SOBOLEV INEQUALITIES AND UNCERTAINTY PRINCIPLES IN MATHEMATICAL PHYSICS. PART I

RUPERT L. FRANK

ABSTRACT. This is a first draft of lecture notes for a course given at the LMU Munich in July 2011.

1. SOBOLEV INEQUALITIES AND UNCERTAINTY PRINCIPLES

1.1. **Example: Stability of the hydrogen atom.** We consider a single particle of mass 1/2 moving in three-dimensional space in the presence of a potential V.

We recall that in classical mechanics the state of such a system (at a given time) is described by a point $(p, x) \in \mathbb{R}^3 \times \mathbb{R}^3$, where x denotes the particle position and p the particle momentum. The energy, which is conserved under the time evolution, is

$$H(p, x) = p^2 + V(x) \,.$$

In quantum mechanics a state is described by a function $\psi \in L_2(\mathbb{R}^3)$, a *wave function*, with

$$\int_{\mathbb{R}^3} |\psi(x)|^2 \, dx = 1 \,. \tag{1.1}$$

In view of this normalization requirement, the function $|\psi|^2$ can be interpreted as the probability density of the position of the particle. On the other hand, by (1.1) and Plancherel's theorem we know that the Fourier transform of ψ,

$$\hat{\psi}(p) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} e^{-ip\cdot x} \psi(x) \, dx \,, \tag{1.2}$$

satisfies a similar normalization condition

$$\int_{\mathbb{R}^3} |\hat{\psi}(p)|^2 \, dp = 1 \,. \tag{1.3}$$

That is, $|\hat{\psi}|^2$ has the interpretation of the probability density of the momentum of the particle. The energy of the system in the state ψ is

$$h[\psi] = \int_{\mathbb{R}^3} p^2 |\hat{\psi}(p)|^2 \, dp + \int_{\mathbb{R}^3} V(x) |\psi(x)|^2 \, dx = \int_{\mathbb{R}^3} \left(|\nabla \psi(x)|^2 + V(x) |\psi(x)|^2 \right) \, dx \,.$$

Again, under certain weak assumptions on V one can construct a time evolution under which this energy is constant. In this lectures, however, we will not consider the time evolution, but we consider the stationary problem of *minimizing the energy*.

The authors wishes to thank T. Ried and V. von Conta for helpful remarks.

Despite a certain formal correspondence between the definitions in the classical and in the quantum case there are fundamental differences. Probably the most important one is the *uncertainty principle.* It is instructive to study the effect of this principle in the simple example of a single electron interacting with a fixed nucleus via attractive Coulomb forces. The classical energy of this system is

$$H^{\text{Coulomb}}(p,x) = p^2 - \kappa |x|^{-1} \,,$$

where the parameter $\kappa > 0$ (with the dimension of 1/length) contains all physical constants. We observe that $H(p,x)$, considered as a function on $\mathbb{R}^3 \times \mathbb{R}^3$, is unbounded from below. There is no minimal energy state! By making $|x|$ small (considering an electron close to the nucleus) and preventing $|p|$ from being too large, we can make $H(p,x)$ arbitrarily negative.

This is not true in quantum mechanics! The reason for this is the uncertainty principle which roughly states that $|p|$ and $|x|$ cannot be simultaneously small. So when $|x|$ we try to make $|x|$ small, as before, we have to make $|p|$ large. While this is intuitively correct, the question of what exactly 'small' and 'large' here means leads to non-trivial mathematics, and this is the content of this course.

We now give to arguments of why the quantum mechanical energy

$$h^{\text{Coulomb}}[\psi] = \int_{\mathbb{R}^3} \left(|\nabla \psi(x)|^2 - \kappa |x|^{-1} |\psi(x)|^2 \right) \, dx \,.$$

is bounded from below. Certainly, at this point the knowing reader will remember that the quantum-mechanical hydrogen problem can be solved explicitly. (... did this already in ..., several years before the advent of quantum mechanics.) `Pauli, Jan. 1926` Our point here is that only very few Schrödinger eigenvalue problems can be solved explicitly. Screening effects, for instance, lead to physical situations with effective potentials, for which the eigenvalues are typically not known in closed form. On the other hand, the comparison with the known result in the Coulomb case give a natural measure of how good our methods are.

Here is the first mathematical formulation of the uncertainty principle.

Theorem 1.1 (Hardy inequality)**.** *For any $\psi \in \dot H^1(\mathbb{R}^3)$ one has*

$$\int_{\mathbb{R}^3} |\nabla \psi|^2 \, dx \geq \frac{1}{4} \int_{\mathbb{R}^3} \frac{|\psi|^2}{|x|^2} \, dx \,. \tag{1.4}$$

The inequality is strict for every $\psi \not\equiv 0$, but the constant $\frac14$ cannot be replaced by a smaller constant.

This inequality says that if ψ is localized close to a point, e.g., $x = 0$, (i.e., the right side is large), then its momentum has to be large (i.e., the left side is large). We also note that a stronger singularity $|x|^{-2}$ can be controlled by the kinetic energy than the $|x|^{-1}$ singularity which arises in our Coulomb problem. The power $|x|^{-2}$ is, of course, dictated by the dimensionality of the gradient on the left side. Indeed, if an inequality of the form $\int |\nabla\psi|^2 \, dx \geq C \int |x|^\alpha |\psi|^2 \, dx$ holds with a constant independent of ψ, then a simple scaling argument (i.e., replacing $\psi(x)$ by $\psi(\lambda x)$) shows that necessarily $\alpha = -2$.

Hardy's inequality (1.4) should be compared with the better known Heisenberg's uncertainty principle,

$$\left(\int_{\mathbb{R}^3} |\nabla\psi|^2\,dx\right)^{1/2} \left(\int_{\mathbb{R}^3} x^2|\psi|^2\,dx\right)^{1/2} \geq \frac{3}{2}\int_{\mathbb{R}^3} |\psi|^2\,dx\,.$$

This inequality says that, if $\int |\nabla\psi|^2\,dx$ is not too big, then $\int x^2|\psi|^2\,dx$ is not too small. This is not too useful in practice, however. The value $\int x^2|\psi|^2\,dx$ can very well be large and, at the same time, ψ can be sharply localized close to the origin, as long as there is an additional bump of ψ very far out. In contrast, Hardy's inequality says that, if $\int |\nabla\psi|^2\,dx$ is not too big, then $\int |x|^{-2}|\psi|^2\,dx$ is not too big either, and this excludes ψ being sharply localized near the origin. This is what we need to prove the quantum-mechanical energy of hydrogen is finite.

Remark 1.2. As an aside, note that the Heisenberg inequality follows (with a non-sharp constant) from the Hardy inequality via Jensen's inequality.

We shall prove Hardy's and Heisenberg's inequalities in the following subsection. Accepting them for the moment, we shall use the former to prove that the quantum mechanical energy of the hydrogen problem is finite. Indeed, (1.4) implies that

$$h^{\mathrm{Coulomb}}[\psi] = \int_{\mathbb{R}^3} \left(\frac{1}{4}|x|^{-2} - \kappa|x|^{-1}\right)|\psi(x)|^2\,dx\,,$$

and therefore, putting $\rho = |\psi|^2$,

$$\inf_{\|\psi\|=1} h^{\mathrm{Coulomb}}[\psi] \geq \inf\{\int_{\mathbb{R}^3} \left(\frac{1}{4}|x|^{-2} - \kappa|x|^{-1}\right)\rho(x)\,dx :\ \rho \geq 0\,, \int_{\mathbb{R}^3}\rho(x)\,dx = 1\}\,.$$

The minimization problem on the right side can easily be solved and leads to the lower bound $-\kappa^2$. (The density ρ wants to be supported only on the sphere $|x| = (2\kappa)^{-1}$.)

The second formulation of the uncertainty principle, from which we will be able to deduce a lower bound on $h^{\mathrm{Coulomb}}[\psi]$, is Sobolev's inequality. In contrast to Hardy's inequality, which is a *linear* inequality, in the sense that it involves the square of ψ on both sides, Sobolev's inequality is a *non-linear* inequality, where ψ appears with a higher power on the left than on the right side. This makes it, intuitively, a stronger inequality; (see, however, the discussion in Remark 5.8.)

Theorem 1.3 (Sobolev inequality)**.** *For any $\psi \in \dot{H}^1(\mathbb{R}^3)$ one has*

$$\int_{\mathbb{R}^3} |\nabla\psi|^2\,dx \geq S_3\left(\int_{\mathbb{R}^3}|\psi|^6\,dx\right)^{1/3} \tag{1.5}$$

with $S_3 = 3(\pi/2)^{4/3}$. Equality holds iff $u(x) = c\,h(b(x-a))$ for some $a \in \mathbb{R}^3$, $b > 0$ and $c \in \mathbb{C}$, where

$$h(x) = (1+x^2)^{-1/2}\,.$$

The complete proof of this theorem will be one of the main results of the first part of this course. A simple proof, with a non-optimal constant, however, will be given in Theorem 1.12. The complete result is in Theorem 4.1.

The non-linearity $|\psi|^6$ in Sobolev's inequality is dictated by scaling reasons in the same way as the function $|x|^{-2}$ is in Hardy's inequality. That is, replacing $\psi(x)$ by $\psi(\lambda x)$ we see that an inequality of the form $\int |\nabla\psi|^2\,dx \geq C\left(\int |\psi|^q\,dx\right)^{2/q}$ can only hold for $q=6$ in $N=3$.

Using Sobolev's inequality (1.5) we arrive at the minimization problem

$$\inf\{h^{\rm Coulomb}[\psi]:\ \|\psi\|=1\} \geq \inf\{S_3\left(\int_{\mathbb{R}^3}\rho(x)\,dx\right)^{1/3} - \kappa\int_{\mathbb{R}^3}|x|^{-1}\rho(x)\,dx:\ \rho\geq 0\,,\int_{\mathbb{R}^3}\rho(x)\,dx=1\}\,.$$

It is an easy exercise to compute the infimum on the right side. (No gradients are involved anymore!) The optimal density is of the form $\rho(x) = C(|x|^{-1}-A\kappa)^{1/2}$ if $|x|<(A\kappa)^{-1}$ and $\rho(x)=0$ if if $|x|\geq (A\kappa)^{-1}$ with two constants C and A. Eliminating C via the constaint $\int\rho\,dx=1$, optimizing in A and recalling the precise value of S_3 from Theorem 1.3 we arrive at

$$\inf\{S_3\left(\int_{\mathbb{R}^3}\rho(x)\,dx\right)^{1/3} - \kappa\int_{\mathbb{R}^3}|x|^{-1}\rho(x)\,dx:\ \rho\geq 0\,,\int_{\mathbb{R}^3}\rho(x)\,dx=1\} = -\frac{1}{3}\kappa^2\,.$$

This is remarkably close to the true answer

$$\inf\{h^{\rm Coulomb}[\psi]:\ \|\psi\|=1\} = -\frac{1}{4}\kappa^2\,.$$

1.2. **Hardy inequalities.** We now explain how to prove Theorem 1.1 and its generalization to higher dimensions.

Theorem 1.4 (Hardy's inequality)**.** *Let $N\geq 3$. Then for all $u\in\dot H^1(\mathbb{R}^N)$*

$$\int_{\mathbb{R}^N}|\nabla u|^2\,dx \geq \left(\frac{N-2}{2}\right)^2\int_{\mathbb{R}^N}\frac{|u|^2}{|x|^2}\,dx\,. \tag{1.6}$$

The inequality is strict for every $u\not\equiv 0$, but the constant $((N-2)/2)^2$ cannot be replaced by a smaller constant.

Lemma 1.5. *Let ω be a positive function satisfying $-\Delta\omega+V\omega\geq 0$ in Ω. Then for all $u\in C_0^1(\Omega)$*

$$\int_{\mathbb{R}^N}\left(|\nabla u|^2+V|u|^2\right)\,dx \geq \int_{\mathbb{R}^N}|\nabla(\omega^{-1}u)|^2\omega^2\,dx\,.$$

Equality holds when $-\Delta\omega+V\omega=0$.

Proof. We use the *ground state substitution* $u=\omega v$ and compute

$$|\nabla u|^2 = |\omega\nabla v+v\nabla\omega|^2 = \omega^2|\nabla v|^2 + \nabla|v|^2\cdot\omega\nabla\omega + |v|^2|\nabla v|^2\,.$$

Integrating this over $\mathbb{R}^N$ yields

$$\begin{aligned}\int_{\mathbb{R}^N} |\nabla u|^2\,dx &= \int_{\mathbb{R}^N} \left(\omega^2|\nabla v|^2 - |v|^2\operatorname{div}(\omega\nabla\omega) + |v|^2|\nabla v|^2\right)\,dx \\ &= \int_{\mathbb{R}^N} \left(\omega^2|\nabla v|^2 - |v|^2\omega(\Delta\omega)\right)\,dx \\ &\geq \int_{\mathbb{R}^N} \left(\omega^2|\nabla v|^2 - V|v|^2\omega\omega\right)\,dx\,.\end{aligned}$$

This is the claimed inequality. □

Proof of Theorem 1.4. A simple computation shows that $\omega(x) = |x|^{-(N-2)/2}$ satisfies $-\Delta\omega + V\omega = 0$ for $V(x) = -((N-2)/2)^2|x|^{-2}$ in $\Omega = \mathbb{R}^N \setminus \{0\}$. Theorem 1.4 follows since $C_0^\infty(\Omega)$ is dense in $\dot{H}^1(\mathbb{R}^N)$. The optimality claim is left as an exercise to the reader. □

Examples 1.6. Here are a few further applications of Lemma 1.5.

(1) *Heisenberg uncertainty principle.* For $\omega(x) = e^{-\alpha x^2/2}$ and $V(x) = \alpha^2 x^2 - \alpha N$ we obtain

$$\int_{\mathbb{R}^N} \left(|\nabla u|^2 + \alpha^2 x^2 |u|^2\right)\,dx \geq \alpha N \int_{\mathbb{R}^N} |u|^2\,dx$$

with equality iff u is proportional to ω. Optimizing in α, we find

$$\left(\int_{\mathbb{R}^N} |\nabla u|^2\,dx\right)^{1/2} \left(\int_{\mathbb{R}^N} |x|^2|u|^2\,dx\right)^{1/2} \geq \frac{N}{2}\int_{\mathbb{R}^N} |u|^2\,dx\,.$$

(2) *Hydrogen uncertainty principle.* For $N \geq 2$, $\omega(x) = e^{-\alpha|x|}$ and $V(x) = -\alpha(N-1)|x|^{-1} + \alpha^2$ we obtain

$$\int_{\mathbb{R}^N} \left(|\nabla u|^2 - \frac{\alpha(N-1)}{|x|}|u|^2\right)\,dx \geq -\alpha^2 \int_{\mathbb{R}^N} |u|^2\,dx$$

with equality iff u is proportional to ω. Optimizing in α, we find

$$\left(\int_{\mathbb{R}^N} |\nabla u|^2\,dx\right)^{1/2} \left(\int_{\mathbb{R}^N} |u|^2\,dx\right)^{1/2} \geq \frac{N-1}{2}\int_{\mathbb{R}^N} \frac{|u|^2}{|x|}\,dx\,.$$

(3) *Linearized Sobolev inequality.* For $N \geq 3$, $\omega(x) = (1+x^2)^{-(N-2)/2}$ and $V(x) = -N(N-2)(1+x^2)^{-2}$ we obtain

$$\int_{\mathbb{R}^N} |\nabla u|^2\,dx \geq N(N-2)\int_{\mathbb{R}^N} \frac{|u|^2}{(1+x^2)^2}\,dx$$

with equality iff u is proportional to ω. This inequality will prove useful in Section 4 below.

Exercise 1.7. half-line inequality

Exercise 1.8. half-space inequality

1.3. **Sobolev inequalities.** The proof of Sobolev's inequality, even with non-sharp constant, is substantially more difficult than that of Hardy's inequality. Below we follow the proof due to Gagliardo and Nirenberg which goes via the quantity $\int |\nabla u|\,dx$ instead of the quantity $\int |\nabla u|^2\,dx$.

Theorem 1.9 (Isoperimetric inequality). *Let $N \geq 2$. Then there is a constant $C_N > 0$ such that for all $u \in \dot W^1(\mathbb{R}^N)$*

$$\int_{\mathbb{R}^N} |\nabla u|\,dx \geq C_N \left(\int_{\mathbb{R}^N} |u|^{N/(N-1)}\,dx \right)^{(N-1)/N} . \tag{1.7}$$

The constant C_N that our proof gives is not sharp. Essentially, we obtain the sharp constant in Theorem 5.10. The reason why we call (1.7) an isoperimetric inequality will be explained in Exercise 1.14. For bibliographical remarks we refer to the discussion of Theorem 5.10.

The Gagliardo–Nirenberg argument of Theorem 1.9 relies on two lemmas. The first one is a *one-dimensional* analogue of the inequality we want to establish. The second one gives a method of how to pass from *smaller* to *higher* dimensions.

Lemma 1.10 (Easiest Sobolev inequality). *If $u \in \dot W^1(\mathbb{R})$, then $\int_{\mathbb{R}} |u'|\,dx \geq 2\sup_{x\in\mathbb{R}} |u(x)|$. Equality holds iff there is a constant $a \in \mathbb{R}$ such that u is non-decreasing on $(-\infty, a)$ and non-increasing on (a,∞).*

Proof. Write $u(x) = \frac12 \left(\int_{-\infty}^x u'(y)\,dy - \int_x^\infty u'(y)\,dy \right)$. □

In the statement of the next lemma, we shall use the following notation for $x \in \mathbb{R}^N$ and $1 \leq j \leq N$,

$$\tilde x_j = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_N) \in \mathbb{R}^{N-1} .$$

Then one has

Lemma 1.11 (Loomis–Whitney inequality). *Let $N \geq 2$ and let $f_1, \ldots, f_N \in L^{N-1}(\mathbb{R}^{N-1})$. Then the function $f(x) := f_1(\tilde x_1) \cdots f_N(\tilde x_N)$ belongs to $L^1(\mathbb{R}^N)$ and*

$$\|f\|_{L^1(\mathbb{R}^N)} \leq \prod_{j=1}^N \|f_j\|_{L^{N-1}(\mathbb{R}^{N-1})} .$$

Proof. Note that this is an equality for $N = 2$. The main idea behind the proof is most clearly seen for $N = 3$, which we assume henceforth. We have

$$\int_{\mathbb{R}^3} |f(x)|\,dx = \iint_{\mathbb{R}\times\mathbb{R}} f_1(x_1,x_2) I(x_1,x_2)\,dx_1\,dx_2 ,$$

where

$$I(x_1,x_2) = \int_{\mathbb{R}} f_2(x_1,x_3) f_3(x_1,x_2)\,dx_3 .$$

By the Schwarz inequality, $I(x_1,x_2) \leq \sqrt{g_2(x_1)}\sqrt{g_3(x_2)}$, where

$$g_2(x_1) = \int_{\mathbb{R}} f_2(x_1,x_3)^2\,dx_3 \quad \text{and} \quad g_3(x_2) = \int_{\mathbb{R}} f_3(x_1,x_2)^2\,dx_3 .$$

Thus, again by Schwarz,

$$\begin{aligned}\int_{\mathbb{R}^3} |f(x)|\,dx &\leq \int_{\mathbb{R}} dx_1 \sqrt{g_2(x_1)} \left(\int_{\mathbb{R}} f_1(x_1,x_2)^2\,dx_2\right)^{1/2} \left(\int_{\mathbb{R}} g_3(x_2)\,dx_2\right)^{1/2} \\ &= \|f_3\|_2 \int_{\mathbb{R}} dx_1 \sqrt{g_2(x_1)}\sqrt{g_1(x_1)}\,,\end{aligned}$$

where g_1 is defined similarly as g_2 and g_3. Applying Schwarz once again, we arrive at the claimed inequality for $N=3$. The case $N \geq 4$ is proved similarly. □

Proof of Theorem 1.9. By Lemma 1.10 we have

$$|u(x)| \leq \frac{1}{2}\int_{\mathbb{R}} |\partial_1 u(y_1, x_2, \ldots, x_N)|\,dy_1 = g_1(\tilde{x}_1)\,.$$

Defining g_2, ..., g_N similarly with respect to the other coordinate directions and multiplying the inequalities we find that

$$|u(x)|^N \leq g_1(\tilde{x}_1)\cdots g_N(\tilde{x}_N)\,,$$

which is the same as

$$|u(x)|^{N/(N-1)} \leq g_1(\tilde{x}_1)^{1/(N-1)}\cdots g_N(\tilde{x}_N)^{1/(N-1)}\,.$$

Applying Lemma 1.11 with $f_j = g_j^{1/(N-1)}$ we infer that

$$\int_{\mathbb{R}^N} |u(x)|^{N/(N-1)}\,dx \leq \prod_{j=1}^N \|g_j^{1/(N-1)}\|_{L^{N-1}(\mathbb{R}^{N-1})} = \prod_{j=1}^N \|g_j\|_{L^1(\mathbb{R}^{N-1})}^{1/(N-1)}\,.$$

The arithmetic-geometric mean inequality,

$$\left(\prod_{j=1}^N a_j\right)^{1/N} \leq \frac{1}{N}\sum_{j=1}^N a_j$$

for $a_j \geq 0$, implies that

$$\left(\int_{\mathbb{R}^N} |u(x)|^{N/(N-1)}\,dx\right)^{(N-1)/N} \leq \frac{1}{N}\sum_{j=1}^N \|g_j\|_{L^1(\mathbb{R}^{N-1})} = \frac{1}{2N}\sum_{j=1}^N \int_{\mathbb{R}^N} |\partial_j u|\,dx\,.$$

Estimating the ℓ_1-norm of the gradient in $\mathbb{C}^N$ in terms of its ℓ_2-norm, we obtain the claimed inequality. □

Corollary 1.12. *Let $2 \leq q \leq \infty$ if $N = 1,2$ and let $2 \leq q \leq 2N/(N-2)$ if $N \geq 3$. Then there is a constant $S_{N,q} > 0$ such that for every $u \in H^1(\mathbb{R}^N)$ one has*

$$\left(\int_{\mathbb{R}^N} |\nabla u|^2\,dx\right)^{\theta} \left(\int_{\mathbb{R}^N} |u|^2\,dx\right)^{1-\theta} \geq S_{N,q}\left(\int_{\mathbb{R}^N} |u|^q\,dx\right)^{2/q}, \tag{1.8}$$

where $N/q = \theta(N-2)/2 + (1-\theta)N/2$.

Of course, the equation determining θ comes from scaling.

An important special case of (1.8) is, for $N \geq 3$,

$$\int_{\mathbb{R}^N} |\nabla u|^2 \, dx \geq S_N \left(\int_{\mathbb{R}^N} |u|^{2N/(N-2)} \, dx \right)^{(N-2)/N} .$$

(Here and in the following we write $S_N = S_{N,2N/(N-2)}$ in the special case $q = 2N/(N-2)$.) This proves Theorem 1.3, except for the value of the constant S_3 and the characterization of optimizers. This inequality, which was stated for functions in the inhomogeneous Sobolev space $H^1(\mathbb{R}^N)$ extends by continuity to the homogeneous space $\dot{H}^1(\mathbb{R}^N)$.

Proof. We first note that by Hölder's inequality, if (1.8) holds for some q, it holds for all smaller values of $q \geq 2$. Hence it suffices to derive (1.8) only for large values of q, and this is what we do in the following.

Case $N \geq 3$ and $q = 2N/(N-2)$. For the proof we recall that if $u \in \dot{W}^1_{1,\mathrm{loc}}$ and $\alpha \geq 1$, then $|u|^\alpha \in \dot{W}^1_{1,\mathrm{loc}}$ and $|\nabla |u|^\alpha| = \alpha |u|^{\alpha-1} |\nabla u|$ in the sense of distributions; see, e.g., [LiLo, Thm. 6.17] for a similar argument. Hence, applying Theorem 1.9 with $|u|^\alpha$ in place of u, we obtain

$$\begin{aligned} \left(\int_{\mathbb{R}^N} |u|^{\alpha N/(N-1)} \, dx \right)^{N/(N-1)} & \leq \alpha C_N^{-1} \int_{\mathbb{R}^N} |u|^{\alpha-1} |\nabla u| \, dx \\ & \leq \alpha C_N^{-1} \left(\int_{\mathbb{R}^N} |\nabla u|^2 \, dx \right)^{1/2} \left(\int_{\mathbb{R}^N} |u|^{2(\alpha-1)} \, dx \right)^{1/2} . \end{aligned}$$

Choosing $\alpha = (N-1)2/(N-2)$ yields (1.8) if we note that $\alpha N/(N-1) = 2N/(N-2) = 2(\alpha-1)$.

Case $N = 2$ and $q \geq 4$. We use the same argument as before, but this time we choose $\alpha = q/2$ depending on q and obtain

$$\left(\int_{\mathbb{R}^2} |u|^q \, dx \right)^2 \leq \frac{2}{q \, C_2} \left(\int_{\mathbb{R}^2} |\nabla u|^2 \, dx \right)^{1/2} \left(\int_{\mathbb{R}^2} |u|^{q-2} \, dx \right)^{1/2} .$$

Since $q \geq 4$, the term $\int |u|^{q-2} \, dx$ can be estimated by Hölder in terms of $\int |u|^q \, dx$ and $\int |u|^2 \, dx$. This yields the desired inequality.

Case $N = 1$ and $q = \infty$. We prefer to record this separately as Corollary 1.13, since in this case we actually obtain the sharp constant.

According to the remark made at the beginning of the proof that large q is enough, the proof of Corollary 1.12 is complete. □

Here is the announced sharp inequality in the one-dimensional case.

Corollary 1.13. *Let $N = 1$ and $q = \infty$. Then*

$$\left(\int_{\mathbb{R}} |u'|^2 \, dx \right)^{1/2} \left(\int_{\mathbb{R}} |u|^2 \, dx \right)^{1/2} \geq \sup_{x \in \mathbb{R}} |u(x)|^2 ,$$

with equality iff $u(x) = c\ h(b(x-a))$ *for some* $a \in \mathbb{R}$, $b > 0$ *and* $c \in \mathbb{C}$, *where* $h(x) = e^{-|x|}$.

Proof. The proof uses the same strategy as that of Corollary 1.12, except that we use Lemma 1.10 instead of Theorem 1.9. Namely, we apply the inequality of Lemma 1.10 to u^2 instead of u and use the Schwarz inequality for $\int |u||u'|\,dx$. We note that this Schwarz inequality is an equality iff $|u|$ and $|u'|$ are proportional. The only functions which satisfy this, together with the condition of Lemma 1.10 are the exponentials stated in Theorem 1.13. □

The same argument as in Corollaries 1.12 and 1.13 yields Sobolev inequalities in $W_p^1(\mathbb{R}^N)$ as well. Namely, if $p \le q \le Np/(N-p)$ for $p < N$ and if $p \le q < \infty$ for $p \ge N$ there is a constant $S_{N,p,q} > 0$ such that for every $u \in W_p^1(\mathbb{R}^N)$ one has

$$\left(\int_{\mathbb{R}^N} |\nabla u|^p\,dx\right)^\theta \left(\int_{\mathbb{R}^N} |u|^p\,dx\right)^{1-\theta} \ge S_{N,p,q} \left(\int_{\mathbb{R}^N} |u|^q\,dx\right)^{p/q},$$

where $N/q = \theta(N-p)/p + (1-\theta)N/p$.

Exercise 1.14. Here is the reason why we call (1.7) an isoperimetric inequality.

(1) Deduce from Theorem 1.9 that for any sufficiently smooth domain $\Omega \subset \mathbb{R}^N$ one has

$$|\partial\Omega| \ge C_N |\Omega|^{(N-1)/N} \tag{1.9}$$

with the same constant C_N as in (1.7). (Here $|\partial\Omega|$ stands for the $(N-1)$-dimensional surface measure of $\partial\Omega$, whereas $|\Omega|$ stands for the N-dimensional volume measure. For the proof, take u to be an approximation to the characteristic function of Ω.

(2) Denote by Ω^* the ball in $\mathbb{R}^N$, centered at the origin, with the same volume as Ω. Deduce that

$$|\partial\Omega| \ge C_N N^{-(N-1)/N} |\mathbb{S}^{N-1}|^{-1/N} |\partial\Omega^*|\,.$$

This is an isoperimetric inequality, which says that the boundary of any set cannot be much smaller (in a controlled way) than that of the ball with the same volume. Actually, in Theorem 5.10 we shall see that balls have the smallest boundary.

(3) Conversely, inequality (1.9) implies inequality (1.7). Deduce by accepting the co-area formula

$$\int_{\mathbb{R}^N} |\nabla u|\,dx = \int_0^\infty |\{|u| = \tau\}|\,d\tau\,,$$

where $|\{|u| = \tau\}|$ denotes the $(N-1)$-dimensional surface measure of the level set $\{|u| = \tau\}$. The proof of this formula (in particular, for Sobolev functions) is rather involved; see `Reference`

2. An isoperimetric problem for the ground state energy of Schrödinger operators

In this section we consider the following 'isoperimetric' problem: How negative can the lowest eigenvalue of a Schrödinger operator $-\Delta + V$ be given an L^p-norm of the potential? Can it be arbitrarily negative or are there p for which there is a universal lower bound? If it exists, are there 'optimal' potentials? Are they unique?

To make this problem more precise, we define

$$\lambda_1(-\Delta+V) = \inf_u \frac{\int_{\mathbb{R}^N} \left(|\nabla u|^2 + V|u|^2\right)\,dx}{\int_{\mathbb{R}^N} |u|^2\,dx}$$

and

$$L^{(1)}_{\gamma,N} := \sup_V \frac{\lambda_1(-\Delta+V)_-^\gamma}{\int_{\mathbb{R}^N} V_-^{\gamma+N/2}\,dx} \tag{2.1}$$

Hence the problems mentioned at the beginning of this section can be rephrased as follows: For which values of γ is $L^{(1)}_{\gamma,N}$ finite? If it is, for which V is the supremum attained?

`Dilation invariance;` $V \mapsto b^2 V(bx)$

The isoperimetric problem for Schrödinger eigenvalues goes back to Keller [Ke], who gave an explicit solution in one-dimension. (The case $p=1$ in $N=1$ is an even older result.) We will discuss Keller's theorem and its extension to higher dimensions. A key ingredient will be a link between this problem and the problem of the sharp constant in the Sobolev interpolation inequalities from Corollary 1.12.

2.1. Duality for lowest eigenvalues of Schrödinger operators.

The main result of this section is an equivalent formulation of the problem of computing $L^{(1)}_{\gamma,N}$ in terms of Sobolev interpolation inequalities. We define

$$S_{N,q} := \inf_u \frac{\left(\int_{\mathbb{R}^N} |\nabla u|^2\,dx\right)^\theta \left(\int_{\mathbb{R}^N} |u|^2\,dx\right)^{1-\theta}}{\left(\int_{\mathbb{R}^N} |u|^q\,dx\right)^{2/q}}\,, \qquad \theta = \frac{N}{2}\left(1-\frac{2}{q}\right). \tag{2.2}$$

Proposition 2.1 (Duality)**.** *Let γ and q be related by*

$$\frac{1}{\gamma+N/2} + \frac{2}{q} = 1\,.$$

We assume that $\gamma \geq 1/2$ if $N=1$ and $\gamma>0$ if $N\geq 2$, that is, we assume that $2<q\leq\infty$ if $N=1$ and $2<q<2N/(N-2)$ if $N\geq 2$. Then problems (2.1) *and* (2.2) *are dual in the sense that*

$$\left(L^{(1)}_{\gamma,N}\right)^{1/(\gamma+N/2)} S_{N,q} = \theta^\theta (1-\theta)^{1-\theta}\,,$$

where θ is given in (2.2) *or, alternatively, $\theta = (N/2)/(\gamma+N/2)$. Moreover, u attains the infimum in* (2.2) *iff $V=-\alpha|u|^{q-2}$ attains the supremum in* (2.1)*.*

Proof. By definition of $L^{(1)}_{\gamma,N}$ one has for every $u \in H^1(\mathbb{R}^N)$

$$\int_{\mathbb{R}^N} \left(|\nabla u|^2 + V|u|^2\right) dx \geq -\left(L^{(1)}_{\gamma,N} \int_{\mathbb{R}^N} V_-^{\gamma+N/2}\, dx\right)^{1/\gamma} \int_{\mathbb{R}^N} |u|^2\, dx$$

We now assume that γ and q are related as in the statement of the theorem, i.e., $(\gamma + N/2)(q-2) = q$, and choose for given u the function $V = -\alpha|u|^{q-2}$. Optimizing the resulting inequality over $\alpha > 0$ we find

$$\left(\int_{\mathbb{R}^N} |\nabla u|^2\, dx\right)^{\theta} \left(\int_{\mathbb{R}^N} |u|^2\, dx\right)^{1-\theta} \geq \left(L^{(1)}_{\gamma,N}\right)^{-1/(\gamma+N/2)} \theta^\theta (1-\theta)^{1-\theta} \left(\int_{\mathbb{R}^N} |u|^q\, dx\right)^{2/q}.$$

By the definition of $S_{N,q}$, this implies

$$S_{N,q} \geq \left(L^{(1)}_{\gamma,N}\right)^{-1/(\gamma+N/2)} \theta^\theta (1-\theta)^{1-\theta}.$$

Conversely, by Hölder's inequality (with $1/(\gamma+N/2)+2/q = 1$) and by the definition of $S_{N,q}$ we can bound for every $u \in H^1(\mathbb{R}^N)$ with $\int |u|^2\, dx = 1$

$$\begin{aligned}\int_{\mathbb{R}^N} V|u|^2\, dx &\geq -\left(\int_{\mathbb{R}^N} V_-^{\gamma+N/2}\, dx\right)^{1/(\gamma+N/2)} \left(\int_{\mathbb{R}^N} |u|^q\, dx\right)^{2/q} \\ &\geq -S_{N,q}^{-1} \left(\int_{\mathbb{R}^N} V_-^{\gamma+N/2}\, dx\right)^{1/(\gamma+N/2)} \left(\int_{\mathbb{R}^N} |\nabla u|^2\, dx\right)^{\theta}.\end{aligned}$$

Thus

$$\int_{\mathbb{R}^N} \left(|\nabla u|^2 + V|u|^2\right) dx \geq T - S_{N,q}^{-1} \left(\int_{\mathbb{R}^N} V_-^{\gamma+N/2}\, dx\right)^{1/(\gamma+N/2)} T^\theta$$

with $T = \int |\nabla u|^2\, dx$. Minimizing the right side over all T we obtain the

$$\int_{\mathbb{R}^N} \left(|\nabla u|^2 + V|u|^2\right) dx \geq -\theta^{\theta/(1-\theta)}(1-\theta) S_{N,q}^{-1/(1-\theta)} \left(\int_{\mathbb{R}^N} V_-^{\gamma+N/2}\, dx\right)^{1/(1-\theta)(\gamma+N/2)}.$$

By the definition of $L^{(1)}_{\gamma,N}$, this implies that

$$\left(L^{(1)}_{\gamma,N}\right)^{1/\gamma} \leq \theta^{\theta/(1-\theta)}(1-\theta) S_{N,q}^{-1/(1-\theta)}.$$

□

Exercise 2.2. Use Exercise `insert!` and a similar argument as in the previous proof to show that $L^{(1)}_{\gamma,N} = \infty$ if $0 < \gamma < 1/2$ and $N = 1$. `Similar question in higher dimensions`

We next discuss the limiting case $\gamma = 0$ of Theorem 2.1. We define

$$L^{(1)}_{0,N} := \sup_{V:\ \lambda_1(-\Delta+V)\geq 0} \frac{1}{\int_{\mathbb{R}^N} V_-^{N/2}\, dx} \tag{2.3}$$

In other words, $L^{(1)}_{\gamma,N}$ is the smallest constant such that $\int_{\mathbb{R}^N} V_-^{N/2}\,dx \leq \left(L^{(1)}_{0,N}\right)^{-1}$ implies that $-\Delta + V \geq 0$. Following the same argument as in the proof of Theorem 2.1 we find a duality principle with the critical Sobolev inequality,

$$S_N := \inf_u \frac{\int_{\mathbb{R}^N} |\nabla u|^2\,dx}{\left(\int_{\mathbb{R}^N} |u|^{2N/(N-2)}\,dx\right)^{(N-2)/N}}\,. \tag{2.4}$$

Theorem 2.3 (Absence of eigenvalues)**.** *For $N \geq 3$ the problems* (2.3) *and* (2.4) *are dual in the sense that*

$$\left(L^{(1)}_{0,N}\right)^{2/N} S_N = 1\,.$$

Moreover, u attains the infimum in (2.2) *iff $V = -\alpha|u|^{4/(N-2)}$ attains the supremum in* (2.1)*.*

2.2. **Sharp constants in the one-dimensional case.** As we shall see now, in one dimension the isoperimetric problem discussed at the beginning of this section can be solved explicitly.

Theorem 2.4. *Let $N = 1$ and $\gamma \geq 1/2$. Then*

$$\lambda_1(-\Delta + V)_-^\gamma \leq L^{(1)}_{\gamma,1} \int_{\mathbb{R}} V_-^{\gamma+1/2}\,dx$$

where $L^{(1)}_{1/2,1} = 1/2$ for $\gamma = 1/2$ and

$$L^{(1)}_{\gamma,1} = \frac{1}{\sqrt{\pi}}\ \frac{(\gamma-\frac12)^{\gamma-\frac12}}{(\gamma+\frac12)^{\gamma+\frac12}}\ \frac{\Gamma(\gamma+1)}{\Gamma(\gamma+\frac12)}$$

for $\gamma > 1/2$. The inequality is strict for $\gamma = 1/2$, but the constant cannot be decreased. For $\gamma > 1/2$ equality holds iff $V(x) = b^2 w(b(x-a))$ for some $a \in \mathbb{R}$ and $b > 0$ where

$$w(x) = -(\gamma^2 - \tfrac14)\cosh^{-2} x\,.$$

`Remark about` δ `functions; reference to Spruch and Keller`

In view of the duality established in Proposition 2.1, this theorem is equivalent to computing the sharp value of the constant $S_{1,q}$. For $q = \infty$ (which corresponds to $\gamma = 1/2$) this has already been done in Corollary 1.13. The following proposition settles the case $q < \infty$ (which corresponds to $\gamma > 1/2$).

Proposition 2.5. *Let $N = 1$ and $2 < q < \infty$. Then*

$$\left(\int_{\mathbb{R}} |u'|^2\,dx\right)^\theta \left(\int_{\mathbb{R}} |u|^2\,dx\right)^{1-\theta} \geq S_{1,q} \left(\int_{\mathbb{R}} |u|^q\,dx\right)^{2/q}\,, \quad \theta = \frac12\left(1 - \frac2q\right)\,,$$

where

$$S_{1,q} = \theta^\theta (1-\theta)^{1-\theta} \frac{q}{2^{2/q}(q-2)^{(q-2)/q}} \left(\frac{\sqrt{\pi}\ \Gamma(\frac{q}{q-2})}{\Gamma(\frac{q}{q-2}+\frac12)}\right)^{(q-2)/q}$$

with equality iff $u(x) = ch(b(x-a))$ *for some* $a \in \mathbb{R}$, $b > 0$ *and* $c \in \mathbb{C}$, *where*

$$h(x) = \cosh^{-2/(q-2)} x \,.$$

Proof. We consider the minimization problem

$$I_q := \inf_u \frac{\left(\int_{\mathbb{R}} |u'|^2 \, dx\right)^\theta \left(\int_{\mathbb{R}} |u|^2 \, dx\right)^{1-\theta}}{\left(\int_{\mathbb{R}} |u|^q \, dx\right)^{2/q}} \,.$$

According to Corollary 1.12 this defines a strictly positive number.

Step 1. The infimum I_q is attained. This can be shown by some compactness arguments. (The one-dimensional case is actually simpler than the multi-dimensional case that we will study in detail below.) Moreover, since replacing u by $|u|$ does not decrease the quotient involved in the definition of I_q, we may henceforth assume that the infimum is attained by a *non-negative* function u. The Euler-Lagrange equation for u reads

$$-u'' - \lambda u^{q-1} = -\mu u \tag{2.5}$$

and one easily shows that both Lagrange multipliers λ and μ are positive. Hence, after a scaling and after multiplication by a positive constant, we can assume that $\lambda = \mu = 1$.

Step 2. We shall show that the only non-negative and non-zero solution in $H^1(\mathbb{R})$ of (2.5) with $\mu = \nu = 1$ is given by

$$u(x) = \left(\frac{q}{2}\right)^{1/(q-2)} \cosh^{-2/(q-2)} \left(\frac{q-2}{2}(x-a)\right) \tag{2.6}$$

for some $a \in \mathbb{R}$. Once this is proved, the value of the constant follows by a straightforward (but tedious) computation, using the fact that

$$\int_{\mathbb{R}} \cosh^{-\alpha} x \, dx = \sqrt{\pi} \frac{\Gamma(\alpha/2)}{\Gamma((\alpha+1)/2)} \,.$$

To prove (2.6), we multiply the equation by u' and find the first integral

$$-\frac{1}{2}(u')^2 - \frac{1}{q} u^q = -\frac{1}{2} u^2 + C$$

for some constant C. Since $u \in H^1(\mathbb{R})$, we have $\lim_{|x|\to\infty} u(x) = 0$ and we deduce from the previous formula that $\lim_{|x|\to\infty} (u'(x))^2$ exists and is given by $-C$. From this we conclude that $C = 0$ and consequently

$$u' = \pm \sqrt{u^2 - \frac{2}{q} u^q} \,.$$

(When solving the quadratic equation for u' a sign ambiguity arises. This ambiguity will disappear later in the proof.) This is a equation with separate variables which, in principle, can be solved in terms of an anti-derivative of $(u^2 - \frac{2}{q} u^q)^{-1/2}$. We proceed somewhat differently and introduce, following, e.g., [Fa], the function

$v(u) = \sqrt{1-\frac{2}{q}u^{q-2}}$. The equation becomes $u' = \pm uv$. On the other hand, we compute

$$\frac{dv}{du} = -\frac{q-2}{q}\frac{u^{q-3}}{v} = -\frac{q-2}{2}\frac{1-v^2}{uv}$$

and obtain

$$\frac{dv}{dx} = \frac{dv}{du}\frac{du}{dx} = \mp\frac{q-2}{2}(1-v^2)\,.$$

Recalling that $(1-v^2)^{-1}$ has anti-derivative arctanh we can integrate this equation and find that

$$x - a = \mp\frac{2}{q-2}\operatorname{arctanh} v\,,$$

that is,

$$\tanh\left(\mp\frac{q-2}{2}(x-a)\right) = v = \sqrt{1-\frac{2}{q}u^{q-2}}\,.$$

Since tanh is odd, we conclude that

$$1-\cosh^{-2}\left(\frac{q-2}{2}(x-a)\right) = \tanh^2\left(\frac{q-2}{2}(x-a)\right) = 1-\frac{2}{q}u^{q-2}\,,$$

which is what we claimed in (2.6). □

The proof of Proposition 2.5 shows, in particular, that the Euler–Lagrange equation (2.5) has a unique (up to translations) positive solution in $H^1(\mathbb{R})$. This can be proved in greater generality.

Proposition 2.6. *Let f be a Lipschitz continuous function on $\mathbb{R}$ with $f(0)=0$ and assume that for $F(z)=\int_0^z f(s)\,ds$ the infimum*

$$\zeta_0 := \inf\{\zeta>0:\ F(\zeta)=0\}$$

is attained and satisfies $\zeta_0>0$ and $f(\zeta_0)>0$. Then the problem

$$-u''=f(u)\,,\qquad \lim_{|x|\to\infty}u(x)=0\,,\qquad u(x_0)>0 \text{ for some } x_0\in\mathbb{R}\,,$$

has a unique solution up to translations. This solution is (after a translation) a positive, even and decreasing (wrt $|x|$) function with $u(0)=\zeta_0$.

This result is stated in [BeLi]. This paper also shows that the above conditions on ζ_0 are not only necessary, but also sufficient.

2.3. **The multi-dimensional case.**

Here we state the multi-dimensional analogue of Theorem 2.4.

Theorem 2.7. *Let $N\geq 2$ and $\gamma>0$. Then there is a negative, radial and increasing function w such that the supremum*

$$L_{\gamma,N}^{(1)} = \sup_V \frac{\lambda_1(-\Delta+V)_-^\gamma}{\int_{\mathbb{R}^N} V_-^{\gamma+N/2}\,dx}$$

is attained iff $V(x)=b^2\,w(b(x-a))$ for some $a\in\mathbb{R}^N$ and $b>0$.

In contrast to the one-dimensional case, the explicit values of the constants are not known. It is remarkable that one can nevertheless prove uniqueness (up to the natural symmetries of the problem).

A complete proof of this result is beyond the scope of these lectures. We only sketch the main steps in the argument. By the duality argument of Proposition 2.1 the theorem follows from

Proposition 2.8. *Let $N \geq 2$ and $2 < q < 2N/(N-2)$. Then there is a positive, radial and decreasing function h such that the infimum*

$$S_{N,q} = \inf_u \frac{\left(\int_{\mathbb{R}^N} |\nabla u|^2 \, dx\right)^\theta \left(\int_{\mathbb{R}^N} |u|^2 \, dx\right)^{1-\theta}}{\left(\int_{\mathbb{R}^N} |u|^q \, dx\right)^{2/q}} , \qquad \theta = \frac{N}{2}\left(1 - \frac{2}{q}\right) .$$

is attained iff $u(x) = c\, h(b(x-a))$ for some $a \in \mathbb{R}^N$ and $b > 0$.

The proof of this proposition consists in three major steps.

Step 1. Existence of a minimizer. This is relatively standard. In the following section we will prove the existence of a minimizer for a minimization problem in $\dot{H}^1(\mathbb{R}^N)$. The case here of a problem in $H^1(\mathbb{R}^N)$ is simpler.

Step 2. Any minimizer is radial. This follows from the method of moving planes. This method goes back to Alexandrov and Serrin and was further developed and popularized by Gidas-Ni-Nirenberg. Note that the result does *not* follow by Schwarz symmetrization for $\int |\nabla u|^2 \, dx$ (see Proposition 5.4), since there is no strict rearrangement inequality.

Step 3. Uniqueness of radial solutions. This is a non-trivial result of Kwong (building upon previous works of `Insert`); see also `Tao's book` for an exposition of the proof. We emphasis that the non-linear ODE, which arises as Euler-Lagrange equation,

$$-u'' - (N-1)r^{-1}u' - u^{q-1} = -u \qquad \text{on } (0,\infty) ,$$

is non-autonomous because of the first order term. This makes the uniqueness proof considerably harder than that in the one-dimensional case, where one deals with the autonomous equation 2.5.

3. A refined Sobolev inequality and its consequences

Our goal is to prove that the variational problem corresponding to the Sobolev inequality

$$S_N = \inf_{u \in \dot{H}^1(\mathbb{R}^N)} \frac{\int_{\mathbb{R}^N} |\nabla u|^2 \, dx}{\left(\int_{\mathbb{R}^N} |u|^q \, dx\right)^{\frac{N-2}{N}}}$$

admits a minimizer. Throughout this section we will assume that $N \geq 3$ and that $q = 2N/(N-2)$.

`History, concentration compactness, symmetrization`

`Proof of Talenti gives existence + sharp constant; existence relies on explicit solution of 1D problem`

3.1. **A refined Sobolev inequality.** We shall deduce the existence of a minimizer for the Sobolev inequality and a form of the concentration compactness principle from a refinement of the Sobolev inequality. In order to state this, we introduce the operator $e^{t\Delta}$ as the integral operator on $\mathbb{R}^N$ with integral kernel $t^{-N/2}G((x-y)/\sqrt{t})$, where $G(y) = (4\pi)^{-N/2}e^{-y^2/4}$, i.e.,

$$\left(e^{t\Delta}u\right)(x) = \int_{\mathbb{R}^N} t^{-N/2}G((x-y)/\sqrt{t})u(y)\,dy\,. \tag{3.1}$$

This notation is consistent with the functional calculus from spectral theory, but the only fact we need in the following is the alternative representation

$$\widehat{e^{t\Delta}u}(p) = e^{-tp^2}\hat{u}(p)\,, \tag{3.2}$$

where $\hat{u}$ is the Fourier transform of u, see (1.2). This follows by computing the Fourier transform of G. We are now ready to state the main result of this subsection.

Theorem 3.1 (Refined Sobolev inequality)**.** *Let $N \geq 3$. Then there is a constant $C_N > 0$ such that for all $u \in \dot{H}^1(\mathbb{R}^N)$*

$$\left(\int_{\mathbb{R}^N} |u|^{\frac{2N}{N-2}}\,dx\right)^{\frac{N-2}{2N}} \leq C_N \left(\int_{\mathbb{R}^N} |\nabla u|^2\,dx\right)^{\frac{N-2}{2N}} \left(\sup_{t>0} t^{(N-2)/4}\|e^{t\Delta}u\|_\infty\right)^{\frac{2}{N}}\,. \tag{3.3}$$

In other words, (3.3) says that the Sobolev inequality holds with a right side which is a certain interpolation between the usual term $\int |\nabla u|^2\,dx$ and a term $I[u] = \sup_{t>0} t^{(N-2)/4}\|e^{t\Delta}u\|_\infty$. We note that both terms behave in the same way under a rescaling of u; they have the same dimension length^{N-2}.

Inequality (3.3) is a refinement of the Sobolev inequality, since $I \leq c_N\|u\|_{2N/(N-2)}$. Indeed, by Hölder's inequality

$$\|e^{t\Delta}u\|_\infty \leq (4\pi t)^{-\frac{N}{2}} \left(\int_{\mathbb{R}^N} e^{-2Nx^2/(N+2)4t}\,dx\right)^{\frac{N+2}{2N}} \|u\|_{2N/(N-2)} = c_N t^{-\frac{N-2}{4}}\|u\|_{2N/(N-2)}\,.$$

The space of tempered distributions u for which $I[u] < \infty$ is the *Besov space* $B^{-(N-2)/2}_{\infty,\infty}$ and $I[u]$ is a norm in this space. However, we will not need the theory of Besov spaces in the following.

Inequality (3.3) appears in [GeMeOr]. Our proof below follows [Le2]. L^p `norms`

Proof. We abbreviate $q = 2N/(N-2)$ and, by homogeneity, we assume that $I[u] \leq 1$, i.e., $|e^{t\Delta}u(x)| \leq t^{-(N-2)/4}$ for all x and $t > 0$. In this way the assertion becomes

$$\int_{\mathbb{R}^N} |u|^q\,dx \leq C_N^q \int_{\mathbb{R}^N} |\nabla u|^2\,dx\,. \tag{3.4}$$

For the proof of this, we recall that

$$|u(x)|^q = \int_0^\infty \chi_{\{|u(x)|^q>\lambda\}}\,d\lambda = q\int_0^\infty \chi_{\{|u(x)|>\tau\}}\,\tau^{q-1}\,d\tau$$

('the layer cake representation') and therefore

$$\int_{\mathbb{R}^N} |u|^q \, dx = q \int_0^\infty |\{|u| > \tau\}| \; \tau^{q-1} \, d\tau \, .$$

Now we decompose $u = (u - e^{t\Delta}u) + e^{t\Delta}u$ for some $t > 0$ to be chosen below and note that

$$|\{|u| > \tau\}| \le |\{|u - e^{t\Delta}u| > \tau/2\}| + |\{|e^{t\Delta}u| > \tau/2\}| \, .$$

In particular, if we pick $t = t_\tau$ depending on τ in such a way that $\tau/2 = t^{-(N-2)/4}$, then by our normalization condition

$$|\{|e^{t\Delta}u| > \tau/2\}| = 0 \, ,$$

and hence

$$\int_{\mathbb{R}^N} |u|^q \, dx \le q \int_0^\infty |\{|u - e^{t_\tau\Delta}u| > \tau/2\}| \; \tau^{q-1} \, d\tau \, . \tag{3.5}$$

The idea now is to use the (non-sharp) inequality

$$\|v - e^{t\Delta}v\|^2 \le t \|\nabla v\|^2 \tag{3.6}$$

for any $v \in \dot{H}^1(\mathbb{R}^N)$. This easily follows in Fourier space from (3.2) and the bound $(1 - e^{-x})^2 \le 1 - e^{-x} \le x$ for all $x \ge 0$, applied to $x = tp^2$.

We now explain a direct application of (3.6) to (3.5), which, however, is not good enough to yield (3.4). The correct argument will be a slight refinement of this idea and will be presented below. The direct way is to use Chebyshev and (3.6) to get

$$\begin{aligned} |\{|u - e^{t_\tau\Delta}u| > \tau/2\}| &\le (\tau/2)^{-2} \int_{\mathbb{R}^d} |u - e^{t_\tau\Delta}u|^2 \, dx \\ &\le (\tau/2)^{-2} (\tau/2)^{-4/(N-2)} \int_{\mathbb{R}^d} |\nabla u|^2 \, dx = (\tau/2)^{-q} \int_{\mathbb{R}^d} |\nabla u|^2 \, dx \, . \end{aligned}$$

Inserting this back into (3.5) we obtain

$$\int_{\mathbb{R}^N} |u|^q \, dx \le q 2^q \int_0^\infty \tau^{-1} \, d\tau \int_{\mathbb{R}^d} |\nabla u|^2 \, dx \, ,$$

which, unfortunately, is infinite! The τ-integral diverges only logarithmically, however, which gives some hope...

The argument can be saved by applying (3.6) not directly to u, but to an approximation to u. This approximation will be chosen in a τ dependent way in order to make the logarithmic divergence disappear. Before giving a precise formula for the approximation u_τ we note that, in any case, the decomposition $u - e^{t\Delta}u = (u_\tau - e^{t\Delta}u_\tau) - e^{t\Delta}(u - u_\tau) + (u - u_\tau)$ leads to

$$|\{|u - e^{t_\tau\Delta}u| > \tfrac{\tau}{2}\}| \le |\{|u_\tau - e^{t_\tau\Delta}u_\tau| > \tfrac{\tau}{4}\}| + |\{|e^{t\Delta}(u - u_\tau)| > \tfrac{\tau}{8}\}| + |\{|u - u_\tau| > \tfrac{\tau}{8}\}| \, . \tag{3.7}$$

We begin by treating the first term and show that a clever choice of u_τ makes the argument above work.

We fix a constant $c\geq 1/16$ and define for every $\tau>0$ a function u_τ on $\mathbb{R}^N$ by

$$u_\tau(x)=\begin{cases}(c-\frac{1}{16})\tau & \text{if } u(x)>c\tau\,,\\ u(x)-\frac{\tau}{16} & \text{if } c\tau\geq u(x)\geq\frac{\tau}{16}\,,\\ 0 & \text{if } \frac{\tau}{16}>u(x)>-\frac{\tau}{16}\,,\\ u(x)+\frac{\tau}{16} & \text{if } -\frac{\tau}{16}\geq u(x)\geq -c\tau\,,\\ -(c-\frac{1}{16})\tau & \text{if } u(x)<-c\tau\,.\end{cases}$$

We note that $u_\tau\in\dot H^1(\mathbb{R}^N)$ with

$$\int_{\mathbb{R}^N}|\nabla u_\tau|^2\,dx=\int_{\tau/16\leq|u|\leq c\tau}|\nabla u|^2\,dx\,.$$

Applying (3.6) to u_τ we obtain, as before,

$$|\{|u_\tau-e^{t_\tau\Delta}u_\tau|>\tfrac{\tau}{4}\}|\leq(\tau/4)^{-2}\int_{\mathbb{R}^d}|u_\tau-e^{t_\tau\Delta}u_\tau|^2\,dx\leq 4(\tau/2)^{-q}\int_{\tau/16\leq|u|\leq c\tau}|\nabla u|^2\,dx\,.$$

The key observation now is that

$$\begin{aligned}\int_0^\infty|\{|u_\tau-e^{t_\tau\Delta}u_\tau|>\tfrac{\tau}{4}\}|\;\tau^{q-1}\,d\tau&\leq 2^{q+2}\int_{\mathbb{R}^N}dx\,|\nabla u|^2\int_{|u|/c}^{16|u|}\frac{d\tau}{\tau}\\&=2^{q+2}\log(16c)\;\int_{\mathbb{R}^N}dx\,|\nabla u|^2\,,\end{aligned}$$

which, indeed, is finite.

Finally, it remains to control the error we made by replacing u by u_τ, that is, we need to bound the last two terms in (3.7). For that purpose, we estimate

$$|u_\tau-u|=|u_\tau-u|\;\chi_{\{|u|\leq c\tau\}}+|u_\tau-u|\;\chi_{\{|u|>c\tau\}}\leq\frac{\tau}{16}+|u|\;\chi_{\{|u|>c\tau\}}\,,\tag{3.8}$$

which again by Chebyshev leads to

$$|\{|u-u_\tau|>\tfrac{\tau}{8}\}|\leq|\{|u|\;\chi_{\{|u|>c\tau\}}>\tfrac{\tau}{16}\}|\leq(\tau/16)^{-1}\int_{\mathbb{R}^N}|u|\;\chi_{\{|u|>c\tau\}}\,dx\,.$$

Similarly, since the heat kernel is positive and has integral one, we deduce from (3.8) that

$$|e^{t\Delta}u_\tau-e^{t\Delta}u|\leq e^{t\Delta}|u_\tau-u|\leq\frac{\tau}{16}+e^{t\Delta}(|u|\;\chi_{\{|u|>c\tau\}})\,,$$

and therefore

$$\begin{aligned}|\{|e^{t\Delta}u-e^{t\Delta}u_\tau|>\tfrac{\tau}{8}\}|\leq|\{e^{t\Delta}(|u|\;\chi_{\{|u|>c\tau\}})>\tfrac{\tau}{16}\}|&\leq(\tau/16)^{-1}\int_{\mathbb{R}^d}e^{t\Delta}(|u|\;\chi_{\{|u|>c\tau\}})\,dx\\&=(\tau/16)^{-1}\int_{\mathbb{R}^d}|u|\;\chi_{\{|u|>c\tau\}}\,dx\,.\end{aligned}$$

We conclude that

$$\int_0^\infty \left(|\{|e^{t\Delta}(u-u_\tau)|>\tfrac{\tau}{8}\}|+|\{|u-u_\tau|>\tfrac{\tau}{8}\}|\right)\tau^{q-1}\,d\tau \le 32\int_{\mathbb{R}^N} dx\,|u|\int_0^{|u|/c} d\tau\,\tau^{q-2}$$
$$= \frac{32}{q-1}c^{-q+1}\int_{\mathbb{R}^N} dx\,|u|^q\,.$$

To summarize, we have shown that

$$\frac{1}{q}\int_{\mathbb{R}^N}|u|^q\,dx \le 2^{q+2}\log(16c)\int_{\mathbb{R}^N} dx\,|\nabla u|^2+\frac{32}{q-1}c^{-q+1}\int_{\mathbb{R}^N} dx\,|u|^q\,.$$

Choosing c sufficiently large, we arrive at (3.4). □

3.2. **Existence of optimizers.**

Corollary 3.2. *Let $N\ge 3$ and let (u_j) be a bounded sequence in $\dot H^1(\mathbb{R}^N)$. Then one of the following alternatives occurs.*

1. *(u_j) converges to zero in $L_{2N/(N-2)}(\mathbb{R}^N)$.*
2. *There is a subsequence (u_{j_m}) and sequences $(a_m)\subset\mathbb{R}^N$ and $(b_m)\subset(0,\infty)$ such that*
$$v_m(x) := b_m^{(N-2)/2}u_{j_m}(b_m(x-a_m))$$
converges weakly in $\dot H^1(\mathbb{R}^N)$ to a function $v\not\equiv 0$. Moreover, (v_m) converges a.e. and in $L_{q,\mathrm{loc}}(\mathbb{R}^N)$, $q<2N/(N-2)$, to v.

Proof. Assume that (1) does not hold, i.e., $\|u_j\|_{2N/(N-2)}\ge\varepsilon>0$ for all sufficiently large j. Hence by the refined Sobolev inequality (3.3) and by the fact that $\|\nabla u_j\|\le A$ for all j, we have

$$\left(\sup_{t>0} t^{(N-2)/4}\|e^{t\Delta}u_j\|_\infty\right)^{\frac{2}{N}} \ge C_N^{-1}A^{-\frac{N-2}{2N}}\varepsilon\,.$$

Hence there are $t_j>0$ and $x_j\in\mathbb{R}^N$ such that $w_j(y) := t_j^{(N-2)/4}u_j(\sqrt{t_j}y+x_j)$ satisfies

$$\left|\int_{\mathbb{R}^N}G(y)w_j(y)\,dy\right| = t_j^{(N-2)/4}\left|\int_{\mathbb{R}^N}t_j^{-N/2}G((x-x_j)/\sqrt{t_j})u_j(x)\,dx\right|$$
$$= t_j^{(N-2)/4}|e^{t\Delta}u_j(x_j)| \ge \frac{1}{2}C_N^{-\frac{N}{2}}A^{-\frac{N-2}{4}}\varepsilon$$

where $G(y)=(4\pi)^{-N/2}e^{-y^2/4}$. Since $\|\nabla w_j\|=\|\nabla u_j\|\le A$, the Banach–Alaoglu theorem implies that w_j has a weakly convergent subsequence in $\dot H^1(\mathbb{R}^N)$. Since $G\in\left(\dot H^1(\mathbb{R}^N)\right)^*$, the dual of $\dot H^1(\mathbb{R}^N)$, we conclude that the limit w is not identically zero. This proves the first part of (ii). The remaining assertions follow from the Rellich–Kondrachov theorem, see, e.g., [LiLo, Thms. 8.6 and 8.7]. □

Recall that Fatou's lemma states that the pointwise limit f of a sequence of non-negative functions (f_j) satisfies

$$\liminf_{j\to\infty}\int f_j\,dx \ge \inf f\,dx\,.$$

Here, in general, one cannot expect equality. The next lemma, however, provides the 'missing term' in Fatou's lemma.

Lemma 3.3 (Brezis–Lieb lemma). *Let (X, dx) be a measure space and let (f_j) be a bounded sequence in $L_p(X)$, $0 < p < \infty$, which converges pointwise a.e. to a function f. Then*

$$\lim_{j\to\infty} \int_X ||f_j|^p - |f_j - f|^p - |f|^p| \, dx = 0 \, .$$

Proof. We write $f_j = f + g_j$ with $g_j \to 0$ pointwise a.e. and estimate

$$\int ||f_j|^p - |f_j - f|^p - |f|^p| \, dx \leq \varepsilon \int |g_j|^p \, dx + \int G_j \, dx$$

with

$$G_j = \left(||f + g_j|^p - |g_j|^p - |f|^p| - \varepsilon |g_j|^p\right)_+ \, .$$

We shall prove that (g_j) is uniformly bounded in L_p and that $\int G_j \, dx \to 0$ for every $\varepsilon > 0$. Since $\varepsilon > 0$ can be taken arbitrarily small, this implies the assertion.

Since $\|f_j\|_p \leq C$ we have $\|f\|_p \leq C$, and therefore

$$\int |g_j|^p \, dx \leq \int |f_j - f|^p \, dx \leq 2^p \int (|f_j|^p + |f|^p) \, dx \leq 2^{p+1} C \, ,$$

as claimed. In order to prove the claim about G_j we first note that for any $\varepsilon > 0$ (and $p > 0$) there is a C_ε such that for all numbers $a, b \in \mathbb{C}$

$$||a + b|^p - |b|^p| \leq \varepsilon |b|^p + C_\varepsilon |a|^p \, .$$

This implies that

$$\left||f + g_j|^p - |g^j|^p - |f|^p\right| \leq \left||f + g_j|^p - |g^j|^p\right| + |f|^p \leq \varepsilon |g_j|^p + (1 + C_\varepsilon)|f|^p \, ,$$

and hence $G_j \leq (1 + C_\varepsilon)|f|^p$. Since $g_j \to 0$ a.e. one has $G_j \to 0$ a.e. as well, and since $|f|^p$ is integrable, dominated convergence implies that $\int G_j \, dx \to 0$. This completes the proof. □

Theorem 3.4. *Let $N \geq 3$. Then the infimum*

$$S_N = \inf \frac{\int_{\mathbb{R}^N} |\nabla u|^2 \, dx}{\left(\int_{\mathbb{R}^N} |u|^{\frac{2N}{N-2}} \, dx\right)^{\frac{N-2}{N}}}$$

is attained.

Proof. Let (u_j) be a minimizing sequence, which we can assume to be normalized in $L_{2N/(N-2)}(\mathbb{R}^N)$. Then (u_j) is bounded in $\dot{H}^1(\mathbb{R}^N)$ and from Lemma 3.2 we infer that, after a translation and a dilation if necessary, (u_j) converges weakly in $\dot{H}^1(\mathbb{R}^N)$ and a.e. to a $u \not\equiv 0$. The weak convergence in $\dot{H}^1(\mathbb{R}^N)$ implies that

$$\int_{\mathbb{R}^N} |\nabla u_j|^2 \, dx = \int_{\mathbb{R}^N} |\nabla (u_j - u)|^2 \, dx + \int_{\mathbb{R}^N} |\nabla u|^2 \, dx + o(1)$$

and the a.e. convergence together with the Brezis–Lieb lemma 3.3 implies that

$$1 = \int_{\mathbb{R}^N} |u_j|^{\frac{2N}{N-2}}\, dx = \int_{\mathbb{R}^N} |u_j - u|^{\frac{2N}{N-2}}\, dx + \int_{\mathbb{R}^N} |u|^{\frac{2N}{N-2}}\, dx + o(1)\,.$$

As a consequence,

$$1 = \left(\int_{\mathbb{R}^N} |u_j|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} \leq \left(\int_{\mathbb{R}^N} |u_j - u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + \left(\int_{\mathbb{R}^N} |u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + o(1)\,.$$

Thus

$$\begin{aligned} S_N + o(1) &= \int_{\mathbb{R}^N} |\nabla u_j|^2\, dx \\ &\geq \frac{\int_{\mathbb{R}^N} |\nabla(u_j - u)|^2\, dx + \int_{\mathbb{R}^N} |\nabla u|^2\, dx + o(1)}{\left(\int_{\mathbb{R}^N} |u_j - u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + \left(\int_{\mathbb{R}^N} |u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + o(1)} \\ &\geq \frac{S_N \left(\int_{\mathbb{R}^N} |u_j - u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + \int_{\mathbb{R}^N} |\nabla u|^2\, dx + o(1)}{\left(\int_{\mathbb{R}^N} |u_j - u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + \left(\int_{\mathbb{R}^N} |u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + o(1)}\,, \end{aligned}$$

which is the same as

$$S_N + o(1) \geq \frac{\int_{\mathbb{R}^N} |\nabla u|^2\, dx + o(1)}{\left(\int_{\mathbb{R}^N} |u|^{\frac{2N}{N-2}}\, dx\right)^{\frac{N-2}{N}} + o(1)}$$

This means that u is a minimizer. □

The above argument shows also that $\int_{\mathbb{R}^N} |u_j - u|^{\frac{2N}{N-2}}\, dx \to 0$, that is, any minimizing sequence has a subsequence which converges in $L_{2N/(N-2)}(\mathbb{R}^N)$ after an appropriate translation and dilation.

3.3. **Bubble decomposition in $\dot{H}^1(\mathbb{R}^N)$.**

Theorem 3.5 (Bubble decomposition)**.** *Let $N \geq 3$ and let (u_j) be a bounded sequence in $\dot{H}^1(\mathbb{R}^N)$. Then for any $k \in \mathbb{N}$ there is a function $\phi^k \in \dot{H}^1(\mathbb{R}^N)$, a sequence $(a_j^k) \subset \mathbb{R}^N$ and a sequence $(\lambda_j^k) \in (0,\infty)$ such that, along a subsequence, one has*

$$u_j(x) = \sum_{k=1}^{K} \left(\lambda_j^k\right)^{-\frac{N-2}{2}} \phi^k((x - a_j^k)/\lambda_j^k) + r_j^K(x) \quad \text{for any } K \in \mathbb{N}\,,$$

where, for $q = 2N/(N-2)$,

$$\limsup_{K\to\infty} \limsup_{j\to\infty} \|r_j^K\|_q = 0\,, \tag{3.9}$$

$$\sup_K \limsup_{j\to\infty} \left| \|\nabla u_j\|^2 - \sum_{k=1}^{K} \|\nabla \phi^k\|^2 - \|\nabla r_j^K\|^2 \right| = 0\,, \tag{3.10}$$

$$\limsup_{K\to\infty}\limsup_{j\to\infty}\left|\|u_j\|_q^q-\sum_{k=1}^K\|\phi^k\|_q^q\right|=0\,, \tag{3.11}$$

$$\liminf_{j\to\infty}\left(\frac{|a_j^k-a_j^{k'}|^2}{\lambda_j^k\lambda_j^{k'}}+\frac{\lambda_j^k}{\lambda_j^{k'}}+\frac{\lambda_j^{k'}}{\lambda_j^k}\right)=\infty \qquad \text{for all } k\neq k'\,, \tag{3.12}$$

$$\left(\lambda_j^k\right)^{\frac{N-2}{2}} r_j^K(\lambda_j^k\cdot-a_j^k)\rightharpoonup 0 \text{ in } \dot H^1(\mathbb{R}^N) \text{ for all } 1\leq k\leq K\,. \tag{3.13}$$

We emphasize that there may be only finitely many non-zero ϕ^k.

The heart of the proof of Theorem 3.5 is the following

Lemma 3.6. *Let $N\geq 3$ and let (u_j) be a sequence in $\dot H^1(\mathbb{R}^N)$ with*

$$\lim_{j\to\infty}\|\nabla u_j\|=A \quad \text{and} \quad \liminf_{j\to\infty}\|u_j\|_q\geq\varepsilon>0\,.$$

Then there is a $0\not\equiv\phi\in\dot H^1(\mathbb{R}^N)$ and sequences $(a_j)\subset\mathbb{R}^N$ and $(\lambda_j)\subset(0,\infty)$ such that, along a subsequence, with $q=2N/(N-2)$

$$(\lambda_j)^{\frac{N-2}{2}}u_j(\lambda_j\cdot-a_j)\rightharpoonup\phi \text{ in } \dot H^1(\mathbb{R}^N)\,, \tag{3.14}$$

$$\lim_{j\to\infty}\left(\|\nabla u_j\|^2-\|\nabla(u_j-(\lambda_j)^{-\frac{N-2}{2}}\phi((\cdot-a_j)/\lambda_j))\|^2\right)=\|\nabla\phi\|^2\geq\text{const }A^2(\varepsilon/A)^{d^2/4}\,, \tag{3.15}$$

$$\limsup_{j\to\infty}\|u_j-(\lambda_j)^{-\frac{N-2}{2}}\phi((\cdot-a_j)/\lambda_j)\|_q^q\leq\varepsilon^{2N/(N-2)}\left(1-\text{const }(\varepsilon/A)^{d(d+2)/2}\right)\,. \tag{3.16}$$

Proof. Since the argument is rather similar to that of Corollary 3.2 and Theorem 3.4 we only sketch the main steps. First of all, by passing to a subsequence, we may assume that $\lim_{j\to\infty}\|u_j\|_q=\varepsilon$. (This will only be important in ... below.) Then, as in the proof of Corollary 3.2,

$$\liminf_{j\to\infty}\left(\sup_{t>0}t^{(N-2)/4}\|e^{t\Delta}u_j\|_\infty\right)^{\frac{2}{N}}\geq C_N^{-1}A^{-\frac{N-2}{2N}}\varepsilon$$

and therefore there are $t_j>0$ and $x_j\in\mathbb{R}^N$ such that $w_j(y):=t_j^{(N-2)/4}u_j(\sqrt{t_j}y+x_j)$ satisfies

$$\left|\int_{\mathbb{R}^N}G(y)w_j(y)\,dy\right|\geq\frac12 C_N^{-\frac{N}{2}}A^{-\frac{N-2}{4}}\varepsilon$$

where $G(y)=(4\pi)^{-N/2}e^{-y^2/4}$. Since $G\in L_{q'}(\mathbb{R}^N)$ the weak limit in $\dot H^1(\mathbb{R}^N)$ of a subsequence of the (w_j) satisfies $\|\nabla w\|^2\geq S_N\|w\|_q^2\geq\text{const }A^{-\frac{N-2}{4}}\varepsilon$. This bound together with the weak convergence property implies (3.15). Once again we use Rellich–Kondrachov (see, e.g., [LiLo, Thms. 8.6 and 8.7]) to obtain a.e. convergence, and then the Brezis–Lieb lemma 3.3 implies

$$\limsup_{j\to\infty}\|t_j^{(N-2)/4}u_j(\sqrt{t_j}\cdot+x_j)-\phi\|_q^q=\varepsilon^q-\|\phi\|_q^q\,,$$

which yields (3.16) and completes the sketch of the proof. ∎

4. The sharp Sobolev inequality

Our goal in this section is to prove the following

Theorem 4.1 (Sharp Sobolev inequality). *For all $u \in \dot H^1(\mathbb{R}^N)$, $N \geq 3$, one has*

$$\int_{\mathbb{R}^N} |\nabla u|^2\,dx \geq \pi N(N-2)\left(\frac{\Gamma(N/2)}{\Gamma(N)}\right)^{2/N}\left(\int_{\mathbb{R}^N} |u|^{2N/(N-2)}\,dx\right)^{(N-2)/N}, \qquad (4.1)$$

with equality if and only if $u(x) = c\,h(b(x-a))$ for some $a \in \mathbb{R}^N$, $b > 0$ and $c \in \mathbb{C}$, where

$$h(x) = (1+x^2)^{-(N-2)/2}\,. \qquad (4.2)$$

In the remainder of this section we will write $q = 2N/(N-2)$.

`Aubin, Talenti, Rosen, Bliss; check whether they prove existence`

Proof of Theorem 4.1. We know from Theorem 3.4 that there is an optimizer U for inequality (4.1).

As a preliminary remark we note that if $u = a + ib$ with a and b real functions, then $\int |\nabla u|^2\,dx = \int (|\nabla a|^2 + |\nabla b|^2)\,dx$. We also note that the right side of (4.1) is $\|a^2+b^2\|_{q/2}$ with $q = 2N/(N-2) > 2$. By the triangle inequality, $\|a^2+b^2\|_{q/2} \leq \|a^2\|_{q/2} + \|b^2\|_{q/2}$. Therefore, if $U = a+ib$ is an optimizer for (4.1), then either one of a and b is identically equal to zero, or else both a and b are optimizers. Hence in any case, we may assume the optimizer U to be real. We may also assume $U \geq 0$ because for any $u \in \dot H^1(\mathbb{R}^N)$, $\partial|u|/\partial x_k = (\operatorname{sgn} u)\partial u/\partial x_k$ in the sense of distributions. (This can be proved similarly to [LiLo, Thm. 6.17].)

It is important for us to know that we may confine our search for optimizers to functions u satisfying a certain 'center of mass condition' with respect to $N+1$ functions $\mathcal{S}_j$, defined by

$$\mathcal{S}_j(x) = \frac{2x_j}{1+x^2} \quad \text{for } j = 1,\ldots,N\,, \qquad \mathcal{S}_{N+1}(x) = \frac{1-x^2}{1+x^2}\,.$$

Lemma 4.2 (Center of mass). *Let $u \in \dot H^1(\mathbb{R}^N)$. Then there is a $v \in \dot H^1(\mathbb{R}^N)$ with*

$$\|\nabla v\| = \|\nabla u\|\,, \qquad \|v\|_q = \|u\|_q$$

and

$$\int_{\mathbb{R}^N} \mathcal{S}_j(x)\ |v(x)|^q\,dx = 0\,, \qquad j = 1,\ldots,N+1\,. \qquad (4.3)$$

Moreover, if u is non-negative, then v is so as well.

We shall prove this lemma at the end of this section. It allows us to assume, without loss of generality, that our optimizer U satisfies the conditions (4.3). Now we shall prove that *the only non-negative optimizers satisfying* (4.3) *are non-negative multiples of the function h*. In other words, the center of mass condition breaks the large symmetry group and leads to uniqueness.

Since U is an optimizer, the second variation of the minimization quotient must be non-negative, i.e.,

$$\frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} \frac{\int_{\mathbb{R}^N} |\nabla(U+\varepsilon v)|^2\,dx}{\left(\int_{\mathbb{R}^N} |U+\varepsilon v|^{2N/(N-2)}\,dx\right)^{(N-2)/N}} \geq 0$$

for all $v \in \dot H^1(\mathbb{R}^N)$. After a quick computation, we deduce that

$$\int_{\mathbb{R}^N} |\nabla v|^2\,dx \int_{\mathbb{R}^N} U^q\,dx - (q-1)\int_{\mathbb{R}^N} |\nabla U|^2\,dx \int_{\mathbb{R}^N} U^{q-2}|v|^2\,dx \geq 0 \tag{4.4}$$

for all v with $\int U^{q-1} v\,dx = 0$.

Because U satisfies condition (4.3) we may choose $v(x) = \mathcal{S}_j(x)U(x)$ in (4.4) and sum over j. Using the fact that

$$\sum_{j=1}^{N+1} |\mathcal{S}_j(x)|^2 = 1 \tag{4.5}$$

we find that

$$\sum_{j=1}^{N+1} \int_{\mathbb{R}^N} |\nabla(\mathcal{S}_j U)|^2\,dx - (q-1)\ \int_{\mathbb{R}^N} |\nabla U|^2\,dx \geq 0\,. \tag{4.6}$$

On the other hand, an integration by parts (similarly as in the proof of Lemma 1.5) leads to

$$\int_{\mathbb{R}^N} |\nabla(\mathcal{S}_j U)|^2\,dx = \int_{\mathbb{R}^N} \left((\mathcal{S}_j)^2|\nabla U|^2 - \mathcal{S}_j \Delta \mathcal{S}_j U^2\right)\,dx\,.$$

Computing $\Delta\mathcal{S}_j$ and summing over j yields that

$$\sum_{j=1}^{N+1} \int_{\mathbb{R}^N} |\nabla(\mathcal{S}_j U)|^2\,dx = \int_{\mathbb{R}^N} |\nabla U|^2\,dx + N\int_{\mathbb{R}^N} \left(\frac{2}{1+x^2}\right)^2 U^2\,dx\,,$$

which, together with (4.6), implies that

$$N\int_{\mathbb{R}^N} \left(\frac{2}{1+x^2}\right)^2 U^2\,dx - (q-2)\ \int_{\mathbb{R}^N} |\nabla U|^2\,dx \geq 0\,.$$

Recalling that $q-2 = \frac{4}{N-2}$, we see that this is the same as

$$\int_{\mathbb{R}^N} |\nabla U|^2\,dx - N(N-2)\int_{\mathbb{R}^N} \frac{2}{1+x^2} U^2\,dx \leq 0\,.$$

This is the *reverse* inequality of what we have shown in Section 1.2. The arguments there imply that the left side can be written as

$$\int_{\mathbb{R}^N} |\nabla(U/\omega)|^2\omega^2\,dx \quad \text{with } \omega(x) = (1+x^2)^{-(N-2)/2}\,.$$

We conclude that U is proportional to ω, as we intended to prove. $\square$

We now turn to the proof of Lemma 4.2. The first observation is that the map $x \mapsto (\mathcal{S}_1(x), \ldots, \mathcal{S}_{N+1}(x))$ is the stereographic projection from $\mathbb{R}^N$ to $\mathbb{S}^{N+1} = \{\omega \in \mathbb{R}^{N+1} : \sum_{j=1}^{N+1} \omega_j^2 = 1\}$. (The fact that the range of this map is contained in the sphere was already used in (4.5).) The Jacobi matrix $DS(x) = (\partial S_j / \partial x_k)_{j,k}$ is an $N \times (N+1)$-matrix, and a tedious but straightforward computation shows that

$$DS(x)^T DS(x) = \left(\frac{2}{1+x^2} \right)^2 (1 - |\mathcal{S}(x)\rangle\langle\mathcal{S}(x)|) \,,$$

where 1 stands for the $(N+1) \times (N+1)$ identity matrix. From this formula we conclude that the Jacobian of $\mathcal{S}$ is

$$\mathcal{J}(x) = \left(\frac{2}{1+x^2} \right)^N \,,$$

and therefore, if $u \in L_q(\mathbb{R}^N)$ and $f \in L_q(\mathbb{S}^N)$ are related by

$$u(x) = \mathcal{J}(x)^{1/q} f(\mathcal{S}(x)) \,, \tag{4.7}$$

then

$$\|u\|_q = \|f\|_q \,. \tag{4.8}$$

Moreover, the formula for $DS(x)^T DS(x)$ together with some straightforward computations yields

$$\int_{\mathbb{R}^N} |\nabla u|^2 \, dx = \int_{\mathbb{S}^N} \left(|\nabla f|^2 + \frac{N(N-2)}{4} |f|^2 \right) d\omega \,. \tag{4.9}$$

The upshot of this discussion is that we have discovered a large group of symmetries of the Sobolev inequality: apart from the translations, dilations and rotation, which are obvious in the $\mathbb{R}^N$ version, we infer that the inequality is also invariant under rotations of the sphere $\mathbb{S}^N$! This will be the key in proving Lemma 4.2.

We shall define a family of maps $\gamma_{\delta,\xi} : \mathbb{S}^N \to \mathbb{S}^N$ depending on two parameters $\delta > 0$ and $\xi \in \mathbb{S}^N$. To do so, we denote dilation on $\mathbb{R}^N$ by $\mathcal{D}_\delta$, that is, $\mathcal{D}_\delta(x) = \delta x$. Moreover, for any $\xi \in \mathbb{S}^N$ we choose an orthogonal $(N+1) \times (N+1)$ matrix O such that $O\xi = (0, \ldots, 0, 1)$ and we put

$$\gamma_{\delta,\xi}(\omega) := O^T \mathcal{S} \left(\mathcal{D}_\delta \left(\mathcal{S}^{-1} (O\omega) \right) \right)$$

for all $\omega \in \mathbb{S}^N \setminus \{-\xi\}$ and $\gamma_{\delta,\xi}(-\xi) := -\xi$. This transformation depends only on ξ (and δ) and not on the particular choice of O. Indeed, a straightforward computation shows that

$$\gamma_{\delta,\xi}(\omega) = \frac{2\delta}{(1 + \omega \cdot \xi) + \delta^2 (1 - \omega \cdot \xi)} \; (\omega - (\omega \cdot \xi) \; \xi) + \frac{(1 + \omega \cdot \xi) - \delta^2 (1 - \omega \cdot \xi)}{(1 + \omega \cdot \xi) + \delta^2 (1 - \omega \cdot \xi)} \; \xi \,.$$

Lemma 4.3. *Let $\rho \in L^1(\mathbb{S}^N)$ with $\int_{\mathbb{S}^N} \rho(\omega) \, d\omega \neq 0$. Then there is a transformation $\gamma_{\delta,\xi}$ of $\mathbb{S}^N$ such that*

$$\int_{\mathbb{S}^N} \gamma_{\delta,\xi}(\omega) \rho(\omega) \, d\omega = 0 \,.$$

In order to derive Lemma 4.2 from Lemma 4.3 we assume that u on $\mathbb{R}^N$ and f on $\mathbb{S}^N$ are related by (4.7). Applying Lemma 4.3 to $\rho = |f|^q$ we obtain a $\gamma = \gamma_{\delta,\xi}$ and then a change of variables shows that the new function $\sigma(\omega) = |\mathcal{J}_{\gamma^{-1}}(\omega)|\rho(\gamma^{-1}(\omega))$ satisfies

$$\int_{\mathbb{S}^N} \omega\sigma(\omega)\,d\omega = 0\,.$$

Finally, we write $\sigma = g^q$ and, if v on $\mathbb{R}^N$ corresponds to g via (4.7), then the center of mass condition for g^q is equivalent to the condition (4.3) for v. Both the L_q and the gradient norm are invariant under this procedure, since γ is a composition of a dilation, a rotation and the stereographic projection, all of which preserve both norms; (cf. (4.8) and (4.9) for the stereographic projection.) Obviously, non-negativity of functions is preserved.

Proof. We may assume that $\rho \in L^1(\mathbb{S}^N)$ is normalized by $\int_{\mathbb{S}^N} \rho(\omega)\,d\omega = 1$. We shall show that the $\mathbb{R}^{N+1}$-valued function

$$F(r\xi) := \int_{\mathbb{S}^N} \gamma_{1-r,\xi}(\omega)\rho(\omega)\,d\omega\,, \qquad 0 < r < 1\,,\ \xi \in \mathbb{S}^N\,,$$

has a zero. First, note that because of $\gamma_{1,\xi}(\omega) = \omega$ for all ξ and all ω, the limit of $F(r\xi)$ as $r \to 0$ is independent of ξ. In other words, F is a continuous function on the open unit ball of $\mathbb{R}^{N+1}$. In order to understand its boundary behavior, one easily checks that for any $\omega \neq -\xi$ one has $\lim_{\delta\to 0}\gamma_{\delta,\xi}(\omega) = \xi$, and that this convergence is uniform on $\{(\omega,\xi) \in \mathbb{S}^N \times \mathbb{S}^N :\ 1 + \omega\cdot\xi \geq \varepsilon\}$ for any $\varepsilon > 0$. This implies that

$$\lim_{r\to 1} F(r\xi) = \xi \qquad \text{uniformly in } \xi\,. \tag{4.10}$$

Hence F is a continuous function on the *closed* unit ball, which is the identity on the boundary.

We claim that there is an $\varepsilon > 0$ such that

$$|x - \varepsilon F(x)| \leq 1 \qquad \text{for all } x \in B\,. \tag{4.11}$$

Once this is shown, Brouwer's fixed point theorem `Reference` applied to $x - \varepsilon F(x)$ implies that there is an x_0 such that $x_0 - \varepsilon F(x_0) = x_0$, that is, $F(x_0) = 0$, as claimed.

In order to prove (4.11), we note that by (4.10) there is an r_0 such that

$$x \cdot F(x) \geq 1/2 \qquad \text{provided } |x| \geq r_0\,.$$

Hence, using that $|x| \leq 1$ and that $|F(x)| \leq \|\rho\|_1$ we see that

$$|x - \varepsilon F(x)|^2 \leq \varepsilon^2\|\rho\|_1^2 - 2\varepsilon x\cdot F(x) + 1 \leq \varepsilon^2\|\rho\|_1^2 - \varepsilon + 1$$

for $|x| \geq r_0$. Hence there is an $\varepsilon_0 > 0$ such that (4.11) is true for any $|x| \geq r_0$ and any $0 < \varepsilon \leq \varepsilon_0$. Decreasing ε_0 if necessary such that $\varepsilon \leq (1 - r_0)\|\rho\|_1^{-1}$, we have for any $|x| \leq r_0$

$$|x - \varepsilon F(x)| \leq r_0 + \varepsilon\|\rho\|_1 \leq 1\,,$$

and hence (4.11) is true for all $x \in B$. This concludes the proof of Lemma 4.3. □

5. Schwarz symmetrization

On physical grounds it is very natural to expect that solutions of certain rotation-invariant minimization problems will be radially symmetric. One effective way of verifying this expectation is a symmetrization procedure called *Schwarz symmetrization*, which we introduce and study in this section. In particular, we shall see that this procedure reduces the kinetic energy of a quantum mechanical (single-particle) state and we shall derive the sharp version of the isoperimetric inequality.

5.1. **Definition and properties.** Let $\Omega \subset \mathbb{R}^N$ be a set with finite measure. We define its Schwarz symmetrization Ω^* to be the ball in $\mathbb{R}^N$ centered at the origin with the same measure as Ω, i.e.,

$$\Omega^* = \{x \in \mathbb{R}^N : \ |x| < \omega_N^{-1/N} |\Omega|^{1/N}\} ,$$

where $\omega_N = |\{x \in \mathbb{R}^N : \ |x| < 1\}|$ is the measure of the unit ball. The Schwarz symmetrization of a measurable function f on $\mathbb{R}^N$ satisfying

$$|\{|f| > \mu\}| < \infty \qquad \text{for all } \mu > 0 \tag{5.1}$$

is defined by

$$f^*(x) = \int_0^\infty \chi_{\{|f|>\tau\}^*}(x)\, d\tau \, . \tag{5.2}$$

This formula should be compared to the 'layer cake representation' of $|f|$,

$$|f(x)| = \int_0^\infty \chi_{\{|f|>\tau\}}(x)\, d\tau \, . \tag{5.3}$$

By definition, f^* is a non-negative, radially symmetric and non-increasing function. It is a rearrangement of f in the sense that $|\{|f^*| > \mu\}| = |\{|f| > \mu\}|$ for every $\mu > 0$. This implies, in particular, that

$$\int_{\mathbb{R}^N} |f^*|^p \, dx = \int_{\mathbb{R}^N} |f|^p \, dx \tag{5.4}$$

for any $p > 0$. Here is another simple property of rearrangement.

Lemma 5.1. *For any non-negative functions f and g on $\mathbb{R}^N$ satisfying* (5.1)*,*

$$\int_{\mathbb{R}^N} fg \, dx \leq \int_{\mathbb{R}^N} f^* g^* \, dx \, . \tag{5.5}$$

Proof. Using (5.3), (5.2) and similar formulas for g we see that it suffices to prove that

$$|A \cap B| = \int_{\mathbb{R}^N} \chi_A(x)\chi_B(x)\, dx \leq \int_{\mathbb{R}^N} \chi_{A^*}(x)\chi_{B^*}(x)\, dx = |A^* \cap B^*| \, ,$$

where $A = \{f > \tau\}$ and $B = \{g > \sigma\}$. This is obvious, however, since $|A \cap B| \leq \min\{|A|, |B|\} = \min\{|A^*|, |B^*|\} = |A^* \cap B^*|$. □

The next result also compares integrals of functions before and after rearrangement. The result, however, is much deeper and its proof is rather involved. The case of one dimension is due to Riesz [Ri] (see also `insert` for a version about sequences, `higher dimensions`). As we shall see, it this lemma has far reaching consequences.

Lemma 5.2 (Riesz lemma)**.** *For any non-negative functions f, g and h on $\mathbb{R}^N$ satisfying* (5.1),

$$\iint_{\mathbb{R}^N\times\mathbb{R}^N} f(x)g(x-y)h(y)\,dx\,dy \le \iint_{\mathbb{R}^N\times\mathbb{R}^N} f^*(x)g^*(x-y)h^*(y)\,dx\,dy\,. \tag{5.6}$$

Proof. Since the proof is rather involved we only sketch the one-dimensional case and refer the reader to [LiLo, Thm. 3.7] for the argument of how to reduce the multi-dimensional case to the one-dimensional one.

As in the proof of Lemma 5.1 we may assume that f, g and h are characteristic functions of sets of finite measure. In order to explain the main idea of the proof we assume that f and g are characteristic functions of intervals centered at the origin and h is the characteristic function of two disjoint intervals. That is, h is of the form

$$h(x) = h_1(x-a_1) + h_2(x-a_2)$$

where h_1 and h_2 are again characteristic functions of intervals centered at the origin and a_1 and a_2 are two real numbers.

We want to show that the left side of (5.6) does not decrease if the two intervals are joint. In order to move the intervals together, we introduce a paramater $0\le t\le 1$ and set

$$I_j(t) = \int_{\mathbb{R}^N} k(y)h_j(y-ta_j)\,dy\,, \qquad k(x) = \int_{\mathbb{R}^N} f(x)g(x-y)\,dx\,.$$

In this way, the left side of (5.6) coincides with $I_1(1)+I_2(1)$. The function k is symmetric decreasing, and therefore a similar argument as in the proof of Lemma 5.1 shows that $I_j(t)$ is non-decreasing as t varies from 1 to 0. This proves that $I_1(t)+I_2(t)$ is non-decreasing as t varies from 1 to 0. As t decreases the two intervals that make up h move together. We stop this movement when their endpoints touch. In this way we have obtained one single interval, and again by Lemma 5.1 this single interval wants to be centered at the origin, as claimed.

This explains the main idea of the proof. In order to treat the general case one uses some approximation arguments to put oneself in the situation where each of f, g and h is the characteristic function of disjoint intervals. Repeated use of the argument we just explained then proves the claim. □

This lemma has been substantially generalized by Brascamp–Lieb–Luttinger [BLL]. Instead of three function f, g and h of two variables x and y, they consider m functions of k variables. While the Riesz lemma will allow us to control single eigenvalues, the BLL extension will allow us to control (exponential) *sums* of eigenvalues.

Lemma 5.3 (BLL lemma)**.** *For any non-negative functions $f_1, \ldots, f_m$ on $\mathbb{R}^N$ satisfying* (5.1) *and any $k \times m$ matrix (b_{ij}) with $k \leq m$,*

$$\int_{\mathbb{R}^N} \cdots \int_{\mathbb{R}^N} \prod_{j=1}^{m} f_j(\sum_{i=1}^{k} b_{ij} x_i)\, dx_1 \ldots\, dx_k \leq \int_{\mathbb{R}^N} \cdots \int_{\mathbb{R}^N} \prod_{j=1}^{m} f_j^*(\sum_{i=1}^{k} b_{ij} x_i)\, dx_1 \ldots\, dx_k\,.$$

For the proof we refer to the original paper [BLL].

5.2. **The Pólya–Szegő principle.** The main reason why Schwarz symmetrization is useful for us is that it decreases (strictly speaking: does not increase) the kinetic energy. In many variational problems this allows one to restrict one's attention to radial function, which makes the problem one-dimensional and therefore considerably easier.

This property of Schwarz symmetrization was popularized in [PoSz] and is often called the *Pólya–Szegő principle*, although the result actually is older `Check Hardy--Littlewood-Polya`. The classical proofs rely on the isoperimetric inequality and delicate regularity considerations. The proof below, which is based on the Riesz lemma and symmetry/monotonicity properties of the heat kernel, is due to Lieb [Li].

Proposition 5.4. *For any u with $|\{|u| > \mu\}| < \infty$ for all $\mu > 0$ and $\int |\nabla u|^2\, dx < \infty$ one has*

$$\int_{\mathbb{R}^N} |\nabla u|^2\, dx \geq \int_{\mathbb{R}^N} |\nabla u^*|^2\, dx\,. \tag{5.7}$$

Inequality (5.7) is, in general, not strict for $u \not\equiv u^*$. For counterexamples and a characterization of the cases of equality we refer to [BrZi].

Proof. We recall formula (3.1) for the heat kernel. Using its representation (3.2) in Fourier space it is easy to prove that for any $u \in H^1(\mathbb{R}^N)$

$$\lim_{t \to 0} t^{-1} \left(u, \left(1 - e^{t\Delta}\right) u\right) = \int_{\mathbb{R}^N} |\nabla u|^2\, dx\,.$$

Since the heat kernel is a symmetric decreasing function, the Riesz lemma together with (5.4) implies that for any $t > 0$

$$\left(u, \left(1 - e^{t\Delta}\right) u\right) \geq \left(u^*, \left(1 - e^{t\Delta}\right) u^*\right)\,.$$

We conclude that for $u \in H^1(\mathbb{R}^N)$ one has

$$\limsup_{t \to 0} t^{-1} \left(u^*, \left(1 - e^{t\Delta}\right) u^*\right) \leq \int_{\mathbb{R}^N} |\nabla u|^2\, dx\,.$$

By monotone converges in Fourier space, one deduces that $u^* \in H^1(\mathbb{R}^N)$ and that its gradient satisfies (5.7).

This proves the inequality for $u \in H^1(\mathbb{R}^N)$. In the general case, we note that we can restrict our attention to non-negative u (since the rearrangement only depends on $|u|$ and the gradients of u and $|u|$ coincide in magnitude as distributions). The function

$u_\tau(x) := \min\{(u(x)-\tau)_+, \tau^{-1}\}$ (compare with the regularization in the proof of Theorem 3.1) belongs to $H^1(\mathbb{R}^N)$ and therefore the H^1-result together with monotone convergence conclude the proof. □

Corollary 5.5. $\lambda_1(-\Delta+V) \geq \lambda_1(-\Delta-(V_-)^*)$

Indeed, one has $\int |\psi|^2\,dx = \int |\psi^*|^2\,dx$ by (5.4) and

$$\int_{\mathbb{R}^N} \left(|\nabla\psi|^2 + V|\psi|^2\right) dx \geq \int_{\mathbb{R}^N} \left(|\nabla\psi^*|^2 - (V_-)^*|\psi|^2\right) dx$$

by (5.7) and (5.5).

Remark 5.6. Proposition 5.4 can be used to give an alternative proof of the sharp Sobolev inequality. Indeed, in view of (5.7) and (5.4) it suffices to consider radial functions in the minimization problem for S_N, i.e.,

$$S_N = |\mathbb{S}^{N-1}|^{2/N} \ \inf \frac{\int_0^\infty |u'|^2 r^{N-1}\,dr}{\left(\int_0^\infty |u|^{2N/(N-2)} r^{N-1}\,dr\right)^{(N-2)/N}} \,.$$

Now we introduce logarithmic coordinates and define for each u on $(0,\infty)$ a function v on $\mathbb{R}$ by

$$u(r) = r^{-(N-2)/2}\, v(\ln r)\,.$$

A quick computation shows that

$$\int_0^\infty |u|^{2N/(N-2)} r^{N-1}\,dr = \int_{\mathbb{R}} |v|^{2N/(N-2)}\,dx$$

and

$$\int_0^\infty |u'|^2 r^{N-1}\,dr = \int_{\mathbb{R}} \left(|v'|^2 + \left(\frac{N-2}{2}\right)^2 |v|^2\right) dx\,.$$

Using scaling it is an easy exercise to express the value of the infimum

$$\inf \frac{\int_{\mathbb{R}} \left(|v'|^2 + \left(\frac{N-2}{2}\right)^2 |v|^2\right) dx}{\int_{\mathbb{R}} |v|^{2N/(N-2)}\,dx}$$

in terms of the value $S_{1,2N/(N-2)}$ from Proposition 2.5. In this way we obtain the sharp constant in Theorem 4.1 and we also obtain the functions $(1+x^2)^{-(N-2)/2}$. (Note that these correspond to a negative power of cosh in logarithmic coordinates.) Since (5.7) is not strict, in general, this argument does *not* imply that these are the only functions (up to translations and dilations) on which the infimum is attained.

Exercise 5.7. Carry out the steps in the previous remark. More generally, find the sharp constant in the inequality

$$\int_{\mathbb{R}^N} |\nabla u|^2\,dx \geq K_{N,b} \left(\int_{\mathbb{R}^N} \frac{|u|^q}{|x|^{bq}}\,dx\right)^{2/q}, \qquad 0 \leq b < 1\,,\ q = \frac{2N}{N-2+2b}\,,$$

which interpolates between the Sobolev inequality and the Hardy inequality. Furthermore, show that this inequality remains valid in the range $-(N-2)/2 < b < 0$ when restricted to radial functions, and compute the corresponding optimal constant. An early reference for this inequality is [Ok]. The sharp constant and the extension to negative values of b appears in [GlMaGrTh].

Remark 5.8. While the Sobolev inequality seems to be stronger than Hardy's inequality (at least its proof is considerably more involved), there is a quick way to deduce the Sobolev inequality from the Hardy inequality via symmetrization. Indeed, if $u = u^*$ is a symmetric decreasing function on $\mathbb{R}^N$, then for any $q > 0$ and any $R > 0$

$$\int_{\mathbb{R}^N} |u|^q \, dx \geq \int_{|x|<R} |u|^q \, dx \geq N^{-1} |\mathbb{S}^{N-1}| \; |u(R)|^q \; R^N \,, \tag{5.8}$$

where, of course, $u(R)$ denotes the common value of $u(x)$ when $|x| = R$. Thus $|u(x)| \leq N^{1/q} |\mathbb{S}^{N-1}|^{-1/q} |x|^{-N/q} \|u\|_q$. We now specialize to the case $q = 2N/(N-2)$ and raise the previous inequality to the power $q - 2 = 4/(N-2)$ to obtain

$$\int_{\mathbb{R}^N} |u|^{2N/(N-2)} \, dx \leq |u(x)| \leq N^{2/N} |\mathbb{S}^{N-1}|^{-2/N} \|u\|_{2N/(N-2)}^{4/(N-2)} \int_{\mathbb{R}^N} \frac{|u|^2}{|x|^2} \, dx \,.$$

Bounding the right side by Hardy's inequality,

$$\int_{\mathbb{R}^N} |u|^{2N/(N-2)} \, dx \leq |u(x)| \leq \frac{(N-2)^2}{4} N^{2/N} |\mathbb{S}^{N-1}|^{-2/N} \|u\|_{2N/(N-2)}^{4/(N-2)} \int_{\mathbb{R}^N} |\nabla u|^2 \, dx \,,$$

we yields the Sobolev inequality for symmeric decreasing functions. By Proposition 5.4 and (5.4), this implies the inequality for arbitrary functions.

Exercise 5.9. In this exercise we derive an improved Sobolev inequality by being less crude in estimate (5.8).

(1) Find a formula for $\int |x|^{-2} (u^*)^2 \, dx$ in $N \geq 3$ in terms of the measure of the sets $\{|u| > \mu\}$ and use Hardy's inequality (Theorem 1.4) to conclude that

$$\int_{\mathbb{R}^N} |\nabla u|^2 \, dx \geq \left(\frac{N-2}{2} \right)^2 \left(\frac{|\mathbb{S}^{N-1}|}{N} \right)^{2/N} \frac{2N}{N-2} \int_0^\infty |\{|u| > \mu\}|^{\frac{N-2}{N}} \mu \, d\mu \,.$$

(2) Here left side is a constant times the square of the norm of u in the Lorentz space $L_{2N/(N-2),2}(\mathbb{R}^N)$. Show that this space is contained in the usual Lebesgue space $L_{2N/(N-2)}(\mathbb{R}^N)$. Give an example of a function in $L_{2N/(N-2),2}(\mathbb{R}^N)$, which does not lie in $L_{2N/(N-2)}(\mathbb{R}^N)$.

(3) Use the inequality from Exercise 5.7 to prove a similar inequality. (The corresponding space will be the Lorentz space $L_{Nq/(N-2),q}(\mathbb{R}^N)$ with $q = \frac{2N}{N-2+2b}$.)

`Reference to Alvino + interpolation people`

5.3. **The isoperimetric inequality.** The isoperimetric inequality in the plane states that among all domains with the same length of the boundary, the ball has the largest area. Equivalently, among all domains of the same area, the boundary of the ball has the shortest length. The following is the multi-dimensional generalization of this result.

Theorem 5.10. *For any sufficiently smooth domain $\Omega \subset \mathbb{R}^N$ of finite measure one has $|\partial\Omega| \geq |\partial\Omega^*|$.*

Of course, $|\partial\Omega|$ stands for the $(N-1)$-dimensional measure of the boundary. This is certainly well-defined for Lipschitz boundaries. The perimeter in the sense of de Giorgi can also be defined for domains with very low regularity, but we do not want to enter this topic.

`Sharp version due to Federer--Fleming and Fleming--Rishel[6] -- they really prove the isoperimetric inequality and then deduce Sobolev`

We have already remarked that the isoperimetric inequality is closely related to the Sobolev inequality $\int |\nabla u|\,dx \geq C_N \left(\int |u|^{N/(N-1)}\,dx\right)^{(N-1)/N}$. Therefore it is not surprising that it can be proved via a rearrangement inequality for $\int |\nabla u|\,dx$, similarly to that of Proposition 5.7. Again, the main tool will be the Riesz lemma 5.2.

The other ingredient is a study of the limit $(\chi_{\Omega^c}, \exp(t\Delta)\chi_\Omega)$ as $t \to 0$. Lemma 5.11 gives an upper bound on this quantity, which Lemma 5.12 proves to be sharp in the case of balls.

Lemma 5.11. *For any $f : \mathbb{R}^N \to \mathbb{R}$ with $\nabla f \in L_1$, one has*

$$\left\|(\exp(t\Delta)f - f)_+\right\|_1 \leq \left(\frac{t}{\pi}\right)^{1/2} \int_{\mathbb{R}^N} |\nabla f|\,dx\,.$$

Lemma 5.12. *If $B \subset \mathbb{R}^N$ is a ball, then*

$$\lim_{t\to 0} t^{-1/2}(\chi_{B^c}, \exp(t\Delta)\chi_B) = \pi^{-1/2}|\partial B|\,.$$

Assuming these two lemmas for the moment, we follow [Le2] for the

Proof of Theorem 5.10. Let Ω by an open set of finite measure with sufficiently smooth boundary. Then by Lemma 5.11

$$(\chi_{\Omega^c}, \exp(t\Delta)\chi_\Omega) = (\chi_{\Omega^c}, \exp(t\Delta)\chi_\Omega - \chi_\Omega) \leq \left\|(\exp(t\Delta)\chi_\Omega - \chi_\Omega)_+\right\|_1 \leq \left(\frac{t}{\pi}\right)^{1/2} |\partial\Omega|\,.$$

On the other hand, by Lemma 5.2

$$\begin{aligned}(\chi_{(\Omega^*)^c}, \exp(t\Delta)\chi_{\Omega^*}) &= |\Omega^*| - (\chi_{\Omega^*}, \exp(t\Delta)\chi_{\Omega^*}) \\ &\leq |\Omega| - (\chi_\Omega, \exp(t\Delta)\chi_\Omega) = (\chi_{\Omega^c}, \exp(t\Delta)\chi_\Omega)\end{aligned}$$

Dividing by $t^{1/2}$ and letting $t \to 0$, Lemma 5.12 yields the assertion. □

It remains to prove the two lemmas.

Proof of Lemma 5.11. Since $\exp(t\Delta)f - f = \int_0^t \Delta \exp(s\Delta) f\,ds$, we have for any g

$$(g, \exp(t\Delta)f - f) = -\int_0^t (\nabla \exp(s\Delta)g, \nabla f)\,ds\,.$$

Using the fact that

$$(\exp(s\Delta)g)(x) = \int_{\mathbb{R}^N} g(x+\sqrt{2s}y) \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}}$$

we find

$$(\nabla \exp(s\Delta)g)(x) = (2s)^{-1/2} \int_{\mathbb{R}^N} yg(x+\sqrt{2s}y) \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}}$$

and hence

$$(g, \exp(t\Delta)f - f) = -\int_0^t ds \int_{\mathbb{R}^N} dx \int_{\mathbb{R}^N} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}} \, (2s)^{-1/2} \overline{g(x+\sqrt{2s}y)}\, y \cdot \nabla f(x) \, .$$

In particular, if $0 \leq g \in L_\infty(\mathbb{R}^N)$ we can bound

$$(g, \exp(t\Delta)f - f) \leq \|g\|_\infty \int_0^t ds \int_{\mathbb{R}^N} dx \int_{\mathbb{R}^N} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}} \, (2s)^{-1/2} \, (y \cdot \nabla f(x))_- \, .$$

For fixed x and s, we perform the y-integration and arrive at

$$(g, \exp(t\Delta)f - f) \leq \|g\|_\infty \int_0^t ds \int_{\mathbb{R}^N} dx \, (4\pi s)^{-1/2} \, |\nabla f(x)| = \|g\|_\infty \sqrt{\frac{t}{\pi}} \int_{\mathbb{R}^N} dx \, |\nabla f(x)| \, .$$

By duality, this implies the assertion. □

Proof of Lemma 5.12. By translation and dilation invariance, we may assume that B is centered at the origin and has radius one. We have

$$\begin{aligned}
(\chi_{B^c}, \exp(t\Delta)\chi_B) &= \int_{|x|>1} dx \int_{\{y:\ |x+\sqrt{2t}y|\leq 1\}} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}} \\
&= \int_{\mathbb{S}^{N-1}} d\omega \int_1^\infty dr\, r^{N-1} \int_{\{y:\ |r\omega+\sqrt{2t}y|\leq 1\}} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}} \\
&= |\mathbb{S}^{N-1}| \int_1^\infty dr\, r^{N-1} \int_{\{y:\ 2t(y')^2+(r+\sqrt{2t}y_d)^2\leq 1\}} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}} \\
&= |\mathbb{S}^{N-1}| \int_{\mathbb{R}^N} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}} \, I_t(y)
\end{aligned}$$

where

$$I_t(y) = \chi_{\{2t(y')^2\leq 1\}} \ \chi_{\{\sqrt{2t}y_d \leq \sqrt{1-2t(y')^2}-1\}} \int_1^\infty dr\, r^{N-1} \, \chi_{\{(r+\sqrt{2t}y_d)^2 \leq 1-2t(y')^2\}} \, .$$

A quick computation shows that

$$\lim_{t\to 0} t^{-1/2} I_t(y) = \sqrt{2}(y_1)_- \, ,$$

and therefore, by dominated convergence,

$$\lim_{t\to 0} t^{-1/2} (\chi_{B^c}, \exp(t\Delta)\chi_B) = \sqrt{2}|\mathbb{S}^{N-1}| \int_{\mathbb{R}^N} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}} (y_1)_- = \pi^{-1/2}|\mathbb{S}^{N-1}| \, .$$

This completes the proof. □

Remark 5.13. Alternatively, one can check that the terms that were thrown away in the proof of Lemma 5.11, that is,

$$\int_0^t ds \int_{\mathbb{S}^{N-1}} d\omega \int_{\{y:|\omega+\sqrt{2s}y|>1\}} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}}\,(2s)^{-1/2}\,(y\cdot\omega)_+ + \int_0^t ds \int_{\mathbb{S}^{N-1}} d\omega \int_{\{y:|\omega+\sqrt{2s}y|\leq 1\}} \frac{e^{-y^2/2}\,dy}{(2\pi)^{N/2}}\,(2s)^{-1/2}\,(y\cdot\omega)_-\,,$$

tend to zero as $t\to 0$.

References

[BeLi] H. Berestycki, P.-L. Lions, *Nonlinear scalar field equations. I. Existence of a ground state*. Arch. Rational Mech. Anal. **82** (1983), no. 4, 313–345.

[BLL] H. J. Brascamp, E. H. Lieb, J. M. Luttinger, *A general rearrangement inequality for multiple integrals*. J. Funct. Anal. **17** (1974), 227–237.

[BrZi] J. Brother, W. P. Ziemer, *Minimal rearrangements of Sobolev functions*. J. Reine Angew. Math. **384** (1988), 153–179.

[GeMeOr] P. Gérard, Y. Meyer, F. Oru, *Inégalités de Sobolev précisées*, Séminaire EDP, Exp. No. IV, École Polytechnique, Palaiseau, 1997.

[GlMaGrTh] V. Glaser, A. Martin, H. Grosse, W.Thirring, *Afamily of optimal conditions for the absence of bound states in a potential*. In: Studies in Mathematical Physics, E. H. Lieb, B. Simon and A. S. Wightman (eds.), Princeton University Press (1976), 169–194.

[Fa] I. Fabbri, *Remarks on some weighted Sobolev inequalities and applications*. PhD Thesis (Università Roma Tre), 2005.

[Ke] J. B. Keller, *Lower bounds and isoperimetric inequalities for eigenvalues of the Schrödinger equation*. J. Math. Phys. **2** (1961), 262–266.

[Le1] M. Ledoux, *Semigroup proofs of the isoperimetric inequality in Euclidean and Gauss space*. Bull. Sci. Math. **118** (1994), 485–510.

[Le2] M. Ledoux, *On improved Sobolev embedding theorems*. Math. Res. Lett. **10** (2003), 659–669.

[Li] E. H. Lieb, *Existence and uniqueness of the minimizing solution of Choquard's nonlinear equation*. Studies in Appl. Math. **57** (1977), 93–105.

[LiLo] E. H. Lieb, M. Loss, *Analysis. Second edition*. Graduate Studies in Mathematics **14**, American Mathematical Society, Providence, RI, 2001.

[Lu] J. M. Luttinger, *Generalized isoperimetric inequalities*. J. Mathematical Phys. **14** (1973), 586–593.

[Ok] G. 0. Okikiolu, *Aspects of the theory of bounded integral operators in L^p-spaces*,M Academic Press, N.Y., 1971.

[PoSz] G. Pólya, G. Szegő, *Isoperimetric inequalities in mathematical physics*. Princeton University Press, Princeton, 1951.

[Ri] F. Riesz, *Sur une inégalité intégral*. J. London Math. Soc. **5** (1930), 162–168.

[Ru1] M. Rumin, *Spectral density and Sobolev inequalities for pure and mixed states*. Geom. Funct. Anal. **20** (2010), 817–844.

[Ru2] M. Rumin, *Balanced distribution-energy inequalities and related entropy bounds*. arXiv: 1008.16741v1.

[Ta] G. Talenti, *Bes constant in Sobolev inequality*. Ann. Mat. Pura Appl. **110** (1976), 353–372.

Rupert L. Frank, Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

E-mail address: rlfrank@math.princeton.edu