequality". In: *Spectral theory, function spaces and inequalities.* Vol. 219. Oper. Theory Adv. Appl. Birkhäuser/Springer Basel AG, Basel, 2012, pp. 55–67. DOI: 10.1007/978-3-0348-0263-5_4.

[GNN79] B. Gidas, Wei-Ming Ni, and L. Nirenberg. "Symmetry and related properties via the maximum principle". In: *Comm. Math. Phys.* 68.3 (1979), pp. 209–243. DOI: 10.1007/BF01221125.

[Heb14] Emmanuel Hebey. *Compactness and stability for nonlinear elliptic equations.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2014, pp. x+291. ISBN: 978-3-03719-134-7. DOI: 10.4171/134.

[Lie83] Elliott H. Lieb. "Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities". In: *Ann. of Math. (2)* 118.2 (1983), pp. 349–374. DOI: 10.2307/2007032.

[Lio84] P.-L. Lions. "The concentration-compactness principle in the calculus of variations. The locally compact case. I". In: *Ann. Inst. H. Poincaré Anal. Non Linéaire* 1.2 (1984), pp. 109–145. DOI: 10.1016/S0294-1449(16)30428-0.

[Lio85] P.-L. Lions. "The concentration-compactness principle in the calculus of variations. The limit case. I". In: *Rev. Mat. Iberoamericana* 1.1 (1985), pp. 145–201. DOI: 10.4171/RMI/6.

[LP87] John M. Lee and Thomas H. Parker. "The Yamabe problem". In: *Bull. Amer. Math. Soc. (N.S.)* 17.1 (1987), pp. 37–91. DOI: 10.1090/S0273-0979-1987-15514-5.

[Neu18] Robin Neumayer. "The Yamabe Problem". Lecture notes. 2018. URL: https://www.math.cmu.edu/~rneumaye/YamabeProblem.pdf.

[Oba71] Morio Obata. "The conjectures on conformal transformations of Riemannian manifolds". In: *Journal of Differential Geometry* 6.2 (1971), pp. 247–258. DOI: 10.4310/jdg/1214430407.

[Str08] Michael Struwe. *Variational methods*. Fourth Edition. Vol. 34. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. Springer-Verlag, Berlin, 2008, pp. xx+302. DOI: 10.1007/978-3-540-74013-1.

[SY94] R. Schoen and S.-T. Yau. *Lectures on differential geometry*. Vol. I. Conference Proceedings and Lecture Notes in Geometry and Topology. International Press, Cambridge, MA, 1994, pp. v+235. ISBN: 1-57146-012-8.

[Tal76] Giorgio Talenti. "Best constant in Sobolev inequality". In: *Ann. Mat. Pura Appl. (4)* 110 (1976), pp. 353–372. DOI: 10.1007/BF02418013.

[Yam60] Hidehiko Yamabe. "On a deformation of Riemannian structures on compact manifolds". In: *Osaka Math. J.* 12 (1960), pp. 21–37. DOI: 10.18910/8577.