

## LECTURE NOTES FOR 247B, WINTER 2020

MOHIT BANSIL, ITAI BAR-NATAN, BOHAN CHEN, ERIN GEORGE, JAMES LENG,  
BERTRAND STONE, MARGARET WARD, MINGTAO XIA, HAOLING XIANG,  
AND JASON ZHAO

### 1. OSCILLATORY INTEGRALS

**Lemma 1.1.** *Let  $A$  and  $B$  be  $d \times d$  real symmetric matrices with  $A$  positive definite. Denoting  $Z = A + iB$ , we have*

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}x \cdot Zx} dx = \sqrt{\det(2\pi Z^{-1})}. \quad (1.1) \quad \boxed{1.\mathbf{gi}}$$

For  $B \neq 0$ , we define the square root via analytic continuation along the map  $s \mapsto A + isB$  for  $s \in [0, 1]$ .

*Remark.* Note  $Z$  is always invertible (so  $Z^{-1}$  has non-zero determinant) since

$$\operatorname{Re}(x \cdot Zx) = x \cdot Ax > 0$$

whenever  $x \in \mathbb{R}^d \setminus \{0\}$ .

We can equivalently define  $\sqrt{\det Z}$  by considering the eigenvalues of  $Z$ . Observe  $A$  and  $B$  are Hermitian with  $A$  positive definite also in the complex sense, so  $z^*Az > 0$  and  $z^*Bz \in \mathbb{R}$  for all  $z \in \mathbb{C}^d \setminus \{0\}$ . If  $\lambda_j$  are the eigenvalues of  $Z$ , repeated by algebraic multiplicity, then

$$\operatorname{Re}(\lambda_j)|z|^2 = \operatorname{Re}(z^*Zz) = \operatorname{Re}(z^*Az + iz^*Bz) = z^*Az > 0$$

for eigenvectors  $z \in \mathbb{C}^n$  of  $\lambda_j$ . Thus  $\operatorname{Re}\lambda_j > 0$ . Letting  $\sqrt{\lambda_j}$  be the root with positive real part, we have

$$\sqrt{\det Z} = \prod_j \sqrt{\lambda_j}.$$

As a side remark, there exist  $A$  and  $B$  such that  $Z$  is not diagonalizable, such as

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

*Proof of lemma.* We first prove the case  $B = 0$ , and conclude the general case via analytic continuation. Let  $\sqrt{A}$  be the positive definite square root of  $A$ , then making the change of variables  $y = \sqrt{A}x$  and noting  $x \cdot Ax = |\sqrt{A}x|^2$  gives

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}x \cdot Ax} dx = \frac{1}{\det \sqrt{A}} \int_{\mathbb{R}^d} e^{-\frac{1}{2}|y|^2} dy = \frac{(2\pi)^{d/2}}{\sqrt{\det A}} = \sqrt{\det(2\pi A^{-1})}.$$

For the general case, define  $Z(z) := A + izB$  and consider the analytic map

$$\Phi(z) := \int_{\mathbb{R}^d} e^{-\frac{1}{2}x \cdot Z(z)x} dx.$$

Observe  $\operatorname{Re} Z(z) = A - (\operatorname{Im} z)B$  is positive definite on the strip

$$|\operatorname{Im} z| < \frac{1}{\|B\| \|A^{-1}\|},$$

where  $\|\cdot\|$  is the usual matrix operator norm induced by the Euclidean metric. It follows from the first case that  $\Phi(z) = \sqrt{\det(2\pi Z(z)^{-1})}$  on the strip whenever  $\operatorname{Re} z = 0$ . By the uniqueness theorem, this result extends globally to the entire strip, particularly when  $z = 1$  which is exactly (1.1).  $\square$

**Lemma 1.2.** *Let  $A$  and  $B$  be  $d \times d$  real symmetric matrices with  $A$  positive definite. Denoting  $Z = A + iB$ , we have*

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}x \cdot Zx} e^{i\xi \cdot x} dx = e^{-\frac{1}{2}\xi \cdot Z^{-1}\xi} \sqrt{\det(2\pi Z^{-1})} \quad (1.2) \quad \boxed{1.\text{fi}}$$

for all  $\xi \in \mathbb{R}^d$ .

*Proof.* For brevity denote the left-hand and right-hand sides of (1.2) by  $L(\xi)$  and  $R(\xi)$  respectively. We claim that

$$\nabla L(\xi) = -Z^{-1}\xi L(\xi), \quad \nabla R(\xi) = -Z^{-1}\xi R(\xi).$$

The previous lemma furnishes  $L(0) = R(0)$ , so  $L$  and  $R$  satisfy an ordinary differential equation under the same initial data condition. By uniqueness, we can conclude the desired equality  $L \equiv R$  from the claim.

To prove the claim, recall

$$\frac{1}{2} \nabla(x \cdot Ax) = Ax$$

for any symmetric matrix  $A$ . The claim therefore follows for  $R$  by the chain rule. To prove the claim for  $L$ , we integrate by parts

$$\begin{aligned} \nabla_\xi L(\xi) &= \int_{\mathbb{R}^d} ixe^{-\frac{1}{2}x \cdot Zx} e^{i\xi x} dx \\ &= -i \int_{\mathbb{R}^d} e^{i\xi x} Z^{-1} (\nabla_x e^{-\frac{1}{2}x \cdot Zx}) dx \\ &= iZ^{-1} \int_{\mathbb{R}^d} (\nabla_x e^{i\xi x}) e^{-\frac{1}{2}x \cdot Zx} dx \\ &= -Z^{-1}\xi \int_{\mathbb{R}^d} e^{-\frac{1}{2}x \cdot Zx} e^{i\xi x} dx = -Z^{-1}\xi L(\xi). \end{aligned}$$

This completes the proof of the claim and thereby the lemma.  $\square$

One application of these lemmas is in proving Isserlis' formula, also known as Wick's probability formula. It allows one to compute higher-order moments of a multi-variate Gaussian distribution

$$d\mu = \frac{1}{\sqrt{\det(2\pi\Sigma)}} e^{-\frac{1}{2}x \cdot \Sigma^{-1}x} dx$$

in terms of its covariance matrix  $\Sigma$ , which is positive definite. Before stating the formula, we introduce some notation. Given a multi-index  $\alpha$ , an  $\alpha$ -pairing is a partition into pairs of

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d} = \overbrace{x_1 \cdots x_1}^{\alpha_1\text{-times}} \overbrace{x_2 \cdots x_2}^{\alpha_2\text{-times}} \cdots \overbrace{x_d \cdots x_d}^{\alpha_d\text{-times}}$$

as a set of  $|\alpha|$ -many symbols. Obviously if  $|\alpha|$  is odd, there are no pairings. As an example, a partition  $P$  into pairs of  $x_1^3x_3x_5^2$  is  $(1,1), (1,3), (5,5)$  where we denote  $(j, k) := (x_j, x_k)$  for brevity.

Denote  $m!!$  to be the product of all integers between 1 and  $m$  with the same parity as  $m$ . A useful combinatorial fact is that, for  $m$  even,

$$(m-1)!! = (m-1)(m-3)\cdots(3)(1) = \frac{m!}{2^{m/2}(m/2)!}$$

is exactly the number of ways to pair off  $m$  symbols.

We are now ready to state Isserlis' formula.

**Corollary 1.3** (Isserlis' formula). *Let  $\Sigma$  be a  $d \times d$  positive definite matrix, then*

$$\frac{1}{\sqrt{\det(2\pi\Sigma)}} \int_{\mathbb{R}^d} x^\alpha e^{-\frac{1}{2}x \cdot \Sigma^{-1}x} dx = \sum_P \prod_{(j,k) \in P} \Sigma_{jk} \quad (1.3)$$

where  $\alpha$  is a multi-index and the sum is over all  $\alpha$ -pairings  $P$ . In particular,

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} x^m e^{-\frac{x}{2\sigma^2}} dx = \begin{cases} 0, & \text{if } m \text{ is odd,} \\ (m-1)!! & \text{if } m \text{ is even.} \end{cases} \quad (1.4)$$

*Proof.* By the previous lemma we can write

$$\begin{aligned} \frac{1}{\sqrt{\det(2\pi\Sigma)}} \int_{\mathbb{R}^d} x^\alpha e^{-\frac{1}{2}x \cdot \Sigma^{-1}x} dx &= \frac{1}{i^{|\alpha|}} \partial^\alpha \Big|_{\xi=0} \int_{\mathbb{R}^d} e^{i\xi \cdot x} \frac{e^{-\frac{1}{2}x \cdot \Sigma^{-1}x}}{\sqrt{\det(2\pi\Sigma)}} dx \\ &= (-i\partial)^\alpha \Big|_{\xi=0} e^{-\frac{1}{2}\xi \cdot \Sigma \xi}. \end{aligned}$$

Observe

$$\partial_j (e^{-\frac{1}{2}\xi \cdot \Sigma \xi}) = -(\Sigma \xi)_j e^{-\frac{1}{2}\xi \cdot \Sigma \xi} = -\sum_{k=1}^d \Sigma_{jk} \xi_k e^{-\frac{1}{2}\xi \cdot \Sigma \xi}.$$

Since we are evaluating at  $\xi = 0$ , all the terms with a factor of  $\xi_k$  vanish. The remaining terms each correspond exactly to an  $|\alpha|$ -pairing, emerging from the product rule by pairing off derivatives. The first in a pair brings down a factor of  $-\Sigma_{jk} \xi_k$  from the exponential, and the second removes the factor of  $\xi_k$ . Thus

$$(-i\partial)^\alpha \Big|_{\xi=0} e^{-\frac{1}{2}\xi \cdot \Sigma \xi} = (-i)^{|\alpha|} (-1)^{|\alpha|/2} \sum_P \prod_{(j,k) \in P} \Sigma_{jk} = \sum_P \prod_{(j,k) \in P} \Sigma_{jk}.$$

This proves Isserlis' formula. □

*Alternative proof of Isserlis' formula.* Given a smooth function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , using integration by parts, we have

$$\begin{aligned} \int_{\mathbb{R}^n} x_i f(x_1, x_2, \dots, x_n) e^{-\frac{1}{2}x \cdot \Sigma^{-1}x} dx &= - \int_{\mathbb{R}^n} f(x_1, x_2, \dots, x_n) e_i \cdot \Sigma \nabla e^{-\frac{1}{2}x \cdot \Sigma^{-1}x} dx \\ &= e_i \cdot \Sigma \int_{\mathbb{R}^n} \nabla f(x_1, x_2, \dots, x_n) e^{-\frac{1}{2}x \cdot \Sigma^{-1}x} dx \\ &= \sum_{j=1}^n \Sigma_{ij} \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_j} e^{-\frac{1}{2}x \cdot \Sigma^{-1}x} dx. \end{aligned}$$

Then let  $x_i f = x^\alpha$ , by induction and using Lemma 1.1, we finish the proof of Isserlis' formula. □

**Lemma 1.4** (Stein's lemma). *Let  $X$  be a random variable such that*

$$\mathbb{E}\{Xf(X) - \sigma^2 f'(X)\} = 0,$$

*for all  $f \in C_c^\infty(\mathbb{R})$ . Then  $X \sim N(0, \sigma^2)$ . I.e. If  $\mu$  is a probability measure and*

$$\int_{\mathbb{R}} xf(x)d\mu = \sigma^2 \int_{\mathbb{R}} f'(x)d\mu, \quad \forall f \in C_c^\infty(\mathbb{R}).$$

*Then  $d\mu = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x/2\sigma^2} dx$ .*

*Proof idea.* Take  $f = e^{-\epsilon x^2 + i\xi x}$ , let  $\epsilon$  go to 0, and use the uniqueness of Fourier transform.  $\square$

## 2. LAPLACE'S METHOD: STIRLING'S FORMULA

The factorial function admits an integral representation:

$$n! = \int_0^\infty x^n e^{-x} dx.$$

This should be compared to the Gamma function  $\Gamma(z) = \int_0^\infty x^z e^{-x} \frac{dx}{x}$ . Note that  $\frac{dx}{x}$  is Haar measure for multiplicative group of real numbers;  $x^n$  is multiplicative character and  $e^{-x}$  is an additive character. This is why Gamma function is so important. Stirling's formula is an accurate approximation for  $n$  factorial. We will use it as our introduction to Laplace's method. Laplace himself used it to find the supremum of a function:

$$\sup_{0 < x < 1} f(x) = \lim_{p \rightarrow \infty} \left( \int_0^1 [f(x)]^p dx \right)^{1/p},$$

for  $f$  nonnegative.

We begin by writing

$$\begin{aligned} n! &= \int_0^\infty e^{-x+n \log(x)} dx \\ &= \int_0^\infty e^{-ns+n \log(s)+n \log(n)} n ds \quad (\text{change of variables } x = ns) \\ &= e^{-n} n^{n+1} \int_0^\infty e^{n(-s+\log(s)+1)} ds. \end{aligned}$$

Then we let  $\phi(s) = -s + \log(s) + 1$ . We find  $\phi'(s) = 1/s - 1$  and  $\phi''(s) = -1/s^2$ . So  $\phi$  achieves its maximum at  $s = 1$ . And near the maximum, we can approximate  $\phi$  by Taylor expansion  $\phi(s) = 0 + 0 \cdot s - \frac{1}{2}(s-1)^2 + \dots$ . Inspired by this, we will make  $\phi$  exactly quadratic: We claim (cf. Morse lemma below) we can make a global smooth change of variables, so to  $\tau$ , so that

$$\begin{aligned} \phi(s) &= -\frac{1}{2}\tau^2, \\ \phi'(s) \frac{ds}{d\tau} &= -\tau. \end{aligned}$$

In particular, for  $\tau \in (-\infty, 0)$ , we define  $s(\tau) \in (0, 1)$  via

$$\begin{aligned} \phi(s(\tau)) &= -\frac{1}{2}\tau^2, \\ s(\tau) &= \phi^{-1}(-\frac{1}{2}\tau^2). \end{aligned}$$

For  $\tau \in (0, \infty)$ , we do the same but with  $s(z) > 1$ .

Alternatively,

$$\tau = \begin{cases} -\sqrt{-2\phi(s)} & s < 1 \\ \sqrt{-2\phi(s)} & s > 1. \end{cases} \quad (2.1)$$

We say  $\tau$  does not have singularity at  $s = 1$  even though it moves from one branch of square root to another branch. We can just consider  $\frac{1}{2}\tau^2 = -\frac{1}{2}(s-1)^2$ . We can see  $\tau = |s-1|$  is a bad definition and we should define  $\tau = s-1$  instead. This definition is consistent with complex analysis when we consider holomorphic function  $f(z) = -\frac{1}{2}z^2 + \dots$  and then let  $f(z) = w^2$ . Then

$$\int_0^\infty e^{n\phi(s)} ds = \int_{-\infty}^\infty e^{-\frac{n}{2}\tau^2} \frac{\phi'(s(\tau))}{\tau} d\tau.$$

And  $\frac{ds}{d\tau} = \frac{\phi'(s(\tau))}{\tau}$  is perfectly smooth. In fact,  $\frac{ds}{d\tau}|_{\tau=0} = 1$ .

**Proposition 2.1.** *If  $g(t) \in C^\infty$  on  $\mathbb{R}$  and is  $O(t^{2l+2})$  as  $t \rightarrow \pm\infty$ . Then*

$$\begin{aligned} \int_{-\infty}^\infty e^{-\frac{1}{2}t^2/\sigma^2} g(t) dt &= \int_{-\infty}^\infty e^{-\frac{1}{2}t^2/\sigma^2} [g(0) + tg'(0) + \dots + \frac{g^{(2l+1)}(0)t^{2l}}{(2l+1)!} + O(t^{2l+2})] dt \\ &= \sqrt{2\pi\sigma^2} \left\{ g(0) + g''(0)\frac{\sigma^2}{2} + g^{(4)}(0)\frac{\sigma^4}{2^2 2!} + \dots + g^{(2l)}(0)\frac{\sigma^{2l}}{2^l l!} + O(\sigma^{2l+2}) \right\}. \end{aligned}$$

In our example, we will want  $\sigma^2 = \frac{1}{n}$  and the dominated term for  $g$  is  $g(0) = 1$ .

So the leading term for  $\int_0^\infty e^{n\phi(s)} ds$  is  $\sqrt{\frac{2\pi}{n}}$ .

We have

$$\int_{-\infty}^\infty e^{-\frac{t^2}{2\sigma^2}} g(t) dt = \sqrt{2\pi\sigma^2} \left\{ g(0) + \frac{g''(0)}{2^1 1!} \sigma^2 + \frac{g^{(4)}(0)}{2^2 2!} \sigma^4 + \dots + \frac{g^{(2\ell)}(0)}{2^\ell \ell!} \sigma^{2\ell} + O(\sigma^{2\ell+2}) \right\}$$

if  $g$  is smooth and  $g(t) = O(t^{2\ell+2})$ . On the other hand,

$$n! = n^{n+1} e^{-n} \int_0^\infty e^{n\phi(s)} ds$$

with

$$\begin{aligned} \phi(s) &= \log(s) - s + 1 \\ &= -\frac{1}{2}(s-1)^2 \left\{ 1 - \frac{2}{3}(s-1) + \frac{2}{4}(s-1)^2 - \frac{2}{5}(s-1)^3 + \dots \right\} \\ &= -\frac{1}{2}(s-1)^2 \left\{ 1 + \sum_{m=1}^{\infty} \frac{2(-1)^m}{m+2} (s-1)^m \right\} \end{aligned}$$

which is convergent for  $|s-1| < 1$ . We change variables via  $s = s(\tau)$  so that

$$\phi(s(\tau)) = -\frac{1}{2}\tau^2$$

which implies

$$\phi'(s(\tau)) \frac{ds}{d\tau} = -\tau.$$

If we wrote

$$s = 1 + \tau + a\tau^2 + b\tau^3 + \dots$$

near  $s = 1, \tau = 0$  we find

$$s(\tau) = 1 + \tau + \frac{1}{3}\tau^2 + \frac{1}{36}\tau^3 + \dots$$

which implies that

$$g(\tau) = \frac{ds}{d\tau} = 1 + \frac{2}{3}\tau + \frac{1}{12}\tau^2 + \dots$$

Since

$$\log(s) - s + 1 = -\frac{1}{2}\tau^2,$$

as  $\tau \rightarrow +\infty$ ,  $s \rightarrow +\infty$  and  $s \approx \frac{1}{2}\tau^2$  and so

$$\frac{ds}{d\tau} = \frac{s\tau}{s-1} \approx O\left(\frac{\tau^3}{\tau^2}\right) = O(\tau).$$

As  $\tau \rightarrow -\infty$ ,  $s \rightarrow 0$ ,  $s \approx e^{-\frac{1}{2}\tau^2}$ , so

$$\frac{ds}{d\tau} = \frac{s\tau}{s-1} \approx \frac{\tau e^{-\frac{1}{2}\tau^2}}{-1} = O(1).$$

Thus,

$$\begin{aligned} n! &= n^{n+1} e^{-n} \int_0^\infty e^{n\phi(s)} ds \\ &= n^{n+1} e^{-n} \int_{-\infty}^\infty e^{-\frac{\sigma^2}{2(\frac{1}{n})}} \frac{ds}{d\tau} d\tau \\ &= n^{n+1} e^{-n} \sqrt{\frac{2\pi}{n}} (1 + \frac{1}{12n} + O(\frac{1}{n^2})), \end{aligned}$$

i.e.,

$$n! = \sqrt{2\pi n} n^n e^{-n} (1 + O(\frac{1}{n})).$$

*Remark.* (a) This asymptotic behaviour can also be expressed through the inequalities

$$\sqrt{2\pi n} n^n e^{-n} \leq n! \leq e\sqrt{n} n^n e^{-n}$$

(b) Our contour (straight along the  $x$ -axis) actually passes through the saddle point  $z = 1$  in the most rapidly decreasing (and non-oscillatory way). Indeed, in the complex plane,

$$\phi(1 + x + iy) = -\frac{1}{2}\{(x^2 - y^2) + 2ixy\} + \dots$$

### 3. LAPLACE'S METHOD: THE PARTITION FUNCTION

Let  $p(n)$  be the number of partitions of  $n$ , e.g.,  $p(5) = 7$  because

$$5 = 5 = 4 + 1 = 3 + 2 = 3 + 1 + 1 = 2 + 2 + 1 = 2 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1.$$

The generating function can be defined by

$$\begin{aligned} P(z) &= \sum_{n=0}^{\infty} p(n) z^n \\ &= \prod_{m=1}^{\infty} \frac{1}{1 - z^m} \\ &= \{1 + z + z^2 + \dots\} \cdot \{1 + z^2 + z^4 + \dots\} \cdot \{1 + z^3 + z^6 + \dots\} \cdot \{1 + z^4 + \dots\}. \end{aligned}$$

The function

$$\prod_{m=1}^{\infty} \frac{1}{1 - z^m}$$

converges on  $\mathbb{D}$ . We can extend  $P$  to the complex plane

$$P(re^{i\theta}) = \sum_{n=1}^{\infty} p(n)r^n e^{in\theta}.$$

So

$$p(n) = \oint \frac{P(z)}{z^{n+1}} \frac{dz}{2\pi i}$$

where the integral is taken around a circle of radius  $0 < R < 1$ . We introduce new variables  $z = e^{-u-i\phi}$

$$p(n) = \int_{-\pi}^{\pi} \exp\left\{\sum_{m=1}^{\infty} -\log(1 - e^{-mu-im\phi}) - nu - in\phi\right\} \frac{d\phi}{2\pi}.$$

Hardy and Ramanajan estimated the term

$$\sum_{k=1}^{\infty} -\log|1 - e^{-ku-ik\phi}|$$

by showing when  $\phi = 0$ , this behaves like  $\frac{\pi^2}{6u}$  as  $u \downarrow 0$ ; when  $\phi = 2\pi\frac{p}{q}$  with  $\gcd(p, q) = 1$  then this behaves like  $\frac{\pi^2}{6q^2u}$  as  $u \downarrow 0$ ; when  $\frac{\phi}{2\pi}$  is irrational, then this terms is  $o(\frac{1}{n})$  as  $u \downarrow 0$ . Saddle points come from finding the critical points of things like

$$\frac{\pi^2}{6q^2u} - nu.$$

#### Lecture 4: The Circle Method (continued)

Given a partition function  $p(n)$  and the corresponding generating function  $P(z)$ ,

$$P(z) = \sum_{n=0}^{\infty} p(n)z^n = \prod_{k=1}^{\infty} \frac{1}{1-z^k} = e^{-\sum_k \log(1-e^{-kz})}$$

The Cauchy integral formula gives

$$p(n) = \frac{1}{2\pi i} \oint \frac{P(z)}{z^{n+1}} dz$$

Let  $z = e^{-w}$ ,  $w = u + i\phi$ , then

$$p(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\Phi(w)} d\phi \quad (3.1) \quad \boxed{\text{hongxia1}}$$

Where

$$\Phi(w) = nw - \sum_{k=1}^{\infty} \log[1 - e^{-kw}] \quad (3.2) \quad \boxed{\text{hongxia2}}$$

Our goal in this lecture is to understand the asymptotic behaviour of equation (3.1). We divide our discussion into three parts:

(1) Behaviour at  $\phi = 0$

$$-\sum_{k=1}^{\infty} \log[1 - e^{-ku}] = \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{m} e^{-kum} = \sum_{m=1}^{\infty} \frac{1}{m} \frac{1}{e^{mu} - 1} \quad (3.3) \quad \boxed{\text{hongxia3}}$$

where we used the summation of geometric series and Taylor series expansion of  $\log(1 - x)$ .

Noting that  $\frac{1}{e^{mu}-1} \leq \frac{1}{mu}$  and, moreover that this is the leading behavior when  $mu$  is small,

$$(3.3) \leq \frac{\pi^2}{6u} \quad \text{and } (3.3) \approx \frac{\pi^2}{6u}$$

Thus  $\Phi(u) \approx nu + \frac{\pi^2}{6u}$ , and the saddle point occurs when

$$\frac{d}{du}\Phi(u) = 0 \Rightarrow n - \frac{\pi^2}{6u^2} = 0 \Rightarrow u = \frac{\pi}{\sqrt{6n}}$$

As  $n \rightarrow \infty$ ,  $u \ll 1$ , close to the unit circle.

Note that

$$p(n) \leq \sum_{k=0}^{\infty} \frac{P(k)e^{-ku}}{e^{-nu}} = \frac{P(e^{-u})}{e^{-nu}} = e^{\Phi(u)}$$

since each term in the summation is positive, and when  $k = n$ , we get the term  $p(n)$ .

When  $u = \frac{\pi}{\sqrt{6n}}$ ,

$$p(n) \leq e^{nu + \frac{\pi^2}{6u}} = e^{\frac{n\pi}{\sqrt{6n}} + \frac{\pi^2}{6} \frac{\sqrt{6n}}{\pi}} = e^{\pi\sqrt{\frac{2n}{3}}}$$

which is what we expect.

*Remark.*

- (a) Bounding a single element by the whole sum seems rather vague. But it is consistent with the Laplace method by replacing integral with summation.
- (b) The sum is dominated by terms where  $k$  is close to  $n$ . All other terms are relatively small, so the sum is not that far off from  $p(n)$ .

(2) Next order:

To obtain more accurate estimate, we need to work with Eq.(3.3) directly.

$$\sum_{m=1}^{\infty} \frac{1}{m} \left[ \frac{1}{e^{mu} - 1} \right] = \sum_{m=1}^{\infty} \left[ \frac{1}{m(e^{mu} - 1)} - \frac{1}{m^2(1 + \frac{mu}{2})u} \right] + \frac{1}{m^2(1 + \frac{mu}{2})u}$$

where the term in the bracket is a Riemann sum for a finite integral, and the other term can be evaluated by the digamma function, and involves the Euler-Mascheroni constant  $\gamma$ .

After some computations, we obtain

$$\begin{aligned} \Phi(u) &= nu + \frac{\pi^2}{6u} + \frac{1}{2} \log\left(\frac{u}{2\pi}\right) + O(u) \\ \Phi''\left(\frac{\pi}{\sqrt{6n}}\right) &\approx \frac{2\sqrt{6}}{\pi} n^{3/2} (1 + o(1)) \quad \text{as } n \rightarrow \infty \end{aligned}$$

Then the Laplace's method predicts the dominant portion of the integral (3.1) is  $\frac{1}{4\sqrt{3n}} e^{\pi\sqrt{2n/3}}$ .

Thus, the asymptotic behaviour of  $p(n)$  is:

$$p(n) \approx \frac{1}{4\sqrt{3n}} e^{\pi\sqrt{2n/3}} (1 + o(1)), \quad \text{as } n \rightarrow \infty$$

(3) What's missing?

(a) How does  $\Phi(w)$  behave for  $w$  close to 0 up to the width of the saddle which is  $\approx n^{-3/4}$ ?

Formula above is valid for width  $n^{-1/2}$ , so it is valid for width  $\approx n^{-3/4}$

(b) How about  $\Phi(w)$  away from the saddle point? Need to show the contribution is negligible.

We want to show

$$\frac{1}{2\pi} \int_{100u < \phi < \pi} |e^{\Phi(u)}| d\phi$$

is small.

Using the fact

$$\left| \frac{1}{e^{mw} - 1} \right| \leq \frac{1}{e^{mu} - 1}, \quad \text{where } w = u + i\phi.$$

we get

$$Re\Phi(u + i\phi) \leq \pi\sqrt{2n/3} - \delta\sqrt{n} \quad \text{for } |\phi| \geq 100u$$

Thus,

$$\frac{1}{2\pi} \int_{100u < |\phi| < \pi} |e^{\Phi(u)}| d\phi \lesssim e^{\pi\sqrt{2n/3}} e^{-\delta\sqrt{n}}$$

which  $\rightarrow 0$  as  $n \rightarrow \infty$  if  $\delta > \sqrt{2/3}\pi$ .

#### 4. NONSTATIONARY PHASE AND THE MORSE LEMMA

**Proposition 4.1** (Nonstationary phase). *Let  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  be smooth,  $a \in C_c^\infty(\mathbb{R}^d)$ , and suppose  $\nabla\phi \neq 0$  on  $\text{supp}(a)$ . Then*

$$\int_{\mathbb{R}^d} e^{i\lambda\phi(x)} a(x) dx = O(\lambda^{-N}) \text{ as } \lambda \rightarrow \infty,$$

for any  $N \in \mathbb{N}$ .

*Proof.* Observe that

$$\frac{\nabla\phi}{|\nabla\phi|^2} \cdot \nabla e^{i\lambda\phi} = i\lambda e^{i\lambda\phi}.$$

Integrating by parts, we thus obtain

$$\begin{aligned} \int_{\mathbb{R}^d} e^{i\lambda\phi} a dx &= \frac{1}{i\lambda} \int_{\mathbb{R}^d} a \frac{\nabla\phi}{|\nabla\phi|^2} \cdot \nabla e^{i\lambda\phi} dx \\ &= \frac{i}{\lambda} \int_{\mathbb{R}^d} e^{i\lambda\phi} \nabla \cdot \left( \frac{a \nabla\phi}{|\nabla\phi|^2} \right) dx. \end{aligned}$$

Note that  $\nabla \cdot (\frac{a\nabla\phi}{|\nabla\phi|^2})$  is  $C_c^\infty$ , so we can repeat the argument. After  $N$  iterations, we get

$$\int_{\mathbb{R}^d} e^{i\lambda\phi} a dx = O(\lambda^{-N}).$$

□

*Remark.* As we iterate, it gets hard to understand how the implicit constant depends on  $\phi$  and  $a$ . Later, we will see the *symbol classes*, which behave well under these kinds of operations.

**Lemma 4.2** (Morse lemma). *Let  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  be smooth, and let  $x_0 \in \mathbb{R}^d$  be a nondegenerate critical point of  $\phi$ , that is,<sup>1</sup>*

$$\nabla\phi(x_0) = 0, \det \partial^2\phi(x_0) \neq 0.$$

*Then, there is a local diffeomorphism  $y = \Phi(x)$  satisfying*

$$\partial\Phi(x_0) = Id$$

(i.e.,  $\frac{\partial y_i}{\partial y_j} = \delta_{ij}$ ), so that

$$\begin{aligned}\phi(x) &= \phi(x_0) + \frac{1}{2}\Phi(x) \cdot \partial^2\phi(x_0)\Phi(x) \\ &= \phi(x_0) + \frac{1}{2} \sum_{i,j} y_i \phi_{,ij}(x_0) y_j.\end{aligned}$$

*Proof.* Using a translation, we may reduce to the case  $x_0 = 0$ . There is also no loss of generality if we assume  $\phi(0) = 0$ . Also, choose coordinates for  $x$  such that  $\partial^2\phi$  is diagonal: that is,  $\phi_{,ij}(0) = \lambda_i \delta_{ij}$ , where all  $\lambda_i$  are nonzero.

Recall Taylor's formula with integral remainder:

$$\phi(x) = \phi(x_0) + \nabla\phi(x_0) \cdot (x - x_0) + \int_0^1 (1-t) \frac{\partial^2}{\partial t^2} \phi(x_0 + t(x - x_0)) dt.$$

To verify this, we can integrate by parts and use the chain rule:

$$\begin{aligned}&\int_0^1 (1-t) \frac{\partial^2}{\partial t^2} \phi(x_0 + t(x - x_0)) dt \\ &= (1-t)(x - x_0) \cdot \nabla\phi(x_0 + t(x - x_0))|_0^1 + \int_0^1 \frac{\partial}{\partial t} \phi(x_0 + t(x - x_0)) dt \\ &= -(x - x_0) \nabla\phi(x_0) + \phi(x) - \phi(x_0),\end{aligned}$$

as desired.

Thus in our case,

$$\phi(x) = \frac{1}{2} \sum_{i,j} x_i x_j m_{ij}(x),$$

where  $m_{ij}(x)$  is the symmetric matrix

$$m_{ij}(x) = 2 \int_0^1 (1-t) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(tx) dt,$$

---

<sup>1</sup>Here  $\partial^2\phi(x_0)$  is the *Hessian* at  $x_0$ , whose  $(i,j)$ -entry is  $\frac{\partial^2\phi}{\partial x_i \partial x_j}(x_0)$ . In the sequel, we will also use the notation  $\phi_{,ij}(x_0) \equiv \frac{\partial^2\phi}{\partial x_i \partial x_j}(x_0)$ .

which satisfies

$$m_{ij}(0) = 2 \int_0^1 (1-t) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(0) dt = \frac{\partial^2 \phi}{\partial x_i \partial x_j}(0) = \lambda_i \delta_{ij}.$$

The main part of the argument is to iteratively change variables to bring  $m(x)$  to a *constant* diagonal matrix  $\partial^2 \phi(0)$  by successively zeroing out the unwanted off-diagonal entries.

Imagine we found a local diffeomorphism  $\Phi$  with  $\partial\Phi|_0 = Id$  such that (writing  $\tilde{y} = \Phi(x)$ ):

$$\phi(x) = \sum_{i=1}^{r-1} \frac{1}{2} \lambda_i \tilde{y}_i^2 + \frac{1}{2} \sum_{j,k \geq r} \tilde{m}_{jk}(\tilde{y}) \tilde{y}_j \tilde{y}_k.$$

We claim that the new choice of variables

$$\begin{cases} y_i = \tilde{y}_i & \text{for } i \neq r \\ y_r = \sqrt{\lambda_r^{-1} \tilde{m}_{rr}(\tilde{y})} \left( \tilde{y}_r + \sum_{j>r} \frac{\tilde{m}_{jr}(\tilde{y})}{\tilde{m}_{rr}(\tilde{y})} y_j \right) \end{cases}$$

will lead to an analogous representation with  $r \mapsto r+1$ . Then we will proceed by induction. In order for this argument to work, we must verify:

- (i) The derivative of this change of variables is 1 at the origin, and
- (ii)  $\tilde{m}_{rr}(\tilde{y}) \neq 0$  in a small neighborhood of the origin.

We will complete the proof in the next lecture.

#### 4.1. Stationary Phase.

**Theorem 4.3** (Nonstationary Phase). *Let  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  be smooth,  $a \in C_c^\infty$ . Then*

$$\int e^{i\lambda\phi(x)} a(x) dx = O(\lambda^{-\infty})$$

*Proof.* Observe that

$$\frac{\nabla \phi}{|\nabla \phi|^2} \cdot \nabla e^{i\lambda\phi} = i\lambda e^{i\lambda\phi}.$$

Thus

$$\int_{\mathbb{R}^d} e^{i\lambda\phi} a(x) dx = \frac{1}{i\lambda} \int_{\mathbb{R}^d} a \frac{\nabla \phi}{|\nabla \phi|^2} \cdot \nabla e^{i\lambda\phi} dx = \frac{1}{i\lambda} \int_{\mathbb{R}^d} e^{i\lambda\phi} \nabla \cdot \left( \frac{a \nabla \phi}{|\nabla \phi|^2} \right) dx.$$

Then we may repeat the argument with  $a = \left( \frac{a \nabla \phi}{|\nabla \phi|^2} \right)$  to obtain that the integral is  $O(\lambda^{-N})$  for all  $N$ .  $\square$

*Remark.* The implicit constant gets really complicated. One class of functions that behaves well under these operations are the symbol classes  $\partial^\alpha a \lesssim (1+x^2)^{\frac{m-|\alpha|}{2}}$ .

**Theorem 4.4** (Morse's lemma). *Let  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  be a smooth function and  $x_0$  a nondegenerate critical point of  $\phi$ . Then there is a local diffeomorphism  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$  such that  $\partial\Phi(x_0) = Id$  so that*

$$\phi(x) = \phi(x_0) + \frac{1}{2} \Phi(x) \cdot \partial^2(x_0) \Phi(x) = \phi(x_0) + \frac{1}{2} \sum_{ij} y_i \phi_{ij} y_j$$

*Remark.* The Morse lemma is used in Morse theory to prove manifolds constructed have certain properties. For example, one can show that a Morse function on a Torus has more than two critical points.

*Proof.* Using a translation, we can assume without a loss of generality that  $x_0 = 0$ . Choose coordinates for  $x$  so that  $\partial^2\phi(x)$  is diagonal, i.e.  $\partial^2\phi_{ij} = \lambda_j\delta_{ij}$ . Recall from Taylor's integral formula:

$$\phi(x) = \phi(x_0) + \nabla\phi(x_0) \cdot (x - x_0) + \int_0^1 (1-t) \frac{\partial^2}{\partial t^2} \phi(x_0 + t(x - x_0)) dt.$$

To derive this formula, we use integration by parts:

$$\begin{aligned} \int_0^1 (1-t) \frac{\partial^2}{\partial t^2} \phi(x_0 + t(x - x_0)) dt &= (1-t)(x - x_0) \cdot \nabla\phi(x_0 + t(x - x_0))|_0^1 \\ &\quad + \int_0^1 \frac{\partial}{\partial t} \phi(x_0 + t(x - x_0)) dt \\ &= -(x - x_0) \cdot \nabla\phi(x_0) + \phi(x) - \phi(x_0) \end{aligned}$$

which is exactly what we needed to check. Thus,

$$\phi(x) = \frac{1}{2} \sum_{i,j} (x - x_0)_i \cdot (x - x_0)_j m_{ij}(x)$$

where

$$m_{ij}(x) = 2 \int_0^1 (1-t) \frac{\partial^2 \phi}{\partial x_i \partial x_j} (x_0 + t(x - x_0)) dt = \frac{\partial^2 \phi}{\partial x_i \partial x_j} (x_0) = \lambda_j \delta_{ij}.$$

The main part of the argument is to iterate so that the diagonal of the Hessian matrix is constant. Suppose we have a local diffeomorphism  $\Phi$  such that  $d\Phi|_0 = Id$  so that writing  $\tilde{y} = \Phi(x)$ , we have

$$\varphi(x) = \sum_{i=1}^{r-1} \frac{1}{2} \lambda_i \tilde{y}_i^2 + \sum_{j,k \geq r} \tilde{m}_{jk}(\tilde{y}) \tilde{y}_j \tilde{y}_k.$$

We claim that there exists a new choice of variables  $y_i = \tilde{y}_i$  such that

$$y_r = \sqrt{\lambda_r^{-1} \tilde{m}_{rr}(\tilde{y})} \left( \tilde{y}_r + \sum_{j>r} \frac{\tilde{m}_{jr}(\tilde{y})}{\tilde{m}_{rr}} y_j \right)$$

will lead to an analogous representation with  $r \mapsto r+1$ , i.e. that we may write  $\varphi$  of the form

$$\varphi(x) = \sum_{i=1}^{r-1} \frac{1}{2} \lambda_i y_i^2 + \sum_{j,k \geq r} m_{jk}(y) y_j y_k$$

for some  $m_{jk}$ . If such a coordinate change exists, we would be done by induction. Then

$$\frac{\partial y_j}{\partial \tilde{y}_i} = \delta_{ij}$$

because  $\tilde{m}_{jk}(0) = \delta_{jk}\lambda_k$ . Since our critical point is nondegenerate,  $\lambda_k$  are all nonzero. Hence, we see that  $\tilde{m}_{rr}(\tilde{y})$  in a small neighborhood of 0. Similarly,

$\frac{\partial y_j}{\partial \tilde{y}_i} = \delta_{ij}$ . The inverse function theorem guarantees that  $y \mapsto \tilde{y}$  is a diffeomorphism in some neighborhood of 0. Observe that

$$\begin{aligned}\varphi(x) &= \sum_{j < r} \frac{1}{2} \lambda_j \tilde{y}_j^2 + \sum_{j, k \geq r} \tilde{m}_{jk}(\tilde{y}) \tilde{y}_j \tilde{y}_k \\ &= \sum_{j < r} \frac{1}{2} \lambda_j \tilde{y}_j^2 - \frac{1}{2} \tilde{m}_{rr}(\tilde{y}) - \tilde{y}_r \sum_{j > r} \tilde{m}_{jr}(\tilde{y}) \tilde{y}_j \\ &\quad - \frac{1}{2} \tilde{m}_{rr}(\tilde{y}) \sum_{j, k \geq r+1} \frac{\tilde{m}_{jr}(\tilde{y}) \tilde{m}_{kr}(\tilde{y})}{\tilde{m}_{rr}(\tilde{y}^2)} + \frac{1}{2} \sum_{j, k \geq r} \tilde{m}_{jk}(\tilde{y}) \tilde{y}_j \tilde{y}_k \\ &= \frac{1}{2} \sum_{j \leq r} \lambda_j \tilde{y}_j^2 + \frac{1}{2} \sum_{j, k \geq r+1} \left( \tilde{m}_{jk}(\tilde{y}) + \frac{\tilde{m}_{jr}(\tilde{y}) \tilde{m}_{rk}(\tilde{y})}{\tilde{m}_{rr}(\tilde{y})} \right) \tilde{y}_j \tilde{y}_k.\end{aligned}$$

Letting

$$m_{jk} = \tilde{m}_{jk}(\tilde{y}) + \frac{\tilde{m}_{jr}(\tilde{y}) \tilde{m}_{rk}(\tilde{y})}{\tilde{m}_{rr}(\tilde{y})}$$

we have our desired result. Hence,  $y \mapsto \tilde{y}$  is a diffeomorphism and we're done.  $\square$

**Theorem 4.5** (Stationary Phase). *Let  $\phi \in C^\infty(\mathbb{R}^d)$  have a nondegenerate critical point  $x_0$ . Suppose  $\psi \in C_c^\infty(\mathbb{R}^d)$  is supported in a sufficiently small neighborhood of  $x_0$  (so that  $x_0$  is the only critical point in that support of  $\psi$ ). Then*

$$I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x)} \psi(x) dx$$

is  $O(\lambda^{-d/2})$  as  $\lambda \rightarrow \infty$ . In fact,

$$I(\lambda) = (2\pi)^{d/2} \lambda^{-d/2} \frac{e^{i\lambda\phi(x_0)}}{\sqrt{|\det(\partial^2\phi(x_0))|} e^{i\frac{\pi}{4}\text{sign}(\partial^2\phi)}} + O(\lambda^{-d/2-1}).$$

where  $\text{sign}(\partial^2\phi) = \#\{\lambda \in \text{eig}(\partial^2\phi) : \lambda > 0\} - \#\{\lambda \in \text{eig}(\partial^2\phi) : \lambda < 0\}$ .

*Proof.* Using the Morse lemma, we can reduce to the case

$$\tilde{I}(\lambda) = \int_{\mathbb{R}^d} e^{\frac{i\lambda}{2} y \cdot Qy} \tilde{\psi}(y) dy$$

Notice that  $\tilde{\psi}$  also incorporates the Jacobian  $\det(\frac{\partial x}{\partial y})$ . Nonetheless,  $\tilde{\psi}(0) = \psi(x_0)$ . This leaves us to evaluate

$$\int_{\mathbb{R}^d} e^{\frac{i\lambda}{2} y \cdot Qy - |y|^2} e^{|y|^2} \tilde{\psi}(y) dy$$

so by Lemma 1.1,

$$\tilde{I}(\lambda) = (2\pi)^{d/2} \prod_{j=1}^d (1 - i\lambda\mu_j)^{-1/2} \psi(x_0) + O(\lambda^{-d/2-1})$$

where  $\mu_j$  are the eigenvalues of the matrix. If  $\mu_j > 0$ , then  $(1 - i\lambda\mu_j)^{-1/2} = \mu_j^{-1/2} \lambda^{-1/2} e^{i\pi/4}$  and if  $\mu_j < 0$ , then  $(1 - i\lambda\mu_j)^{-1/2} = \mu_j^{-1/2} \lambda^{-1/2} e^{-i\pi/4}$ . This gives the desired result.  $\square$

We give an alternative proof that  $I(\lambda) = O(\lambda^{-d/2})$ :

*Proof.* For  $\phi$  of the type described, we have  $|\nabla\phi| \lesssim \varepsilon$  in a small “ball” around  $x_0$  of radius  $O(\varepsilon)$ . For simplicity we will assume  $x_0 = 0$ . Outside of this ball  $|\nabla\phi| \gtrsim \varepsilon$ . The Morse Lemma guarantees this is true locally, but typically we may be able to show this more globally.

Let’s prove the result under these hypotheses:

*Step 1:* Smoothly excise a ball of radius approximately  $\lambda^{-1/2}$  by introducing a smooth bump function  $\chi$ . We will have  $\chi$  be 1 when  $|x|$  is in  $[0, \frac{1}{2}]$ , 0 in  $[2, \infty)$ , and have it smoothly transition between these two values on  $(\frac{1}{2}, 2)$ . We then multiply by

$$1 = \chi(\sqrt{\lambda}x) + [1 - \chi(\sqrt{\lambda}x)].$$

Bringing the absolute values inside we see

$$\left| \int e^{i\lambda\phi(x)} \psi(x) \chi(\sqrt{\lambda}x) dx \right| \lesssim \|\psi\|_{L^\infty} \lambda^{-d/2}.$$

*Step 2:* Let

$$\tilde{I}(\lambda) = \int e^{i\lambda\phi(x)} \psi(x) [1 - \chi(\sqrt{\lambda}x)].$$

Note that where  $[1 - \chi(\sqrt{\lambda}x)]$  is non-zero,  $|\nabla\phi| \gtrsim \lambda^{-1/2}$ .

We should integrate by parts using

$$-\frac{i\nabla\phi(x)}{\lambda|\nabla\phi(x)|^2} \cdot \nabla e^{i\lambda\phi(x)} = e^{i\lambda\phi(x)}.$$

The coefficient out front is not really  $O(\frac{1}{\lambda})$  because  $|\nabla\phi|$  can be similar to  $\lambda^{-1/2}$ .

If we think instead in terms of  $\sqrt{\lambda}\phi$  then we have

$$\begin{aligned} \tilde{I} &= \int e^{i\sqrt{\lambda}\cdot\sqrt{\lambda}\phi(x)} \psi(x) [1 - \chi(\sqrt{\lambda}x)] dx \\ e^{i\lambda\phi} &= \frac{i\sqrt{\lambda}\nabla\phi}{\sqrt{\lambda}|\sqrt{\lambda}\nabla\phi|^2} \nabla e^{i\lambda\phi}. \end{aligned}$$

The quantity

$$\frac{\sqrt{\lambda}\nabla\phi}{|\sqrt{\lambda}\nabla\phi|^2}$$

is bounded, which will allow us to continue our estimate.

First, however, we need to discuss symbolic estimates. If  $a$  and  $b$  are both  $C_c^\infty(\mathbb{R} \setminus \{0\})$  functions that satisfy

$$\partial_x^\alpha a(x) \lesssim_\alpha |x|^{-|\alpha|} \tag{4.1} \quad \boxed{7.\text{symbol}}$$

uniformly in  $\lambda$ , then so does  $a \cdot b$ :

$$\begin{aligned} \partial_x^\alpha a \cdot b &= \sum_{\alpha_1 + \alpha_2 = \alpha} \binom{\alpha}{\alpha_1} \partial_x^{\alpha_1} a \cdot \partial_x^{\alpha_2} b \\ |\partial_x^\alpha a \cdot b| &\lesssim \sum_{\alpha_1 + \alpha_2 = \alpha} |x|^{-|\alpha_1|} \cdot |x|^{-|\alpha_2|} \lesssim |x|^{-|\alpha|}. \end{aligned}$$

Analogously, if  $a$  obeys (4.1) and  $a \gtrsim 1$ , then  $\frac{1}{a}$  obeys (4.1).

$$\frac{1}{a} \rightarrow^\partial -\frac{\partial a}{a^2} \rightarrow^\partial -\frac{\partial^2 a}{a^2} + 2 \frac{(\partial a)^2}{a^3} \rightarrow \dots$$

as the numerator is always  $\lesssim |x|^{-|\alpha|}$  and the denominator is always  $\gtrsim 1$ . Of course, this means that if  $a$  and  $b$  are symbols and  $a \gtrsim 1$ , then  $\frac{b}{a}$  is also a symbol. Now let's return to our problem:

$[1 - \chi(\sqrt{\lambda}x)]$  is a symbol because

$$\partial^\alpha \chi(\sqrt{\lambda}x) = \lambda^{|\alpha|/2} (\partial^\alpha \chi)(\sqrt{\lambda}x)$$

and is supported where  $\frac{1}{2}\lambda^{-1/2} \leq x \leq 2\lambda^{-1/2}$ . Hence

$$|\partial^\alpha \chi(\sqrt{\lambda}x)| \lesssim |x|^{-|\alpha|}.$$

Multiplying by  $\psi$  makes this remain a symbol.  $\square$

### Lecture 9: Bounding Oscillatory Integrals with Symbol Classes, continued

Recall that we have not finished showing that

$$I = \int_{\mathbb{R}^d} e^{i\lambda\phi(x)} \psi(x) [1 - \chi(\sqrt{\lambda}x)] dx$$

satisfies  $|I| \lesssim \lambda^{-d/2}$  where  $\psi \in C_c^\infty$  and  $\chi$  is a cutoff function ( $\chi \in C_c^\infty$  and  $\chi = 0$  near 0). To do this, we write

$$e^{i\lambda\phi} = -\frac{i\nabla\phi}{\lambda|\nabla\phi|^2} \cdot \nabla e^{i\lambda\phi}$$

and integrate by parts.

**Lemma 4.6.** *If  $a_k : \mathbb{R}^d \rightarrow \mathbb{C}$  is a symbol supported on  $\lambda^{-1/2} \lesssim |x| \lesssim 1$ , then*

$$\nabla \cdot \left( \frac{i\nabla\phi}{\lambda|\nabla\phi|^2} \cdot \lambda^{-k/2} |x|^{-k} a_k(x) \right) = \lambda^{-(k+1)/2} |x|^{-k-1} \cdot a_{k+1}(x)$$

where  $a_{k+1}$  is also a symbol supported on  $|x| \gtrsim \lambda^{-1/2}$ , with implicit constants independent of  $\lambda$ .

Start with  $a_0(x) = \psi(x)[1 - \chi(\sqrt{\lambda}x)]$  repeatedly applying integration by parts and the above claim, we get

$$I = \int e^{i\lambda\phi} \lambda^{-k/2} |x|^{-l} a_k(x) dx$$

for arbitrarily large  $k$ . We may suppose  $k > d$ . Then we can estimate

$$\begin{aligned} |I| &\lesssim_k \int_{|x| \geq c_k \lambda^{-1/2}} \lambda^{-k/2} |x|^{-k} dx \\ &\lesssim_k \int_{c_k \lambda^{-1/2}}^\infty \lambda^{-k/2} r^{-k} r^{d-1} dr \\ &\lesssim_k \lambda^{-k/2} (1/\sqrt{\lambda})^{d-k} = \lambda^{-d/2}, \end{aligned}$$

which is the estimate we wanted to prove.

*Remark.* Why do we need that  $k > d$ ? After all, the integrand  $a_k$  has compact support, so we should be allowed to replace the  $\infty$  on the upper end of the radial integral with a constant  $C$ , and so the integral converges for every  $k$ . That is true, but for  $k < d$  we then get an upper bound of

$$\int_{c_k \lambda^{-1/2}}^C \lambda^{-k/2} r^{-k} r^{d-1} dr = \frac{1}{d-k} \lambda^{-k/2} r^{d-k} \Big|_{c_k \lambda^{-1/2}}^C$$

$$\begin{aligned} &\lesssim \lambda^{-k/2}((\lambda^{-1/2})^{d-k} + 1^{d-k}) \\ &\lesssim \lambda^{-k/2}. \end{aligned}$$

In other words, the right boundary term at  $r \sim 1$  dominates the left boundary term at  $r \sim \lambda^{-1/2}$  and we get a weaker bound.

*Proof of Lemma.* Rewrite

$$\frac{\nabla\phi}{\lambda|\nabla\phi|^2} = \frac{1}{\sqrt{\lambda}} \cdot \underbrace{\frac{\nabla\phi/|x|}{|\nabla\phi/|x||^2}}_{\text{symbols}} \frac{1}{\sqrt{\lambda}|x|}$$

where I claim that the underbraced expressions are symbols where  $\lambda^{-1/2} \lesssim |x| \lesssim 1$ . Indeed, since  $|x|^{-1}$  is homogeneous then  $\partial^\alpha|x|^{-1}$  must also be homogeneous with  $\partial^\alpha|x|^{-1} \approx |x|^{-|\alpha|-1}$ . It follows that

$$\partial^\alpha \frac{1}{\sqrt{\lambda}|x|} \approx \frac{1}{\sqrt{\lambda}}|x|^{-|\alpha|-1} \lesssim |x|^{-|\alpha|}$$

since we are assuming  $|x| \gtrsim \lambda^{-1/2}$ . Hence  $\frac{1}{\sqrt{\lambda}|x|}$  is a symbol.

Next, I claim that  $\nabla\phi/|x|$  is a symbol. By the product rule it suffices to show that

$$|\partial^\alpha \nabla\phi \partial^\beta |x|^{-1}| \lesssim |x|^{-|\alpha|-|\beta|}$$

for all multi-indices  $\alpha, \beta$ . There are two cases:

- $\alpha = 0$ : Then observe that  $|\nabla\phi| \lesssim |x|$ , and so we have

$$|\nabla\phi \cdot \partial^\beta |x|^{-1}| \lesssim |x| \cdot |x|^{-|\beta|-1} = |x|^{-|\beta|}.$$

- $|\alpha| \geq 1$ : Then  $|\partial^\alpha \nabla\phi| \lesssim 1$ , and so

$$|\partial^\alpha \nabla\phi \cdot \partial^\beta |x|^{-1}| \lesssim 1 \cdot |x|^{-|\beta|-1} \lesssim |x|^{-|\alpha|-|\beta|},$$

since we assumed that  $|\alpha| \geq 1$  and  $|x| \lesssim 1$ .

So  $|\nabla\phi/|x|$  is a symbol. In addition,  $|\nabla\phi| \lesssim |x|$  for small  $x$ , so  $|\nabla\phi/|x|| \gtrsim 1$ , and hence  $\frac{1}{|\nabla\phi/|x||}$  is also a symbol. This implies  $\frac{\nabla\phi/|x|}{|\nabla\phi/|x||^2}$  is a symbol too.

Thus

$$\nabla \cdot \left( \frac{i\nabla\phi}{\lambda|\nabla\phi|^2} \cdot |x|^{-k} \cdot a_k \right) = \nabla \cdot \left( \frac{1}{\sqrt{\lambda}} |x|^{-k} \cdot \tilde{a}_k \right)$$

where  $\tilde{a}_k$  is a symbol, and so

$$\partial^\alpha \nabla \cdot (|x|^{-k} \tilde{a}_k)$$

is bounded by terms of the form

$$\begin{aligned} &|\partial^\beta (|x|^{-k} \tilde{a}_k)| \quad (|\beta| = |\alpha| + 1) \\ &\lesssim_\beta \sum_{\beta=\beta'+\beta''} \underbrace{|\partial^{\beta'}|x|^{-k}|}_{\lesssim |x|^{-k-|\beta'|}} \cdot \underbrace{|\partial^{\beta''} \tilde{a}_k|}_{\lesssim |x|^{-|\beta''|}} \\ &\lesssim \sum_{\beta=\beta'+\beta''} |x|^{-k-|\beta'|-|\beta''|} \\ &\lesssim |x|^{-k-|\beta|} = |x|^{-k-1} |x|^{-|\alpha|} \end{aligned}$$

Therefore

$$a_{k+1} = |x|^{k+1} \nabla \cdot (|x|^{-k} \tilde{a}_k(x))$$

is a symbol.  $\square$

## 5. SOLUTIONS OF SCHRÖDINGER'S EQUATION

Before we finish the proof, we consider an application of this. The (linear) Schrödinger equation

$$i\partial_t \psi = -\frac{1}{2} \Delta \psi.$$

describes the evolution of a complex field  $\psi : \mathbb{R}^d \rightarrow \mathbb{C}$ .

Clearly  $i\partial_t \hat{\psi} = \frac{1}{2} |\xi|^2 \hat{\psi}$  so

$$\hat{\psi}(t, \xi) = e^{-\frac{i}{2}|\xi|^2 t} \hat{\psi}(0, \xi).$$

Thus,

$$\begin{aligned} \psi(t, x) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{i}{2}t|\xi|^2 + ix \cdot \xi} \hat{\psi}(0, \xi) d\xi \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\frac{i}{2}|\xi|^2 t + i\xi \cdot (x-y)} \hat{\psi}(0, y) dy d\xi \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\frac{i}{2}|\xi|^2 t + i\xi \cdot (x-y) - \varepsilon|\xi|^2} \hat{\psi}(0, y) dy d\xi. \end{aligned}$$

Reversing the order of integration (which we can do by Fubini's Theorem), we are led to consider

$$\int e^{-\frac{1}{2}(2\varepsilon+it)|\xi|^2 + i\xi \cdot (x-y)} d\xi = \frac{(2\pi)^{d/2}}{(2\varepsilon+it)^{d/2}} e^{-|x-y|^2/(4\varepsilon+2it)}.$$

Sending  $\varepsilon \downarrow 0$  we get

$$(2\pi)^{d/2} |t|^{-d/2} e^{-i\pi d/4 \operatorname{sign}(t)} e^{i|x-y|^2/(2t)}.$$

Thus,

$$\Psi(t, x) = \int \frac{e^{-i\pi d/4 \operatorname{sign}(t)}}{(2\pi|t|)^{d/2}} e^{i|x-y|^2/(2t)} \hat{\psi}(0, y) dy.$$

$$e^{it\Delta/2} \psi_0 = (2\pi)^{-d/2} \int e^{i\xi x - i|\xi|^2 t/2} \hat{\psi}_0(\xi) d\xi = (2\pi|t|)^{-d/2} e^{-i\pi d \operatorname{sign}(t)/4} \int e^{\frac{i|x-y|^2}{2t}} \psi_0(y) dy$$

is the solution to the PDE

$$i\partial_t \psi = -\frac{1}{2} \Delta \psi$$

with the initial condition  $\psi(0, x) = \psi_0(x)$ . We remark that this PDE has a unique solution that is continuous in time and  $L^2$  in  $x$ . This claim is easy to prove. It's equivalent to

$$i\partial_t \hat{\psi} = -\frac{1}{2} |\xi|^2 \hat{\psi},$$

with initial condition  $\hat{\psi}(t=0, \xi) = \hat{\psi}_0(\xi)$ , having a unique continuous solution for each  $\xi \in \mathbb{R}^d$ .

We now ask what are the long-asymptotics, assuming  $\psi_0 \in \mathcal{S}$  (or  $\psi_0 \in C_c^\infty$ ). Writing

$$\xi x - \frac{1}{2} |\xi|^2 t = -\frac{t}{2} |\xi - \frac{x}{t}|^2 + \frac{|x|^2}{2t},$$

stationary phase predicts that as  $t \rightarrow \pm\infty$  with  $v = \frac{x}{t}$  fixed,

$$\psi(t, x) \sim \frac{e^{\frac{-i\pi d \operatorname{sign}(t)}{4} + i\frac{|x|^2}{2t}}}{|t|^{d/2}} \hat{\psi}_0\left(\frac{x}{t}\right) + O(t^{-\frac{d}{2}-1})$$

The  $|t|^{d/2}$  downstairs comes from conserving the  $L^2$  norm. Note that  $v$  is distance over time and so has the units of speed. This suggests that  $\xi$  is secretly speed. That's true to the quantum mechanical interpretation. Actually it technically should be momentum but in quantum mechanics we use  $-\frac{1}{2m}\Delta$  instead of  $-\frac{1}{2}\Delta$  and so we are essentially assuming that  $m = 1$  (here  $m$  is mass). Hence the momentum and velocity are the same for us.

Note that in the long-term your position is dictated more by your speed than by your starting point.

**Proposition 5.1** (Fraunhofer formula). *For  $\psi_0 \in L^2$ , let*

$$\begin{aligned} \psi(t) &= e^{it\Delta/2} \psi_0 \\ \psi_{sc}(t) &= \frac{e^{\frac{-i\pi d \operatorname{sign}(t)}{4} + i\frac{|x|^2}{2t}}}{|t|^{d/2}} \hat{\psi}_0\left(\frac{x}{t}\right). \end{aligned}$$

Then  $\|\psi - \psi_{sc}\|_{L^2} \rightarrow 0$  as  $t \rightarrow \infty$ .

*Proof.* We have

$$\begin{aligned} \psi - \psi_{sc} &= \frac{e^{\frac{-i\pi d \operatorname{sign}(t)}{2}}}{(2\pi|t|)^{d/2}} \int \left[ e^{\frac{i|x-y|^2}{2t}} - e^{i\frac{|x|^2}{2t} - i\frac{x}{t} \cdot y} \right] \psi_0(y) dy \\ &= \frac{e^{\frac{-i\pi d \operatorname{sign}(t)}{2}}}{(2\pi|t|)^{d/2}} \int \left[ e^{\frac{i|x-y|^2}{2t}} (1 - e^{-i\frac{|y|^2}{2t}}) \right] \psi_0(y) dy \end{aligned}$$

and so

$$\|\psi - \psi_{sc}\|_{L^2} = \|(1 - e^{-i\frac{|y|^2}{2t}})\psi_0(y)\|_{L^2} \rightarrow 0$$

where the convergence is by the Dominated Convergence Theorem. Note that the proof is only technically valid if  $\psi_0 \in L^1$  in addition to being in  $L^2$ . The proof can be extended to general  $\psi_0$  by a standard approximation argument.  $\square$

We now discuss the wave equation. Solving it leads to the study of

$$\begin{bmatrix} \cos(t|\xi|) & \frac{\sin(t|\xi|)}{|\xi|} \\ -|\xi| \sin(t|\xi|) & \cos(t|\xi|) \end{bmatrix}$$

and hence to  $e^{it|\xi|}$ . The graph of  $\xi \mapsto |\xi|$  is a cone. Is a cone curved? Sort of. It isn't because you can fold a flat piece of paper into a cone. It is because you can look at it extrinsically and it is curved.

We use polar coordinates

$$\int e^{itr - ixr \cos(\theta)} \hat{\psi}_0(r, \omega) r^{d-1} dr d\Omega$$

where  $\theta$  is the angle between  $x$  and  $\xi$ . Also  $d\Omega$  is solid angle measure.

For the  $r$  integral the phase is non-stationary unless  $x \cos \theta = t$  but the  $\theta$  integral is amenable to stationary phase.

We end by continuing our discussion of the Schrödinger equation. Recall we derived that a solution to the equations

$$\begin{aligned} i\partial_t \psi &= \frac{1}{2}\Delta \psi \\ \psi(0, x) &= \psi_0(x) \end{aligned}$$

can be given by the explicit formula

$$\psi(t, x) = \int_{\mathbb{R}^d} \frac{e^{-i\pi d \operatorname{sign}(t)/4}}{(2\pi|t|)^{d/2}} e^{i|x-y|^2/2t} \psi_0(y) dy.$$

For fixed  $t$  it is a common abbreviation to write the function  $x \mapsto \psi(t, x)$  as  $e^{it\Delta/2}\psi_0$ , in analogy to how for a matrix  $A$ , the vector equation

$$\begin{aligned} \partial_t v &= Av, \\ v(0) &= v_0, \end{aligned}$$

is solved by  $v(t) = e^{tA}v$ . The explicit formula has the following immediate corollary:

**Corollary 5.2** (Dispersive estimate). *If  $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ , then we have*

$$\|e^{it\Delta/2}\psi_0\|_{L^\infty} \leq \frac{1}{(2\pi|t|)^{d/2}} \|\psi_0\|_{L^1}.$$

*Proof.* For any fixed  $x$ , we can use the exact formula to calculate

$$|e^{it\Delta/2}\psi_0(x)| \leq \left\| \frac{e^{-i\pi d \operatorname{sign}(t)/4}}{(2\pi|t|)^{d/2}} e^{i|x-y|^2/2t} \right\|_{L_y^\infty} \|\psi_0\|_{L^1} = \frac{1}{(2\pi|t|)^{d/2}} \|\psi_0\|_{L^1}$$

□

Note that the  $L_x^2$  norm of  $\psi(t, x)$  remains constant. In order for the  $L_x^2$  norm to remain constant while the  $L_x^\infty$  norm goes like  $|t|^{-d/2}$ , the mass of  $\psi$  must “spread out” to a volume of roughly  $|t|^d$ . That is why this corollary is called the dispersive estimate.

## 6. REARRANGEMENT

**Definition 6.1.** Given  $A \subseteq \mathbb{R}^d$ , we define  $A^* = B(0, r)$  (open ball), where  $r$  is chosen so that  $|A| = |A^*|$ . This set is called the Schwarz rearrangement of  $A$ .

**Definition 6.2.** A measurable function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is said to vanish at infinity if  $\forall \epsilon > 0$ ,

$$|\{x : |f(x)| > \epsilon\}| < \infty$$

*Remark.* Clearly any  $f \in L^p$ ,  $1 \leq p < \infty$ , lies in the class of function that vanish at infinity.

**Definition 6.3.** For a finite measure  $\mu$ , we say  $\mu$  is tight if  $\forall \epsilon > 0$ ,  $\exists$  a compact set  $K_\epsilon$  so that  $|\mu|(K_\epsilon^c) < \epsilon$ .

For example, every Borel measure on  $C([0, 1])$  is tight.

**Definition 6.4.** Given  $f : \mathbb{R}^d \rightarrow \mathbb{C}$ , vanishing at infinity, we define the radially-symmetric (spherically-symmetric) decreasing (or Schwarz) rearrangement  $f^*$  as follows:

$$f^*(x) = \int_0^\infty \chi_{\{y : |f(y)| > \lambda\}}(x) d\lambda, \quad (6.1)$$

I.e. we Schwartz-rearrange the super level sets.

Properties of Schwarz rearrangement:

- (1)  $f^*(x)$  is spherically symmetric: the value depends only on  $|x|$ .
- (2) It is a non-increasing function of radius, because if  $|x_0| \leq |x_1|$ , then

$$x_1 \in \{y : |f(y)| > \lambda\}^* \implies x_0 \in \{y : |f(y)| > \lambda\}^*.$$

So at  $x_0$  we integrate over a set of  $\lambda$ 's. That is at least as large as the integral at  $x_1$ .

- (3) If  $|g(x)| \geq |f(x)|$ , then for any  $\lambda \geq 0$ ,  $|\{y : |g(y)| > \lambda\}| \geq |\{y : |f(y)| > \lambda\}|$  and so

$$\{y : |g(y)| > \lambda\}^* \supseteq \{y : |f(y)| > \lambda\}^*.$$

Hence  $g^*(x) \geq f^*(x)$ .

- (4)  $f$  and  $f^*$  are equi-measurable:

$$|\{x : |f(x)| > \lambda\}| = |\{x : f^* > \lambda\}|$$

Indeed

$$\{x : |f(x)| > \lambda\}^* = \{x : f^*(x) > \lambda\} = \{x : |f^*(x)| > \lambda\},$$

then

$$|\{x : |f(x)| > \lambda\}| = |\{x : |f(x)| > \lambda\}^*| = |\{x : |f^*(x)| > \lambda\}|$$

This means that for  $1 \leq p < \infty$ ,

$$\|f\|_{L^p} = \|f^*\|_{L^p},$$

because

$$\|f\|_{L^p}^p = \int_0^\infty p\lambda^{p-1} |\{f > \lambda\}| d\lambda = \int_0^\infty p\lambda^{p-1} |\{f^* > \lambda\}| d\lambda = \|f^*\|_{L^p}^p.$$

Ditto for  $L^\infty$  and also for Lorentz spaces.

**Lemma 6.5.** *If measurable functions  $f, g : \mathbb{R}^d \rightarrow [0, \infty)$  vanish at infinity, then*

$$\int_{\mathbb{R}^d} f(x)g(x)dx \leq \int_{\mathbb{R}^d} f^*(x)g^*(x)dx. \quad (6.2)$$

Moreover if  $f$  is strictly decreasing, which is to say  $f(x) = f^*$  and  $\forall |x| < |y|, f(x) > f(y)$ , then the equality in (6.2) holds if and only if  $g = g^*$  a.e.

*Remark.* The restriction for the case of equality is necessary. If we consider a ball  $B$  centering at 0 and  $f = \chi_B$ , then for any measurable function  $g$  that  $\text{supp}(g) \subseteq B$ , the equality holds.

*Proof.* Note that for  $A, B \subseteq \mathbb{R}^d$ ,

$$|A \cap B| \leq \min\{|A|, |B|\} = \min\{|A^*|, |B^*|\} = |A^* \cap B^*|, \quad (6.3)$$

where the last equality holds because  $A^*$  and  $B^*$  are both balls centering at 0.

$$\begin{aligned} \text{LHS of (6.2)} &= \int_{\mathbb{R}^d} \int_0^\infty \int_0^\infty \chi_{\{f>\lambda\}}(x) \chi_{\{g>\mu\}}(x) d\lambda d\mu dx \\ &= \int_0^\infty \int_0^\infty \int_{\mathbb{R}^d} \chi_{\{f>\lambda\}}(x) \chi_{\{g>\mu\}}(x) dx d\lambda d\mu \\ &= \int_0^\infty \int_0^\infty |\{x : f > \lambda\} \cap \{x : g > \mu\}| d\lambda d\mu \end{aligned}$$

$$\begin{aligned} &\leq \int_0^\infty \int_0^\infty |\{x : f > \lambda\}^* \cap \{x : g > \mu\}^*| d\lambda d\mu \\ &= \int_0^\infty \int_0^\infty \int_{\mathbb{R}^d} \mathcal{X}_{\{f>\lambda\}^*}(x) \mathcal{X}_{\{g>\mu\}^*}(x) dx d\lambda d\mu = \text{RHS of (6.2)}, \end{aligned}$$

where we can exchange the order of calculating the integral because both  $f$  and  $g$  vanish at infinity and the inequality step is because of 6.3.

Notice the equality in (6.2) means that

$$|\{x : f > \lambda\} \cap \{x : g > \mu\}| = |\{x : f > \lambda\}^* \cap \{x : g > \mu\}^*|, \text{ for a.e. } \lambda, \mu \in [0, \infty).$$

□

*Proof.* Suppose now that  $f$  is strictly decreasing radially and we have equality in (6.2). We claim that every ball can be realized as a super-level set of  $f$ . Equivalently, writing

$$\{x : f(x) > \lambda\} = B(0, r(\lambda)),$$

we want to show that  $r(\lambda)$  achieves every positive value. Since  $f$  is strictly decreasing radially, we know  $r(\lambda)$  is non-increasing and  $r(\lambda) \rightarrow \infty$  as  $\lambda \rightarrow 0$ . Moreover, we know  $f$  is finite everywhere and vanishes at infinity, so by monotone convergence

$$|\{x : f(x) > \lambda\}| \xrightarrow{\lambda \rightarrow \infty} 0$$

and therefore  $r(\lambda) \rightarrow 0$  as  $\lambda \rightarrow \infty$ . Observe continuity of  $r(\lambda)$  would furnish the claim. Indeed, since it is monotonic, every discontinuity at  $\lambda$  would be a jump discontinuity, i.e.

$$\lim_{\epsilon \rightarrow 0} r(\lambda + \epsilon) - r(\lambda - \epsilon) > 0.$$

However, by monotone convergence, this would imply

$$|\{x : f(x) = \lambda\}| = \lim_{\epsilon \rightarrow 0} |\{x : f(x) > \lambda - \epsilon\}| - |\{x : f(x) > \lambda + \epsilon\}| > 0.$$

However,  $f$  is strictly decreasing radially, so it cannot be constant on a set of positive measure. This implies  $r(\lambda)$  is continuous, completing the proof of the claim. It follows from the claim that for a.e.  $\mu$

$$|\{x \in B(0, r) : g(x) > \mu\}| = |\{x \in B(0, r) : g^*(x) > \mu\}|$$

for a dense set of  $r > 0$ . As  $\{g^* > \mu\}$  is a ball, we obtain

$$|\{x : g(x) > \mu\} \Delta \{x : g^*(x) > \mu\}| = 0$$

for a.e.  $\mu$ . This allows us to conclude  $g = g^*$  a.e. Indeed, let  $g_n(x) = \min\{n, g(x)\}$ , so  $(g_n)^* = (g^*)_n$  and

$$g_n(x) - g_n^*(x) = \int_0^n (\mathbf{1}_{\{g>\mu\}}(x) - \mathbf{1}_{\{g^*>\mu\}}(x)) d\mu.$$

Thus,

$$\begin{aligned} \int_{\mathbb{R}^d} |g_n(x) - g_n^*(x)| e^{-x^2} dx &\leq \int_0^n \int_{\mathbb{R}^d} |\mathbf{1}_{\{g>\mu\}}(x) - \mathbf{1}_{\{g^*>\mu\}}(x)| e^{-x^2} dx d\mu \\ &\leq \int_0^n \int_{\mathbb{R}^d} \mathbf{1}_{\{g>\mu\} \Delta \{g^*>\mu\}}(x) e^{-x^2} dx d\mu = 0. \end{aligned}$$

Note we introduce the Gaussian weight so that a priori we know the integral converges and we are not multiplying zero by infinity. This proves that  $g_n = g_n^*$  a.e. and, sending  $n \rightarrow \infty$ , that  $g = g^*$  a.e. □

**Corollary 6.6.** *Let  $f, g \in L^2(\mathbb{R}^d)$  be non-negative, then*

$$\|f^* - g^*\|_{L^2} \leq \|f - g\|_{L^2}.$$

*Proof.* We have

$$\begin{aligned} \|f^* - g^*\|_{L^2}^2 &= \|f^*\|_{L^2}^2 + \|g^*\|_{L^2}^2 - 2 \int_{\mathbb{R}^d} f^*(x)g^*(x)dx \\ &\leq \|f\|_{L^2}^2 + \|g\|_{L^2}^2 - 2 \int_{\mathbb{R}^d} f(x)g(x)dx = \|f - g\|_{L^2}^2. \end{aligned}$$

□

This estimate holds for more general  $L^p$ -spaces. Essentially, rearrangement brings functions together in  $L^p$ . In fact, we have the following more general proposition.

**Proposition 6.7.** *Let  $\phi : \mathbb{R} \rightarrow [0, \infty)$  be convex with  $\phi(0) = 0$  and suppose that  $f, g : \mathbb{R}^d \rightarrow [0, \infty)$  vanish at infinity. Then*

$$\int_{\mathbb{R}^d} \phi(f^* - g^*)dx \leq \int_{\mathbb{R}^d} \phi(f - g)dx.$$

*Proof.* Write  $\phi = \phi_+ + \phi_-$ , where

$$\phi_+(\lambda) = \begin{cases} \phi(\lambda), & \text{if } \lambda \geq 0 \\ 0, & \text{if } \lambda < 0 \end{cases}, \quad \phi_-(\lambda) = \begin{cases} 0, & \text{if } \lambda \geq 0, \\ \phi(\lambda), & \text{if } \lambda < 0 \end{cases}.$$

Both  $\phi_+$  and  $\phi_-$  are convex. It suffices to prove the inequality for  $\phi_+$ . Recall that convex functions are Lipschitz continuous on compact domains and therefore absolutely continuous. Thus there exists a left-continuous function  $\phi'_+$  such that

$$\phi_+(\lambda) = \int_0^\lambda \phi'_+(\mu)d\mu.$$

Then if  $f(x) > g(x)$ , we have

$$\begin{aligned} \phi_+(f(x) - g(x)) &= \int_0^{f(x)} \phi'_+(f(x) - s)\mathbf{1}_{\{g \leq s\}}(x)ds \\ &= \int_{g(x)}^{f(x)} \phi'_+(f(x) - s)ds \\ &= -\phi_+(0) + \phi_+(f - g) \\ &= \int_0^\infty \phi'_+(f(x) - s)\mathbf{1}_{\{g \leq s\}}(x)ds. \end{aligned}$$

If  $g(x) > f(x)$ , we see this continues to hold with everything vanishing. Thus

$$\int_{\mathbb{R}^d} \phi_+(f(x) - g(x))dx = \int_{\mathbb{R}^d} \int_{\mathbb{R}} \phi'_+(f(x) - s)\mathbf{1}_{\{g \leq s\}}dsdx.$$

□

$$\int_{\mathbb{R}^d} \phi_+(f(x) - g(x))dx = \int_{\mathbb{R}^d} \int_R \phi'_+(f(x) - s)\chi_{g \leq s}(s)dsdx.$$

As  $\phi'_+$  is left continuous and increasing, by the lemma below, we have

$$[\phi'_+(f(x) - s)]^* = \phi'_+(f^*(x) - s).$$

Writing  $\chi_{g \leq s} = 1 - \chi_{g > s}$  and using equi-measurability and the previous lemma, we get

$$\int_{R^d} \phi_+(f(x) - g(x)) dx \geq \int_R \int_R^d \phi'_+(f^*(x) - s) \chi_{g \leq s}^*(s) dx ds.$$

The new RHS is  $\int_R \int_{R^d} \phi'_+(f^*(x) - s) \chi_{g^* \leq s}(s) dx ds$ . If we reverse the earlier algebra, the RHS will be equal to

$$\int_{R^d} \phi_+(f^* - g^*) dx.$$

**Lemma 6.8.** *If  $F$  is left continuous and increasing, then*

$$F(f)^* = F(f^*),$$

for every  $f : R^d \rightarrow [0, \infty)$ .

*Proof.* We need that

$$\{x : F(f(x)) > \lambda\} = \{x : f(x) > \mu\}$$

for  $\mu = "F^{-1}(\lambda)"$ .

Without loss of generality, we assume  $F(0) \leq \lambda < F(\infty)$ .

If  $\lambda$  belongs to the range of  $F$ , we set  $\mu = F^{-1}(\lambda)$ .

If  $\lambda \in [\lim_{x \rightarrow x_0^-} F(x), \lim_{x \rightarrow x_0^+} F(x)]$ , we set  $\mu = x_0$ .  $\square$

**Theorem 6.9** (Riesz rearrangement in 1D). *Let  $f, g, h : R \rightarrow [0, \infty)$  be decaying at infinity. Then*

$$\int_R \int_R f(x) g(x-y) h(y) dx dy \leq \int_R \int_R f^*(x) g^*(x-y) h^*(y) dx dy. \quad (*)$$

Riesz ineq

If  $g = g^*$  is strictly decreasing, the equality occurs only if  $f(x) = f^*(x-a)$  a.e. and  $h(x) = h^*(x-a)$  a.e. for a common translation  $a$ .

*Remark.* The same is true in  $R^d$  but we need this 1D result to prove that. A  $R^3$  example of the meaning of this is that the fact that planets are approx spherical.

$$E = \int_{R^3} \int_{R^3} \rho(x) \frac{1}{4\pi|x-y|} \rho(y) dx dy.$$

Here  $\rho$  is the density of matter and  $\frac{1}{4\pi|x-y|}$  is Newton's Law of gravity.  
Riesz says

$$E(\rho) \geq E(\rho^*),$$

with equality iff  $\rho = \rho^*$ .

*Proof.* By performing a layer cake decomposition,

$$f(x) = \int_0^\infty \chi_{\{f > \lambda\}}(x) d\lambda,$$

to each of the functions. We see that it suffices to prove  $(*)$  where each function is just an indicator of a set of finite measure. I.e.

$$\int_R \int_R \chi_A(x) \chi_B(x-y) \chi_C(y) dx dy \leq \int_R \int_R \chi_{A^*}(x) \chi_{B^*}(x-y) \chi_{C^*}(y) dx dy.$$

Suppose now we could prove this only when  $A, B, C$  are open sets. This would suffice. Outer regularity guarantees that  $A, B, C$  can be approximated by larger open sets on measure, i.e. there is open decreasing sequence  $\{A_n\}$  s.t.  $| \cap A_n \setminus A | = 0$ . Every open set is a countable union of disjoint open intervals. As  $A$  has finite measure, we can approximate  $A$  by a mere finite sum of open intervals  $\|\chi_{A_n} - \chi_A\|_{L^1} = |A_n \Delta A| \rightarrow 0$  as  $n$  goes to infinity and each  $A_n$  is finite sum of open intervals (by D.C.T.).

Let's now turn to the problem of proving (\*) when

$$A = \cup_{i=1}^n A_i, B = \cup_{j=1}^m B_j, C = \cup_{k=1}^l C_k,$$

where  $\{A_i\}, \{B_j\}, \{C_k\}$  are disjoint families of open intervals. we define mid points:

$$A_i = (a_i - |A_i|/2, a_i + |A_i|/2).$$

and likewise  $b_j$  and  $c_k$ . Then we define  $A_i(t) = (a_i t - |A_i|/2, a_i t + |A_i|/2)$  and likewise  $B_j(t), C_k(t)$ .

$$\begin{aligned} I_{ijk}(x) &= \int_R \int_R \chi_{A_i(t)}(x) \chi_{B_j(t)}(x-y) \chi_{C_k(t)}(y) dx dy \\ &= \int_R \int_R \chi_{(-\frac{|A_i|}{2}, \frac{|A_i|}{2})}(x-a_i t) \chi_{(-\frac{|B_j|}{2}, \frac{|B_j|}{2})}(x-y-b_j t) \chi_{(-\frac{|C_k|}{2}, \frac{|C_k|}{2})}(y-c_k t) dx dy \\ &= \int_R \int_R \chi_{(-\frac{|A_i|}{2}, \frac{|A_i|}{2})}(x) \chi_{(-\frac{|B_j|}{2}, \frac{|B_j|}{2})}(x-y) \chi_{(-\frac{|C_k|}{2}, \frac{|C_k|}{2})}(y+t(a_i - b_j - c_k)) dx dy \end{aligned} \tag{**}$$

The last line is using change of variable.

The  $x$  integral in (\*\*) is just a convolution of indicators of  $(-\frac{|A_i|}{2}, \frac{|A_i|}{2}), (-\frac{|B_j|}{2}, \frac{|B_j|}{2})$  since  $\chi_{(-\frac{|B_j|}{2}, \frac{|B_j|}{2})}$  is even. And we could compute this convolution

$$\phi_{i,j}(y) = \int_R \chi_{(-\frac{|A_i|}{2}, \frac{|A_i|}{2})}(x) \chi_{(-\frac{|B_j|}{2}, \frac{|B_j|}{2})}(x-y) dx = \begin{cases} y + z_{ij} + |A_i| \wedge |B_j| & y \in [-z_{ij} - |A_i| \wedge |B_j|, -z_{ij}] \\ |A_i| \wedge |B_j| & y \in [-z_{ij}, z_{ij}] \\ -y + z_{ij} + |A_i| \wedge |B_j| & y \in [z_{ij}, z_{ij} + |A_i| \wedge |B_j|] \\ 0 & \text{else,} \end{cases}$$

where  $z_{ij} = \frac{|A_i| + |B_j| - 2|A_i| \wedge |B_j|}{2}$ .  $\square$

$$I_{ijk} = \int \phi_{ij}(y) \chi_{(-\frac{|C_k|}{2}, \frac{|C_k|}{2})}(y+t[a_i - b_j - c_k]) dy, \tag{6.4}$$

where  $\phi_{ij}(y)$  is the convolution of centered intervals depicted above. As  $\phi_{ij}(y)$  is symmetric non-increasing, we see that  $I_{ijk}(t)$  is a non-increasing function of  $t$ . Thus we may decrease  $t$ . However we stop if any pair of intervals, say  $A_i$  &  $A_{i'}$  (or  $B_j$  &  $B_{j'}$ , or  $C_k$  &  $C_{k'}$ ) touch. At that moment we replace the two intervals by a simple interval. Then we restart the analysis of  $\phi_{ijk}(t)$  with these new intervals and  $t$  set back to 1.

The procedure finishes in finite time because each such conjunction reduces the total number of intervals. Eventually, we are reduced to single  $A, B \& C$  intervals. Taking  $t \downarrow 0$ , we deduce that

$$I(\chi_A, \chi_B, \chi_C) \leq I(\chi_{A^*}, \chi_{B^*}, \chi_{C^*}). \tag{6.5}$$

Suppose now  $g = g^*$  and the equality holds, then the equality must hold for

$$\chi_{f>\lambda}, \chi_{(-\frac{r}{2}, \frac{r}{2})}, \chi_{g>\lambda} \quad (6.6)$$

for a.e.. To finish, we must show that

$$I(\chi_A, \chi_{(-\frac{r}{2}, \frac{r}{2})}, \chi_C) = I(\chi_{A^*}, \chi_{(-\frac{r}{2}, \frac{r}{2})}, \chi_{C^*}), \quad (6.7)$$

then

$$\|\chi_A(x) - \chi_{A^*}(x-a)\|_{L^2} + \|\chi_C(x) - \chi_{C^*}(x-a)\|_{L^2} = 0. \quad (6.8)$$

If  $r > |A| + |C|$ , then

$$I(\chi_{A^*}, \chi_{(-\frac{r}{2}, \frac{r}{2})}, \chi_{C^*}) = \int \int_{|x-y|<\frac{r}{2}} \chi_{A^*}(x) \chi_{C^*}(y) dx dy \quad (6.9)$$

$$= |A^*||C^*| \quad (6.10)$$

$$= |A||C| \quad (6.11)$$

because the (Minkowski) different set  $\{x-y : x \in A^*, y \in C^*\}$  is a subset of  $(-\frac{r}{2}, \frac{r}{2})$ .

Likewise, if we consider the essential differences set

$$(A - B)^{ess} = \{x - y : x \in A \text{ \& } y \in C \text{ \& both are Lebesgue points}\} \quad (6.12)$$

the equality in Eqn. (\*) guarantees that

$$(A - C)^{ess} \subseteq [-\frac{r}{2}, \frac{r}{2}]. \quad (6.13) \quad \boxed{\text{essent}}$$

This in turn guarantees that the essential diameters obey

$$ess\_diam(A) + ess\_diam(C) = |A| + |C|, \quad (6.14)$$

where

$$ess\_diam(A) = \sup\{x - x' : x, x' \in A \text{ are Lebesgue points}\}. \quad (6.15)$$

The equality Eqn. (6.13) (as opposed to  $\geq$ ) guarantees that  $A$  &  $C$  are both (a.e.) intervals. Retuning to the full strength of Eqn (6.13) we deduce that  $A$  &  $C$  have the same center.

**Definition 6.10.** Given  $A \subseteq \mathbb{R}^d$  and  $e$  to be a unit vector in  $\mathbb{R}^d$ , we define the Steiner symmetrization of  $A$  by applying the usual rearrangement on each line parallel to  $e$ .

We center each such crosssector of the line parallel to  $e$  and  $A$  on the “plane”  $e^\perp$ . I.e.,  $e^\perp$  serves as the origin we are doing the arrangement at each line. Note: if  $A$  is measurable, then so is  $A^{*e}$ . The key observations are those:

- (i)  $m(x) = \int \chi_A(x, y) dy$  is measurable.
- (ii) The (sub)graph of a measurable function is measurable

$$\{(x, y) : -\frac{m(x)}{2} \leq y \leq \frac{m(x)}{2}\}. \quad (6.16)$$

This explains the measurability when  $e = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . Rotations preserve measurability.

From the one-dimensional Riesz theorem, we see that

$$I(\chi_A, \chi_B, \chi_C) \leq I(\chi_{A^*}, \chi_{B^*}, \chi_{C^*}) \quad (6.17)$$

which we could iterate using “lots” of different directions. The hard part about proving Riesz theorem in  $\mathbb{R}^d$  is showing that repeated Steiner symmetrization leads to complete symmetrization  $A \rightarrow \dots \rightarrow A^*$ .

Following the previous model, we define the **Steiner symmetrization** of a measurable function  $f$  with respect to some direction  $e$  in  $R^n$  by:

$$f^{*e} = \int_0^\infty \chi_{\{y:f(y)>\lambda\}}^{*e}(x) d\lambda$$

**Theorem 6.11** (Riesz's rearrangement in 2-D). *Let  $f, g$  and  $h$  be three nonnegative functions on  $R^2$ , and decaying at infinity. Then*

$$I(f, g, h) = \iint f(x)g(x-y)h(y) dx dy \quad (6.18) \quad \boxed{\text{Covid-19}}$$

satisfies

$$I(f, g, h) \leq I(f^*, g^*, h^*) \quad (6.19) \quad \boxed{\text{Covid-20}}$$

Moreover, if  $g$  is strictly decreasing and equality holds, then

$$f(x) = f^*(x-a), \quad h(x) = h^*(x-a)$$

a.e. for some  $a \in R^2$ .

*Proof.* As in the 1-D case, we first reduce to the case of sets, via a layer-cake decomposition.

Let  $X$  denote the Steiner symmetrization parallel with the x-axis, and  $Y$  the Steiner symmetrization parallel with the y-axis. Let  $R$  be an anticlockwise rotation by some angle  $\alpha \notin 2\pi\mathbb{Q}$ .<sup>2</sup> By rotating a function  $f$ , we mean rotating the argument, i.e.  $Rf(x) = f(Rx)$ .

Apply the 1-D Riesz rearrangement inequality to  $X$  and  $Y$ , and by the linearity of  $R$ , we have

$$I(\chi_A, \chi_B, \chi_C) \leq I(YXR\chi_A, YXR\chi_C, YXR\chi_C)$$

To iterate this, let  $A_k = (YXR)^k A$ . Likewise for  $B_k$  and  $C_k$ . To prove the inequality (6.19), we just need to show that  $A_k$  converges to  $A^*$  in a suitable sense, e.g.

$$\|\chi_{A_k} - \chi_{A^*}\|_{L^2} \rightarrow 0$$

And similarly for  $B$  and  $C$ .

First, notice that  $I$  is continuous in this topology since

$$\begin{aligned} |I(\chi_A, \chi_B, \chi_C) - I(\chi_{A'}, \chi_{B'}, \chi_{C'})| &\leq |I(\chi_A - \chi_{A'}, \chi_B, \chi_C)| + |I(\chi_{A'}, \chi_B - \chi_{B'}, \chi_C)| + |I(\chi_{A'}, \chi_B, \chi_C - \chi_{C'})| \\ &\leq \|\chi_A - \chi_{A'}\|_{L^2} \|\chi_B\|_{L^1} \|\chi_C\|_{L^2} + \dots \\ &= \|\chi_A - \chi_{A'}\|_{L^2} \|\chi_B\|_{L^2} \|\chi_C\|_{L^2} + \dots \end{aligned}$$

By the same argument, we may restrict the sets  $A$ ,  $B$ , and  $C$  to be bounded.

Given a bounded set  $A$ , let

$$\chi_{A_1} = YXR\chi_A$$

Then  $A_1$  is a set with the same measure as  $A$ . It is reflection symmetric about the  $x-$  and  $y-$ axes, and can be represented by

$$A_1 = \{(x, y) : |y| \leq \omega(|x|)\}$$

---

<sup>2</sup>Use an irrational rotation to ensure we get a complete symmetrization of a set.

for some non-increasing, symmetric and lower semicontinuous function  $\omega$ . Note that since  $A$  is bounded, so is  $\omega$  and  $\text{supp}(\omega)$ . More generally, we can define  $\omega_k$  so that

$$A_k = \{(x, y) : |y| \leq \omega_k(|x|)\}$$

Since  $A$  is bounded, which implies that  $\omega_k$  is uniformly bounded, we can apply Bolzano-Weierstrass to get a subsequence in  $k$  so that  $\omega_k$  converges at each  $x$ . Then we can get a subsequence of the subsequence so it converges at all rational points by the diagonalisation argument. Since  $\omega_k$  are all non-increasing, this guarantees convergence a.e. In fact, it guarantees convergence except at the countably many jumps of the limiting monotone function.

Let  $\tilde{A}$  be the subsequential limit:

$$\tilde{A} = \{(x, y) : |y| \leq \omega_\infty(|x|) = \lim_{k \rightarrow \infty} \omega_k(|x|)\}$$

This implies  $\chi_{A_k} \rightarrow \chi_{\tilde{A}}$  a.e. for some set  $\tilde{A}$ .

*Remark.*

Once we prove  $|\tilde{A} - A^*| = 0$ , we may deduce that we need never have passed to a subsequence since if a sequence converges, its limit is unique.

To conclude  $\tilde{A} = A^*$ , it suffices to show  $|\tilde{A}|$  is a ball.

Let us consider the strictly symmetric decreasing Gaussian on  $R^2$ ,

$$\gamma(x) = e^{-|x|^2}$$

Let  $a_k = \|\gamma - \chi_{A_k}(x)\|_{L^2}$ . Note  $R\gamma = \gamma$ ,  $X\gamma = \gamma$ , and  $Y\gamma = \gamma$ . So

$$\begin{aligned} a_{k+1} &= \|\gamma - \chi_{A_{k+1}}(x)\|_{L^2} \\ &= \|\gamma - YXR\chi_{A_k}(x)\|_{L^2} \\ &= \|YXR\gamma - YXR\chi_{A_k}(x)\|_{L^2} \\ &\leq \|XR\gamma - XR\chi_{A_k}\|_{L^2} \\ &\leq \|R\gamma - R\chi_{A_k}\|_{L^2} = \|\gamma - \chi_{A_k}\|_{L^2} = a_k \end{aligned}$$

Thus,  $\{a_k\}$  is a decreasing sequence. Because  $a_k \geq 0$ ,  $a = \lim_{k \rightarrow \infty} a_k$  exists. Next, note that

$$\begin{aligned} |a_{k+1} - \|\gamma - YXR\chi_{\tilde{A}}\|_{L^2}| &= \|\|\gamma - \chi_{A_{k+1}}\|_{L^2} - \|\gamma - YXR\chi_{\tilde{A}}\|_{L^2}\| \\ &\leq \|\chi_{A_{k+1}} - YXR\chi_{\tilde{A}}\|_{L^2} \\ &\leq \|YXR\chi_{A_k} - YXR\chi_{\tilde{A}}\|_{L^2} \\ &\leq \|\chi_{A_k} - \chi_{\tilde{A}}\|_{L^2} \rightarrow 0. \end{aligned}$$

which implies

$$a_{k+1} \rightarrow \|\gamma - YXR\chi_{\tilde{A}}\|_{L^2}$$

But  $\{a_k\} \rightarrow a$ , so we conclude that

$$a = \|\gamma - YXR\chi_{\tilde{A}}\|_{L^2}$$

□

**Lecture 15: Riesz rearrangement theorem in  $\mathbb{R}^2$  (continued) and in  $\mathbb{R}^d$** 

Last time we showed that

$$\|\gamma - YXR\chi_{\tilde{A}}\|_{L^2} = a = \|\gamma - \chi_{\tilde{A}}\|_{L^2},$$

where  $\gamma(x) = e^{-|x|^2}$  and  $a = \lim_{k \rightarrow \infty} \|\gamma - \chi_{A_k}\|_{L^2}$ , etc.

Note that

$$\|\gamma - \chi_{\tilde{A}}\|_{L^2} = \|\gamma - R\chi_{\tilde{A}}\|_{L^2}$$

since  $\gamma$  is radially symmetric. It follows that

$$\begin{aligned} \|\gamma - YXR\chi_{\tilde{A}}\|_{L^2} &= \|Y\gamma - YXR\chi_{\tilde{A}}\|_{L^2} \\ &\leq \|\gamma - XR\chi_{\tilde{A}}\|_{L^2} = \|X\gamma - XR\chi_{\tilde{A}}\|_{L^2} \\ &\leq \|\gamma - R\chi_{\tilde{A}}\|_{L^2} = \|\gamma - YXR\chi_{\tilde{A}}\|_{L^2}, \end{aligned}$$

so in fact, equality holds everywhere in this chain of inequalities. By the equality case in the 1-dimensional Riesz rearrangement theorem, we deduce that  $R\chi_{\tilde{A}} = XR\chi_{\tilde{A}}$  and  $YXR\chi_{\tilde{A}} = XR\chi_{\tilde{A}}$ . Thus we have  $YR\chi_{\tilde{A}} = R\chi_{\tilde{A}}$ , which says that  $R\chi_{\tilde{A}}$  is symmetric about the  $x$ -axis: that is,

$$R\chi_{\tilde{A}}(z) = 1 \iff R\chi_{\tilde{A}}(\bar{z}) = 1 \text{ a.e.} \quad (6.20) \quad \boxed{\text{rxa-symm}}$$

By construction,  $\chi_{\tilde{A}}$  has the same property:  $\chi_{\tilde{A}}(z) = 1 \iff \chi_{\tilde{A}}(\bar{z}) = 1$  (indeed, each  $A_k$  has this property). This implies

$$z \in \tilde{A} \iff R\chi_{\tilde{A}}(e^{i\alpha}z) = 1 \iff R\chi_{\tilde{A}}(e^{-i\alpha}\bar{z}) = 1 \iff e^{-2i\alpha}\bar{z} \in \tilde{A} \iff e^{2i\alpha}z \in \tilde{A},$$

where the second equivalence follows from (6.20) and the last follows from the fact that  $\tilde{A}$  is symmetric about the  $x$ -axis.

We have thus shown that

$$\|R^2\chi_{\tilde{A}} - \chi_{\tilde{A}}\|_{L^2} = 0,$$

but we want to prove

$$\mu(\theta) = \|R_\theta\chi_{\tilde{A}} - \chi_{\tilde{A}}\|_{L^2} = 0,$$

where  $R_\theta$  is a rotation by  $\theta \in [0, 2\pi)$ . But we have shown that  $\mu(2n\alpha) = 0$  for  $n = 0, 1, 2, \dots$ . By a theorem of Weyl (see below), these numbers are dense in  $[0, 2\pi)$  (or, in  $\mathbb{R}/2\pi\mathbb{Z}$ ). Moreover,  $\mu$  is continuous, as we can check using the density of  $C_c^\infty(R^d)$  in  $L^2(\mathbb{R}^d)$  (approximate  $\chi_{\tilde{A}}$  by smooth functions). This proves the inequality.

If  $g$  is strictly decreasing and equality holds, then we can use the equality case in the 1-dimensional Riesz theorem to see that all our sets must have agreed a.e. with their rearrangements  $A_k, C_k$  up to a common translation. But since  $A_k \rightarrow A^*, C_k \rightarrow C^*$  (that is, the characteristic functions converge in  $L^2$ ), we have  $|A \Delta A^*| = |C \Delta C^*| = 0$ .  $\square$

**6.1. Aside: Weyl's proof of equidistribution for an irrational rotation.**

**Definition 6.12.** A sequence  $(x_n)$  in  $\mathbb{R}/2\pi\mathbb{Z}$  is said to be *equidistributed* if

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) \rightarrow \frac{1}{2\pi} \int_0^{2\pi} f(\theta) d\theta$$

for all  $f \in C^\infty(\mathbb{R}/2\pi\mathbb{Z})$  [or  $f \in C(\mathbb{R}/2\pi\mathbb{Z})$ ].

Evidently, equidistribution implies that  $(x_n)$  is dense.

**Theorem 6.13** (Weyl's equidistribution theorem). *Suppose  $\alpha$  is an irrational angle (in the sense that  $\alpha \notin 2\pi\mathbb{Q}$ ). Then the sequence  $(n\alpha)$  is equidistributed in  $\mathbb{R}/2\pi\mathbb{Z}$ .*

**weyl-eq**

*Proof.* Since trigonometric polynomials are dense in  $f \in C^\infty(\mathbb{R}/2\pi\mathbb{Z})$  in the supremum norm, it is enough to show the result for such functions. But this further reduces to checking functions of the form  $e^{ikx}$  for  $k \in \mathbb{Z}$ . That is, we need to show

$$\frac{1}{N} \sum_{n=1}^N e^{ikx_n} \rightarrow \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} d\theta = \delta_{k0}$$

as  $N \rightarrow \infty$ , for  $k \in \mathbb{Z}$ . The left-hand side is

$$\frac{1}{N} \sum_{n=1}^N e^{ikn\alpha} = \begin{cases} \frac{1}{N} \frac{e^{ik\alpha} - e^{ik\alpha(N+1)}}{1 - e^{ik\alpha}} & \text{if } k \neq 0 \\ 1 & \text{if } k = 0. \end{cases}$$

Taking  $N \rightarrow \infty$ , we can see that the latter converges to 0 if  $k \neq 0$ , and to 1 if  $k = 0$ , which is the desired result.  $\square$

*Remark.* In fact, Weyl showed that  $2\pi P(n)$  is equidistributed if  $P$  is a nonconstant polynomial with at least one irrational coefficient other than the constant term.

**Theorem 6.14** (Riesz rearrangement theorem in  $\mathbb{R}^d$ ). *If  $f, g, h : \mathbb{R}^d \rightarrow [0, \infty)$  decay at  $\infty$ , we have*

$$I(f, g, h) \leq I(f^*, g^*, h^*).$$

*Moreover, if  $g = g^*$  is strictly decreasing, then equality holds only if  $f(x) = f^*(x-a)$  and  $h(x) = h^*(x-a)$  a.e. for some common translation  $a \in \mathbb{R}^d$ .*

**riesz-rd**

*Proof (by induction on  $d$ ).* We may reduce to bounded sets  $A, B, C$  as previously. Now let  $R$  be some rotation that maps  $e_d \mapsto e_{d-1}$ . Let  $X$  be a Schwarz rearrangement taken separately on each plane perpendicular to  $e_d$  (using the  $e_d$  axis as the origin), and let  $Y$  be Steiner symmetrization parallel with the  $e_d$  axis. The action of  $X$  (on a single slice of the red set) is illustrated below:

Then by the induction hypothesis, we have

$$I(\chi_A, \chi_B, \chi_C) \leq I(YXR\chi_A, YXR\chi_B, YXR\chi_C).$$

We will continue the proof in the next lecture.  $\square$

**Theorem 6.15.** *Let  $f, g, h : \mathbb{R}^d \rightarrow \mathbb{R}$  be non-negative and decreasing at infinity. Then*

$$\int f(x)g(x-y)h(y)dxdy \leq \int f^*(x)g^*(x-y)h^*(y)dxdy \quad (1)$$

*If  $g = g^*$  which is strictly decreasing, then equality occurs only if  $f(x) = f^*(x-a)$  a.e. and  $h(x) = h^*(x-a)$  a.e. for some fixed  $a$ .*

*Proof.* We prove this by induction on  $d$ . As in previous cases, it suffices to show this for  $f, g, h = \chi_A, \chi_B, \chi_C$  indicator functions of measurable sets. Let  $Y$  be symmetrization parallel to the  $e_d$ -axis,  $X$  be Schwarz rearrangement in each plane perpendicular to  $e_d$ , and  $R$  be rotation that takes  $e_d$  to  $e_{d-1}$ . As in the two dimensional case, it follows that

$$I(\chi_A, \chi_B, \chi_C) \leq I(YXR\chi_A, YXR\chi_B, YXR\chi_C).$$

As in the two dimensional case, it suffices to show (1) for  $A, B, C$  are bounded. Noticing that for all  $A$  bounded,  $\chi_{A_{k+1}} = YXR\chi_{A_k}$  all have the form  $A_k = \{(x^\perp, x_n) : |x_n| \leq \omega_k(|x^\perp|)\}$  with  $\omega_k$  measurable and nonincreasing in radius. Since

$A_k$  are bounded,  $\omega_k$  are also bounded, so by Helly selection principle, we may pass to a subsequence to get convergence from  $A_k$  to some  $\tilde{A}$ . Noticing that  $\tilde{A}$  has rotational symmetry over  $e_n$  and reflection symmetry around  $x_n = 0$ , taking inner products with  $\gamma = e^{-|x|^2}$  as in the two dimensional case, it follows that

$$YXR\chi_{\tilde{A}} = XR\chi_{\tilde{A}} = R\chi_{\tilde{A}}.$$

This shows that  $R\chi_{\tilde{A}}$  has rotational symmetry about the  $e_d$  axis and reflection symmetry about  $x_d = 0$ . Thus,  $\tilde{A}$  is rotationally symmetric about  $e_d$  and  $e_{d-1}$ . We will now show that  $\tilde{A}$  is spherically symmetric. As it admits a representation of  $\tilde{A} = \{(x^\perp, x_d) : |x_d| \leq \omega(|x_\perp|)\}$ ,  $\tilde{A}$  being spherically symmetric implies that it is a ball.

Replacing  $\chi_{\tilde{A}}$  with  $\varphi_\epsilon * \chi_{\tilde{A}}$ , where  $\varphi_\epsilon = \epsilon^{-d}\varphi(\cdot/\epsilon)$ ,  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,  $\varphi \geq 0$  and radial, it suffices to show the following: suppose  $u, v \in C_c^\infty(\mathbb{R}^d)$  satisfy

$$u(\sqrt{\rho^2 + x_{d-1}^2}, x_d) = v(\sqrt{\rho^2 + x_d^2}, x_{d-1}) \quad \forall \rho \geq 0, x_d, x_{d-1} \in \mathbb{R}$$

where  $\rho = \sqrt{x_1^2 + \dots + x_{d-2}^2}$ , then  $\psi(x) := u(\sqrt{\rho^2 + x_{d-1}^2}, x_d) = v(\sqrt{\rho^2 + x_d^2}, x_{d-1})$  is spherically symmetric. Then setting  $x_d = 0$ , we see that  $v(\rho, x_{d-1}) = u(\sqrt{\rho^2 + x_{d-1}^2}, 0)$  and hence  $v(\sqrt{\rho^2 + x_d^2}, x_{d-1}) = u(\sqrt{\rho^2 + x_d^2 + x_{d-1}^2}, 0)$ . Hence,  $\psi$  is spherically symmetric. It remains to prove the equality case. If  $I(f, g, h) = I(f^*, g^*, h^*)$  where  $g = g^*$  is strictly decreasing in radius, then by induction,

$$I(f, g, h) \leq I(Xf, g, Xh) \leq I((Xf)^*, g, (Xh)^*) = I(f^*, g, h^*).$$

Hence, by the case of equality in the  $d - 1$ -dimensional Riesz theorem, we see that there exists some  $a \in \mathbb{R}^{d-1}$  such that  $f$  and  $h$  are symmetric about  $\{(a, t) : t \in \mathbb{R}\}$ . Likewise, equality guarantees rotational symmetry about the line parallel with  $e_{d-1}$ . Notice that our two lines must actually intersect, since the orbits of rotating by the two axes must be bounded (since  $f, g, h$  are decreasing at infinity) but orbits of rotating by skewed axes are unbounded (by using a pair of symmetries, we may find infinite sequence of axes of rotation symmetry “marching off to infinity”). Hence, if equality holds,  $f$  and  $h$  must have cylindrical symmetry about two perpendicular axes and hence must be radial.  $\square$

**Lemma 6.16.** Suppose  $f \in L^1_{loc}(\mathbb{R}^d; \mathbb{C})$  and  $\nabla f \in L^1_{loc}(\mathbb{R}^d; \mathbb{C}^d)$ . Then

$$|\nabla f| \geq |\nabla|f||$$

almost everywhere.

*Proof.*  $\nabla|f|$  may be represented as

$$\nabla|f| = \begin{cases} \frac{u\nabla v + v\nabla u}{\sqrt{u^2 + v^2}} & f = u + iv \neq 0 \\ 0 & f = 0 \end{cases}.$$

Let  $\varphi \in C_c^\infty(\mathbb{R}^d)$  be a standard bump function with  $0 \leq \varphi \leq 1$ ,  $\int \varphi = 1$ , and define  $f_\epsilon = f * \varphi_\epsilon$  with  $\varphi_\epsilon = \epsilon^{-d}\varphi(\cdot/\epsilon)$  and given  $\delta > 0$ , we will consider  $\sqrt{\delta^2 + |f_\epsilon|^2}$  in place of  $f$ . The inequality follows from the analogous inequality on  $\sqrt{\delta^2 + |f_\epsilon|^2}$  and by sending  $\delta$  and  $\epsilon$  to zero.  $\square$

*Proof.* Let  $f = u + iv$  for  $u$  and  $v$  real-valued functions. Denote  $u_\varepsilon = u * \phi_\varepsilon$  and similarly for  $v_\varepsilon$ . Linearity of convolution gives  $f_\varepsilon = u + iv_\varepsilon$ .

For every  $\psi \in C_c^\infty$ , we have

$$\int -\nabla \psi \sqrt{\delta^2 + |f_\varepsilon|^2} = \int \psi \frac{u_\varepsilon \nabla u_\varepsilon + v_\varepsilon \nabla v_\varepsilon}{\sqrt{\delta^2 + |f_\varepsilon|^2}} \quad (6.21) \quad \boxed{17.\text{ibp}}$$

using integration by parts. We want to send  $\varepsilon \rightarrow 0$  and then  $\delta \rightarrow 0$ . We will start with the left hand side of (6.21)

We see that

$$\begin{aligned} \left| \int (\nabla \psi) \left[ \sqrt{\delta^2 + |f_\varepsilon|^2} - \sqrt{\delta^2 + |f|^2} \right] dx \right| &\leq \int |\nabla \psi| \frac{| |f_\varepsilon|^2 - |f|^2 |}{\sqrt{\delta^2 + |f_\varepsilon|^2} + \sqrt{\delta^2 + |f|^2}} dx \\ &\leq \int |\nabla \psi| \cdot |f_\varepsilon - f| \frac{|f| + |f_\varepsilon|}{\sqrt{\delta^2 + |f_\varepsilon|^2} + \sqrt{\delta^2 + |f|^2}} dx \\ &\leq \|\nabla \psi\|_{L^\infty} \cdot \|f_\varepsilon - f\|_{L^1}. \end{aligned}$$

This goes to 0 as  $\varepsilon$  goes to 0. Hence the left hand side of (6.21) goes to  $\int -\nabla \psi \sqrt{\delta^2 + |f|^2}$  as  $\varepsilon$  goes to 0. This in turn goes to  $\int -\nabla \psi |f|$  as  $\delta$  goes to 0 by the dominated convergence theorem (since  $\int \nabla \psi (1 + |f|)$  is absolutely integrable).

Now we turn to the right hand side of (6.21).

$$\int \psi \frac{u_\varepsilon \nabla u_\varepsilon + v_\varepsilon \nabla v_\varepsilon}{\sqrt{\delta^2 + |f_\varepsilon|^2}} dx = \int \psi \frac{u_\varepsilon [\nabla u_\varepsilon - \nabla u] + v_\varepsilon [\nabla v_\varepsilon - \nabla v]}{\sqrt{\delta^2 + |f_\varepsilon|^2}} dx + \int \frac{u_\varepsilon \nabla u + v_\varepsilon \nabla v}{\sqrt{\delta^2 + u_\varepsilon^2 + v_\varepsilon^2}} dx.$$

The first term vanishes as  $\varepsilon$  goes down to 0. This is because  $\nabla f_\varepsilon = (\nabla f)_\varepsilon$  converges to  $\nabla f$  in  $L^1$  as  $\varepsilon$  goes to 0 and because

$$\left| \frac{f_\varepsilon}{\sqrt{\delta^2 + |f_\varepsilon|^2}} \right| \lesssim 1$$

uniformly in  $\varepsilon$ . The second term converges as  $\varepsilon$  goes down to 0 to

$$\int \psi \frac{u \nabla u + v \nabla v}{\sqrt{\delta^2 + u^2 + v^2}} dx$$

and then as  $\delta$  goes down to 0 to

$$\int \psi \nabla |f| dx$$

with  $\nabla |f|$  defined above. Both of these convergences are by the Dominated Convergence Theorem. For the first, the integrand is bounded above by uniformly in  $\varepsilon$  by  $|\nabla u| + |\nabla v| \in L^1$ . For the second, the integrand in absolute values is bounded above by  $|\psi| \nabla |f|$  which is integrable because  $\nabla f \in L^1_{loc}$ .

We have shown

$$\int -\nabla \psi |f| = \int \psi \nabla |f|$$

for all  $\psi \in C_c^\infty$ . This is what it means for  $\nabla |f|$  to be a distributional derivative of  $|f|$ .  $\square$

A few remarks on this lemma:

- For an intuitive idea for this theorem, suppose  $f = re^{i\theta}$ . Then  $\nabla f = (\nabla r)e^{i\theta} + i\nabla\theta re^{i\theta}$  while  $\nabla|f| = \nabla r$ . We cut out a perpendicular component (in some sense) by considering  $|f|$  instead of  $f$ , so the norm of the derivative can only decrease.
- We would be able to prove this result using maximal functions if we knew that  $Mf_\varepsilon \in L^1_{\text{loc}}$ . This is true by a result of Gagliardo-Nirenberg:

$$\nabla f \in L^1 \implies f \in L^{d/(d-1)} \implies Mf \in L^{d/(d-1)}.$$

**Theorem 6.17** (Polya-Szegő). *Let  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  be vanishing at infinity. Then*

$$\|\nabla f^*\|_{L^2} \leq \|\nabla f\|_{L^2}.$$

*Remark.* Polya-Szegő proved this in  $L^p$  for all  $1 \leq p \leq \infty$ .

*Proof.* Let us first suppose that  $f \in L^2$  as well. We can use the Fourier transform:

$$\|\nabla f\|_{L^2} = \int |\xi|^2 |\hat{f}(\xi)|^2 d\xi$$

By the Monotone Convergence Theorem, the right hand side of the above is

$$\lim_{t \downarrow 0} \int \frac{1 - e^{-t|\xi|^2}}{t} |\hat{f}(\xi)|^2 d\xi$$

which can be rearranged to become

$$\lim_{t \downarrow 0} \frac{1}{t} \left[ \|f\|_{L^2}^2 - \langle \hat{f}, e^{-t|\xi|^2} \hat{f} \rangle \right].$$

The inner product on the right hand side is equivalent to  $\langle f, e^{t\Delta} f \rangle$ . This allows us to rewrite the entire expression using the heat kernel:

$$\lim_{t \downarrow 0} \frac{1}{t} \left[ \|f\|_{L^2}^2 - \iint \overline{f(x)} \frac{1}{(4\pi t)^{d/2}} e^{-|x-y|^2/(4t)} f(y) dx dy \right].$$

Now we use the Riesz Rearrangement Theorem on the above to conclude that

$$\|\nabla f\|_{L^2} \geq \limsup_{t \downarrow 0} \frac{1}{t} \left[ \|f^*\|_{L^2}^2 - \iint f^*(x) \frac{1}{(4\pi t)^{d/2}} e^{-|x-y|^2/(4t)} f^*(y) dx dy \right].$$

We then undo the previous steps to get

$$\|\nabla f\|_{L^2}^2 \geq \|\nabla f^*\|_{L^2}^2.$$

Of course, we don't actually have that  $f \in L^2$  in general. In the general case, we will proceed as follows.  $\square$

It remains to treat  $f$  vanishing at infinity but not (necessary)  $L^2$ .

First, as  $|\nabla|f|| \leq |\nabla f|$ , we may restrict attention to  $f : \mathbb{R}^d \rightarrow [0, \infty)$ . If we consider

$$\phi(f) = \begin{cases} 0, & 0 \leq f < \epsilon, \\ f - \epsilon, & \epsilon \leq f < A, \\ A - \epsilon, & A \leq f, \end{cases}$$

then

$$\nabla \phi(f)(x) = \begin{cases} \nabla f(x), & \epsilon < f(x) < A, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover  $\phi(f)^* = \phi(f^*)$ . Now as  $f$  vanishes at  $\infty$ , so  $\phi(f)$  is  $L^2$ . Thus

$$\|\nabla[\phi(f)^*]\|_{L^2} \leq \|\nabla \phi(f)\|_{L^2}. \quad (6.22) \quad \boxed{\text{eq:2_17_1}}$$

By monotone convergence theorem, when  $\epsilon \rightarrow 0$  and  $A \rightarrow \infty$ ,

$$\begin{aligned}\|\nabla\phi(f)\|_{L^2} &\rightarrow \|\nabla f\|_{L^2}, \\ \|\nabla[\phi(f)^*]\|_{L^2} &= \|\nabla[\phi(f^*)]\|_{L^2} \rightarrow \|\nabla f^*\|_{L^2}.\end{aligned}$$

Hence by (6.22), we have

$$\|\nabla f^*\|_{L^2} \leq \|\nabla f\|_{L^2}.$$

*Remark.* It is not true that the inequality is strict unless  $f(x) = e^{i\theta}|f(x - x_0)|$ .

For example, a skewed wedding cake.

Brothers-Ziemer showed that if  $|\{x : \nabla f^* = 0\}| = 0$ , then one may conclude that

$$\|\nabla f\| = \|\nabla f^*\| \implies f(x) = e^{i\theta}f^*(x - x_0).$$

One may also consider  $f \rightarrow \hat{f} \rightarrow \hat{f}^* \rightarrow \tilde{f}$ . We may call  $\tilde{f}$  the Fourier symmetrized  $f$ . Clearly

$$\|\nabla \tilde{f}\|_{L^2}^2 = \|\xi|\hat{f}^*\|_{L^2}^2 \leq \|\xi|\hat{f}\|_{L^2}^2 = \|\nabla f\|_{L^2}^2.$$

Warning: This rearrangement is not equimeasurable. Nevertheless the Riesz rearrangement inequality shows

$$\|f\|_{L^4}^4 = [\hat{f} * \hat{f} * \bar{\hat{f}} * \bar{\hat{f}}](0) \leq [\hat{f}^* * \hat{f}^* * \hat{f}^* * \hat{f}^*](0).$$

So  $\|f\|_{L^4}^4 \leq \|\tilde{f}\|_{L^4}^4$ . The analogous result holds for any even  $p$ . It is false for other  $p$ . Cf. Littlewood's majorant conjecture.

For example, the elastic energy stored in a drum. Function  $u(x)$  denotes the height at location  $x$ . The energy is  $\int |\nabla u|^2$ .

The energy in a electro-magnetic field is

$$\int \frac{1}{2}|E|^2 + \frac{1}{2}|B|^2 = \int \frac{1}{2}|\nabla\phi|^2 + \frac{1}{2}|\nabla \times A|^2 dx,$$

where  $\phi$  and  $A$  are electromagnetic potentials.

In quantum mechanics  $\psi : \mathbb{R}^d \rightarrow \mathbb{C}$ , kinetic energy  $= \int \frac{1}{2m}|\nabla\psi|^2 dx$ , where  $m$  is the mass.

One simple application of our rearrangement inequalities is to show that Hydrogen atoms are spherical.

$$\text{Energy}(\psi) = \frac{1}{2} \int_{\mathbb{R}^3} |\nabla\psi|^2 - \frac{1}{|x|}|\psi|^2 dx,$$

where  $\int |\nabla\psi|^2$  is the kinetic energy and  $\int \frac{1}{|x|}|\psi|^2$  refers to the potential energy. We wish to minimize energy subject to the constraint  $\int |\psi|^2 dx = 1$ .

Notice if we replace  $\psi$  by  $\psi^*$  then the energy goes down (unless  $\psi = \psi^*e^{i\theta}$  anyway) and the constraint is still satisfied.

We didn't actually show that our problem admits an optimizer, merely that (a) if there is an optimizer, it is radial,; and (b) in constructing an optimizer, we need only consider radial optimizing sequence.

However if  $\psi_0$  was an optimizer, then

$$\left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \text{Energy} \left( \frac{\psi_0 + \epsilon\varphi}{\sqrt{\int |\psi_0 + \epsilon\varphi|^2}} \right) = 0, \quad \forall \varphi \in C_c^\infty, \varphi \perp \psi_0$$

Integral by parts, we have

$$\begin{aligned} \int \left( -\Delta \psi_0 - \frac{1}{|x|} \psi_0 \right) \bar{\varphi} dx &= 0, \quad \forall \varphi \perp \psi_0 \\ \implies -\Delta \psi_0 - \frac{1}{|x|} \psi_0 &= \lambda \psi_0, \end{aligned}$$

where  $\lambda$  is called "Lagrange multiply" or "eigenvalue". Note that  $\psi_0$  is radial so this is actually an ODE

$$\left( -\frac{\partial}{\partial r^2} - \frac{2}{r} \frac{\partial}{\partial r} - \frac{1}{r} - \lambda \right) \psi_0 = 0. \quad (6.23) \quad \boxed{\text{eq:2_17_2}}$$

No nonzero solution of such an ode can have  $\partial_r \psi_0 = 0$  on a set of positive measure. If it did, there'd be a  $r_0 > 0$  so  $\partial_r \psi_0(r_0) = 0$  and  $\partial_r^2 \psi_0(r_0) = 0$ . According to the ODE (6.23), there should be  $\psi_0(r_0) = 0$ . But by uniqueness for ODE, this implies  $\psi_0 \equiv 0$ .

**Proposition 6.18** (Diamagnetic inequality). *If  $\psi \in H^1(\mathbb{R}^d)$  and  $A : \mathbb{R}^d \rightarrow \mathbb{R}^d$  then*

$$|(\nabla + iA)\psi| \geq |\nabla|\psi||$$

where by the operator  $iA$  we mean  $[(iA)\psi](x) = i\psi(x)A(x)$ .

This is proved similarly to our proof of the  $A = 0$  case.

The reason this proposition is called this way comes from physics: In a magnetic field  $B$  given by  $B = \nabla \times A$ , the Schrödinger equation for an electron takes the form

$$i\partial_t \psi = (\nabla + iA)^*(\nabla + iA)\psi + V\psi$$

The diamagnetic inequality implies that

$$\langle (\nabla + iA)\psi, (\nabla + iA)\psi \rangle \geq \langle \nabla|\psi|, \nabla|\psi| \rangle.$$

It follows that if  $\psi$  is the lowest energy state for an electron in a magnetic field then the expected energy for the wavefunction  $|\psi|$  when there isn't magnetic field is lower than the expected energy of the wavefunction  $\psi$  when there is one. It follows that the lowest energy state with a magnetic field is higher than the lowest energy state without a magnetic field. In physics a material whose energy rises when there is a magnetic field is called "diamagnetic". The diamagnetic inequality seems to imply that every material should be diamagnetic, but this isn't actually true. That is because this model doesn't account for the intrinsic spin of the electron and its interaction with the magnetic field.

When  $B = \nabla \times A = 0$ , then physically we should expect the system to behave like one with  $A = 0$  as they both have zero magnetic field. Indeed, if  $\nabla \times A = 0$  we must have  $A = \nabla\phi$  for some  $\phi$ , and then

$$(\nabla + iA)\psi = e^{-i\phi} \nabla(e^{i\phi}\psi).$$

It follows that  $\psi \mapsto e^{i\phi}\psi$  is a correspondence that relates the Schrödinger equation for this value of  $A$  and  $A = 0$ . This is an example of a gauge symmetry.

7. COMPACTNESS IN  $L^p$  SPACES.

**Theorem 7.1** (Fréchet-Kolmogorov). *For  $1 \leq p < \infty$ , a subset  $\mathcal{F} \subseteq L^p(\mathbb{R}^d)$  is precompact (i.e. its closure is compact) or (equivalently) totally bounded if and only if it satisfies the following:*

- (1)  $\mathcal{F}$  is bounded:

$$\sup_{f \in \mathcal{F}} \|f\|_{L^p} < \infty.$$

- (2)  $\mathcal{F}$  is equicontinuous: For all  $\epsilon > 0$  there exists a  $\delta > 0$  so that

$$\sup_{f \in \mathcal{F}} \sup_{|h| \leq \delta} \|f(x) - f(x-h)\|_{L_x^p} < \epsilon.$$

- (3)  $\mathcal{F}$  is tight: For all  $\epsilon > 0$  there exists a  $R > 0$  so that

$$\sup_{f \in \mathcal{F}} \|f\|_{L^p(|x| \geq R)} < \epsilon.$$

For example, if  $0 \not\equiv f \in L^p$ , then the sets

$$\begin{aligned}\mathcal{F}_0 &= \{x \mapsto f(x-y) : y \in \mathbb{R}^d\} \\ \mathcal{F}_1 &= \{x \mapsto f(x)e^{i\xi \cdot x} : \xi \in \mathbb{R}^d\} \\ \mathcal{F}_2 &= \{x \mapsto \lambda^{\frac{d}{p}} f(\lambda x) : 0 < \lambda \leq 1\} \\ \mathcal{F}_3 &= \{x \mapsto \lambda^{\frac{d}{p}} f(\lambda x) : 1 \leq \lambda < \infty\}\end{aligned}$$

are all bounded, but none are compact.  $\mathcal{F}_0$  and  $\mathcal{F}_2$  are equicontinuous but not tight, while  $\mathcal{F}_1$  and  $\mathcal{F}_3$  are tight but not equicontinuous.

*Proof.* Precompactness  $\implies$  (a): This is obvious.

Precompactness  $\implies$  (b) & (c): Since  $\mathcal{F}$  is precompact, for every  $\epsilon > 0$  there are functions  $f_1, \dots, f_N$  such that

$$\mathcal{F} \subseteq \bigcup_{i=1}^N B(f_i, \frac{\epsilon}{10}).$$

Now, for each fixed  $i$ , we have

$$\|f_i\|_{L^p(|x| \geq R)} \rightarrow 0 \text{ as } R \rightarrow \infty$$

by the Dominated Convergence Theorem. Thus there exists an  $R$  so that

$$\|f_i\|_{L^p(|x| \geq R)} \leq \frac{\epsilon}{10} \text{ for all } 1 \leq i \leq N.$$

Then for all  $f \in \mathcal{F}$ , there is an  $i$  so that

$$\|f\|_{L^p(|x| \geq R)} \leq \|f - f_i\|_{L^p} + \|f_i\|_{L^p(|x| \geq R)} \leq \frac{\epsilon}{10} + \frac{\epsilon}{10} < \epsilon.$$

Similarly, for each  $i$  we have

$$\|f_i(x+h) - f_i(x)\|_{L_x^p} \rightarrow 0 \text{ as } h \rightarrow 0$$

by a Theorem of Lebesgue (which can be proven by approximating  $f_i$  with a  $C_c^\infty$  function). Using the same trick, for each  $f \in \mathcal{F}$  we have

$$\|f(x+h) - f(x)\|_{L_x^p} \leq \|f_i(x+h) - f_i(x)\|_{L_x^p} + 2\|f - f_i\|_{L^p} < \epsilon$$

so we see that  $\mathcal{F}$  is equicontinuous.

It remains to show that (a), (b), and (c) imply precompactness. We will show that  $\mathcal{F}$  is totally bounded, by exhibiting a finite  $\epsilon$ -covering where  $\epsilon > 0$  is arbitrary. By property (c) there is an  $R$  such that

$$\sup_{f \in \mathcal{F}} \|f - f\phi_R\|_{L^p} \leq \frac{\epsilon}{10}$$

where  $\phi_R$  is a fixed smooth cutoff satisfying

$$\begin{aligned}\phi_R(x) &= 1 && \text{if } |x| \leq R, \\ \phi_R(x) &= 0 && \text{if } |x| \geq 2R.\end{aligned}$$

Notice that  $\{f\phi_R : f \in \mathcal{F}\}$  inherits properties (a) and (b).

Now consider mollification by  $\phi_\delta(x) := \delta^d \phi(x/\delta)$  with, say,  $0 \leq \phi \in C^\infty(\mathbb{R}^d)$  with  $\text{supp}(\phi) \subseteq B(0, 1)$  and  $\int \phi = 1$ .

Now,

$$\begin{aligned}\|f - \phi_\delta * f\|_{L^p} &= \left\| \int [f(x) - f(x+h)] \phi_\delta(h) dh \right\|_{L_x^p} \\ &\leq \int \|f(x) - f(x+h)\|_{L_x^p} \phi_\delta(h) dh \quad (\text{supported on } |h| < \delta) \\ &\leq \sup_{|h| < \delta} \|f(x+h) - f(x)\|_{L_x^p}.\end{aligned}$$

Thus property (b) guarantees that

$$\sup_{f \in \mathcal{F}} \| [f\phi_R] - [f\phi_R] * \phi_\delta \| \rightarrow 0 \quad \text{as } \delta \rightarrow 0.$$

Thus there exists a  $\delta > 0$  such that this quantity is less than  $\frac{\epsilon}{10}$ .

Finally, observe that  $\mathcal{F}_{R,\delta} = \{[f \cdot \phi_R] * \phi_\delta\}$  consists of continuous functions supported on  $B(0, 2R + \delta)$ , and we can apply the Arzela-Ascoli Theorem to show that  $\mathcal{F}_{R,\delta}$  is totally bounded in the  $C(B(0, 2R + \delta))$  metric. It is hence also totally bounded in the  $L^p(B(0, 2R + \delta))$  metric since

$$\|h\|_{L^p(B(0, 2R + \delta))} \lesssim \|h\|_{C(B(0, 2R + \delta))}.$$

Moreover, it also satisfies that for every  $f \in \mathcal{F}$  there is a  $g \in \mathcal{F}_{R,\delta}$  such that

$$\|f - g\|_{L^p} \leq \frac{\epsilon}{10} + \frac{\epsilon}{10}.$$

Thus, a  $\frac{\epsilon}{10}$ -covering of  $\mathcal{F}_{R,\delta}$  can be transported into a  $\epsilon$ -covering of  $\mathcal{F}$ . This shows that  $\mathcal{F}$  is totally bounded.  $\square$

**Corollary 7.2.**  $\mathcal{F} \subset L^2(\mathbb{R}^d)$  is precompact if and only if

- a:  $\mathcal{F}$  is bounded.
- b:  $\{\hat{f} : f \in \mathcal{F}\}$  is tight.
- c:  $\mathcal{F}$  is tight.

*Proof.* We do the forward direction first. By the above theorem we have (a) and (c). If  $\mathcal{F}$  is precompact then so is  $\hat{\mathcal{F}} = \{\hat{f} : f \in \mathcal{F}\}$  (since the Fourier transform is an isometry) and so we get that this set is tight.

Next we do the backward direction. By the above theorem we just need to show that (a), (b), (c) imply equicontinuity. For  $f \in \mathcal{F}$  we have

$$\|f(x+h) - f(x)\|_{L^2}^2 = \|e^{i\xi h} \hat{f}(\xi) - \hat{f}(\xi)\|_{L^2}^2$$

$$\begin{aligned}
&= \int |e^{i\xi h} - 1|^2 |\hat{f}(\xi)|^2 d\xi \\
&\leq \int_{|\xi| \leq N} |\xi|^2 h^2 |\hat{f}(\xi)|^2 + \int_{|\xi| \geq N} 4 |\hat{f}(\xi)|^2 d\xi \\
&\leq \int_{|\xi| \leq N} N^2 h^2 |\hat{f}(\xi)|^2 + \int_{|\xi| \geq N} 4 |\hat{f}(\xi)|^2 d\xi
\end{aligned}$$

Given  $\epsilon > 0$  condition (b) says

$$\sup_{f \in \mathcal{F}} \int_{|\xi| \geq N} |\hat{f}(\xi)|^2 d\xi < \frac{\epsilon^2}{10}$$

provided that  $N \geq N_0(\epsilon)$ . By (a)

$$\sup_{f \in \mathcal{F}} \int_{|\xi| \leq N} N^2 h^2 |\hat{f}(\xi)|^2 d\xi = O(h^2 N^2)$$

in particular given  $N$  there exists  $\delta > 0$  so that

$$\sup_{f \in \mathcal{F}} \int_{|\xi| \leq N} h^2 N^2 |\hat{f}(\xi)|^2 d\xi \leq \frac{\epsilon}{10}$$

whenever  $|h| < \delta$ . □

A linear map  $T : X \rightarrow Y$ , where  $X, Y$  are Banach spaces, is said to be *compact* if  $T(B_X(0, 1))$  is precompact in  $Y$ . The map is called *completely continuous* if and only if for every sequence  $x_n$ , that is weakly convergent in  $X$ , the sequence  $T(x_n)$  is norm convergent in  $Y$ . (Warning, in old books the names are confused).

**Lemma 7.3.** a: *Compact implies completely continuous.*

b: *If  $X$  is reflexive then completely continuous implies compact.*

c: *The identity  $i : l^1 \rightarrow l^1$  is completely continuous but not compact.*

*Proof.* a: If  $x_n \rightharpoonup x$  then it is bounded by the Principle of Uniform Boundedness. Thus  $\{T(x_n)\}$  lives in a compact set. In particular every subsequence has a convergent subsubsequence. Moreover given  $l \in Y^*$  we have

$$\langle l, T(x_n) \rangle = \langle T^* l, x_n \rangle \rightarrow \langle T^* l, x \rangle$$

Thus  $T(x_n)$  converges weakly and consequently all convergent subsequences converge to the same limit (Hahn-Banach). Thus (Urysohn principle)  $T(x_n)$  is norm convergent.

b: Let  $x_n$  be a sequence in  $B(0, 1)$ . We must prove that  $T(x_n)$  admits a norm-convergent subsequence.

Notice that  $\text{span}\{x_n\}$  is a separable subspace of  $X$ . Thus the Banach-Alaoglu theorem says that  $x_n$  admits a weakly convergent subsequence,  $y_n$  (as  $X$  is reflexive the weak and weak-\* topologies are equivalent). Then as  $T$  is completely continuous  $T(y_n)$  converges in norm.

c: This is a result of Schur; see below. As  $l^1$  is infinite dimensional the unit ball is not compact. To show that  $i$  is completely continuous, we must show that weakly convergent sequences in  $l^1$  are norm convergent. Clearly weak convergence in  $l^1$  implies entry-wise convergence, the obstacle to overcome is to show that weak-convergence implies tightness... . □

**Proposition 7.4** (Schur). *The space  $\ell^1(\mathbb{N})$  has SCHUR'S PROPERTY: every weakly convergent sequence is norm convergent.*

*Proof.* By linearity it suffices to show that if  $\{a^{(k)}\}_{k \in \mathbb{N}} \subseteq \ell^1(\mathbb{N})$  converge weakly to the zero sequence, then it converges in norm. Assume towards a contradiction otherwise, i.e. there exists  $\epsilon > 0$  such that

$$\sum_{n \in \mathbb{N}} |a_n^{(k)}| \geq 3\epsilon$$

for infinitely many  $k \in \mathbb{N}$ . Pass to a subsequence and re-index such that the above holds for all  $k \in \mathbb{N}$ . We follow the sliding hump argument to exhibit a subsequence that does not converge weakly to zero, a contradiction. Define the bounded sequence  $\{b_n\}_{n \in \mathbb{N}} \subseteq \{-1, 1\}$  inductively: by summability we can find  $n_0 \in \mathbb{N}$  such that

$$\sum_{n > n_0} |a_n^{(0)}| < \epsilon.$$

Set  $b_n = \operatorname{sgn} a_n^{(0)}$  for  $1 \leq n \leq n_0$ , then observe that

$$\left| \sum_{n \in \mathbb{N}} a_n^{(0)} b_n \right| \geq \sum_{n \leq n_0} |a_n^{(0)}| - \sum_{n > n_0} |a_n^{(0)}| \geq \epsilon.$$

By weak convergence and summability, there exists  $k_1 \in \mathbb{N}$  and  $n_1 > n_0$  such that

$$\sum_{n \leq n_0} a_n^{(k_1)} b_n < \epsilon, \quad \sum_{n > n_1} |a_n^{(k_1)}| < \epsilon.$$

Set  $b_n = \operatorname{sgn} a_n^{(k_1)}$  for  $n_0 < n \leq n_1$ , then

$$\left| \sum_{n \in \mathbb{N}} a_n^{(k_1)} b_n \right| \geq \sum_{n_0 < n \leq n_1} |a_n^{(k_1)}| - \sum_{n \leq n_0} |a_n^{(k_1)}| - \sum_{n > n_1} |a_n^{(k_1)}| \geq \epsilon.$$

In essence, we use weak convergence to extract an index  $k$  where the head of  $a^{(k)}$  is controlled, use summability to control the tail, and choose signs to collect the remaining mass in the middle. Proceeding inductively gives a subsequence  $\{a^{(k_j)}\}_{j \in \mathbb{N}}$  and a bounded sequence  $\{b_n\}_{n \in \mathbb{N}}$  such that

$$\left| \sum_{n \in \mathbb{N}} a_n^{(k_j)} b_n \right| \geq \epsilon$$

for all  $j \in \mathbb{N}$ , so the subsequence does not converge weakly to zero. This completes the proof.  $\square$

We will mostly be interested in the inclusions:

$$i : H^1 \hookrightarrow L^p$$

Actually this inclusion is seldom compact and we will employ powerful ideas to remedy that.

**Proposition 7.5** (Gagliardo-Nirenberg). *The inclusion*

$$i : H^1 \hookrightarrow L^p$$

*is bounded if and only if*

- 1:  $d = 1$  and  $2 \leq p \leq \infty$ .
- 2:  $d = 2$  and  $2 \leq p < \infty$ .
- 3:  $d \geq 3$  and  $2 \leq p \leq 2^* = \frac{2d}{d-2}$ .

*Proof.* For dimensions  $d \geq 3$ , this follows from Sobolev embedding:

$$\|f\|_{L^{\frac{2d}{d-2}}} \lesssim \|\nabla f\|_{L^2}$$

together with interpolation.

Excepting the Sobolev endpoint and  $p = 2$ , a simpler argument suffices.

$$\begin{aligned} \|f\|_{L^p} &\leq \|f_{\leq N}\|_{L^p} + \sum_{M \geq N} \|f_M\|_{L^p} \\ &\leq N^{\frac{d}{2} - \frac{d}{p}} \|f\|_{L^2} + \sum_{M \geq N} M^{\frac{d}{2} - \frac{d}{p} - 1} \cdot M \|f_M\|_{L^2} \\ &\lesssim N^{\frac{d}{2} - \frac{d}{p}} \|f\|_{L^2} + \sum_{M \geq N} M^{\frac{d}{2} - \frac{d}{p} - 1} \cdot \|\nabla f\|_{L^2} \end{aligned}$$

As  $2 < p < \frac{2d}{d-2}$  so  $\frac{d}{2} - \frac{d}{p} > 0$  and  $\frac{d}{2} - \frac{d}{p} - 1 < 0$ . Hence continuing from above we get

$$\|f\|_{L^p} \lesssim N^{\frac{d}{2} - \frac{d}{p}} \|f\|_{L^2} + \sum_{M \geq N} M^{\frac{d}{2} - \frac{d}{p} - 1} \cdot \|\nabla f\|_{L^2} \lesssim N^{\frac{d}{2} - \frac{d}{p}} \|f\|_{L^2} + N^{\frac{d}{2} - \frac{d}{p} - 1} \cdot \|\nabla f\|_{L^2}$$

After optimizing the choice of  $N$  we get

$$\|f\|_{L^p} \lesssim \|f\|_{L^2}^{\frac{2d-p(d-2)}{2p}} \|\nabla f\|_{L^2}^{\frac{d(p-2)}{2p}} \lesssim \|f\|_{H^1}.$$

□

For  $d \geq 3$ , the Sobolev embedding lemma

$$\|f\|_{L^{\frac{2d}{d-2}}} \lesssim \|\nabla f\|_{L^2}$$

guarantees that  $\dot{H}^1(\mathbb{R}^d)$  is a space of functions. However, this fails to hold for  $d = 1, 2$  as we can construct sequences of continuous functions which diverge to infinity a.e. however the gradients converge to zero in  $L^2(\mathbb{R}^d; \mathbb{R}^d)$ .

**Proposition 7.6.** *For  $d = 1, 2$ , there exist continuous functions  $f_n : \mathbb{R}^d \rightarrow \mathbb{R}$  of compact support such that  $f_n \rightarrow \infty$  a.e. and*

$$\|\nabla f_n\|_{L^2} \rightarrow 0.$$

*Proof.* Let  $h > 0$  and  $0 < R_1 < R_2$ , to be chosen later. For  $d = 1$ , we consider the “trapezoid” function

$$f(x) = \begin{cases} \frac{h}{R_2 - R_1}(x + R_2), & \text{if } x \in [-R_2, -R_1], \\ h, & \text{if } x \in [-R_1, R_1], \\ \frac{h}{R_1 - R_2}(x - R_2), & \text{if } x \in [R_1, R_2], \end{cases}$$

Then

$$\|\nabla f\|_{L^2}^2 = \left| \frac{h}{R_2 - R_1} \right|^2 \cdot 2|R_2 - R_1| = \frac{2h^2}{R_2 - R_1}.$$

Choosing, for example,  $R_1 = h^3$  and  $R_2 = 2h^3$ , we see that  $f$  satisfies the desired properties passing the limit as  $h \rightarrow \infty$ .

For  $d = 2$ , we consider the radial piecewise harmonic function

$$f(x) = \begin{cases} h, & \text{if } |x| < R_1, \\ 0, & \text{if } |x| > R_2, \\ h \frac{\log(|x|/\log R_2)}{\log(R_1/R_2)}, & \text{if } R_1 \leq |x| \leq R_2. \end{cases}$$

We compute the norm of the Laplacian converting the polar coordinates,

$$\|\nabla f\|_{L^2}^2 = 2\pi \int_{R_1}^{R_2} \left| \frac{h}{\log(R_1/R_2)} \right|^2 \frac{1}{r} dr = 2\pi \frac{h^2}{\log(R_2/R_1)}.$$

Choosing  $h = \log(R_2/R_1)^{\frac{1}{3}}$  and  $R_2 = R_1^2$ , observe that  $f$  satisfies the desired properties passing the limit as  $R_1 \rightarrow \infty$ .  $\square$

*Remark.* One proving the 2-dimensional case, one might naively attempt to generalize from the 1-dimensional case by symmetrizing the “trapezoid” function, i.e. choosing the frustum of a cone. Unfortunately, this idea does not work, as the crux of the proof depends on the (piecewise)-harmonicity of  $f$ . In one dimension, straight lines are harmonic, however  $|x|$  fails to be harmonic in two dimensions. In the proof, we are essentially interested in minimizing the *Dirichlet energy*

$$\|u\|_{L^2(\Omega)}^2 = \int_{\Omega} |\nabla u|^2 dx,$$

of which harmonic functions are minimizers, with given appropriate boundary conditions on the domain  $\Omega \subseteq \mathbb{R}^2$ . Indeed, arguing by first variation, for any  $\phi \in C_c^\infty(\Omega)$  we have

$$\begin{aligned} \frac{d}{dt} \Big|_{t=0} \int_{\Omega} |\nabla(u + t\phi)|^2 dx &= \frac{d}{dt} \Big|_{t=0} \int_{\Omega} (\nabla u \cdot \nabla u + 2t\nabla u \cdot \nabla\phi + t^2 \nabla\phi \cdot \nabla\phi) dx \\ &= 2 \int_{\Omega} \nabla u \cdot \nabla\phi dx = -2 \int_{\Omega} \Delta u \phi dx. \end{aligned}$$

Thus if  $u$  is a minimizer of the Dirichlet energy with given boundary conditions, then

$$\int_{\Omega} \Delta u \phi dx = 0$$

for all  $\phi \in C_c^\infty(\Omega)$ , i.e.  $u$  is distributionally harmonic and therefore, by Weyl’s lemma, smooth and harmonic.

**Theorem 7.7** (Gagliardo-Nirenberg inequality). *If  $d = 1$  and  $2 \leq p \leq \infty$ , or  $d = 2$  and  $2 \leq p < \infty$ , or  $d \geq 3$  and  $2 \leq p \leq \frac{2d}{d-2}$ , then*

$$\|f\|_{L^p} \lesssim \|f\|_{L^2}^{\frac{2d-p(d-2)}{2p}} \|\nabla f\|_{L^2}^{\frac{d(p-2)}{p}}.$$

In particular, if  $d = 2$  and  $2 < p < \infty$  or  $d \geq 3$  and  $2 < p < \frac{2d}{d-2}$ , then

$$H_{rad}^1(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$$

compactly, where  $H_{rad}^1(\mathbb{R}^d)$  are the radial functions in  $H^1(\mathbb{R}^d)$ .

*Remark.*

- The case  $d = 1$  must be omitted. If we consider  $f_n(x) = \varphi(x-n) + \varphi(-x-n)$  where  $\varphi \in C_C^\infty$ , then such a sequence is clearly bounded in  $H^1(\mathbb{R})$  and is radial since it is even. But is not compact in  $L^p$  for any  $1 \leq p \leq \infty$ .
- $p = 2$  must be excluded in all dimensions. Pick  $\varphi \in C_C^\infty$  and  $\varphi$  is radial and not zero function. Then let  $f_\lambda(x) = \lambda^{-d/2} \varphi(\frac{x}{\lambda})$  which satisfies  $\int |f_\lambda|^2 dx = \int \varphi^2 dx$  and  $\int |\nabla f_\lambda|^2 dx = \frac{1}{\lambda^2} \int |\nabla \varphi|^2 dx$ . The family  $\{f_\lambda : \lambda \geq 1\}$  is not tight in  $L^2$  since  $\int_{|x| \geq R} |f_\lambda(x)|^2 dx = \int_{|y| \geq R/\lambda} |\varphi(y)|^2 dy$ . So it is not compact but  $\|f_\lambda\|_{H^1}^2 = \|\varphi\|_{L^2}^2 + \lambda^{-2} \|\nabla \varphi\|_{L^2}^2 \lesssim 1$  as  $\lambda \geq 1$ .
- $p = \infty, d = 2$  is forbidden because  $H_{rad}^1$  does not embed into  $L^\infty$ .

- $p = \frac{2d}{d-2}$ ,  $d \geq 3$  is forbidden by considering  $\{f_\lambda : \lambda \leq 1\}$  and define

$$f_\lambda(x) = \lambda^{\frac{1-d}{2}} \varphi(x/\lambda),$$

where  $\varphi \in C_C^\infty$  and  $\varphi$  is radial and not zero function. This family is bounded in  $H^1$  but not equicontinuous. We see  $f_\lambda$  is supported on a ball of radius  $O(|\lambda|)$ , so if  $|\lambda| \ll h$ , then

$$\|f_\lambda(x-h) - f_\lambda(x)\|_{L^{\frac{2d}{d-2}}}^{\frac{2d}{d-2}} = 2\|f_\lambda\|_{L^{\frac{2d}{d-2}}}^{\frac{2d}{d-2}}$$

**Lemma 7.8** (Radial Sobolev embedding). *Suppose  $f \in H^1 \cap L^q(R^d)$  is radial for some  $1 \leq q < \infty$ . Then*

$$|x|^{\frac{2(d-1)}{2-q}} |f(x)| \lesssim \|\nabla f\|_{L^2}^{\frac{q}{2+q}} \|f\|_{L^q}^{\frac{2}{2+q}}.$$

For  $d \geq 2$ , this shows that  $f$  decays pointwise as  $|x| \rightarrow \infty$ .

*Proof.* Let  $f \in C_C^\infty$  and write it as  $f(r), r = |x|$

$$\begin{aligned} r^{d-1} |f(r)|^{\frac{q+2}{2}} &\leq r^{d-1} \int_r^\infty |f'(s)|(1 + \frac{q}{2}) |f(s)|^{\frac{q}{2}} ds && \text{(FTC)} \\ &\lesssim \int_r^\infty |\nabla f(s)| |f(s)|^{\frac{q}{2}} s^{d-1} ds \\ &\lesssim \|\nabla f\|_{L^2} \|f\|_{L^q}^{\frac{q}{2}}. && \text{(Cauchy-Schwartz)} \end{aligned}$$

So

$$r^{\frac{2(d-1)}{2-q}} |f| \lesssim \|\nabla f\|_{L^2}^{\frac{q}{2+q}} \|f\|_{L^q}^{\frac{2}{2+q}}$$

□

*Proof of compactness in radial Gagliardo–Nirenberg.* We need to show the unit ball in  $H_{rad}^1$  is compact in  $L^p$ . This requires three things:

- Boundedness follows from Gagliardo–Nirenberg inequality.
- Equicontinuity: Let  $f \in C_C^\infty$ , then

$$\begin{aligned} |f(x+h) - f(x)| &= \left| \int_0^1 h \cdot \nabla f(x+\theta h) d\theta \right| \\ &\leq \int_0^1 |h| |\nabla f(x+\theta h)| d\theta \end{aligned}$$

Choose  $q$  so that  $p < q < \frac{2d}{d-2}$ ,

$$\begin{aligned} \|f(x+h) - f(x)\|_{L^p} &\leq \left\| \int_0^1 |h| |\nabla f(x+\theta h)| d\theta \right\|_{L^p}^p \|f(x+h) - f(x)\|_{L^q}^{1-\varphi} \\ &\leq \left[ \int_0^1 |h| |\nabla f(x+\theta h)|_{L_x^2}^\varphi [2\|f\|_{L^q}]^{1-\varphi} d\theta \right]^\frac{1}{p} && \text{(Minkowski)} \\ &\lesssim |h|^\varphi \|\nabla f\|_{L^2}^\varphi \|f\|_{H^1}^{1-\varphi} && \text{(Gagliardo–Nirenberg)} \\ &\lesssim |h|^\varphi \|f\|_{H^1} \end{aligned}$$

- Tightness: As  $2 < p < \infty$ ,

$$\begin{aligned} \|f\|_{L^p(|x| \geq R)} &\leq \|f\|_{L^2(|x| \geq R)}^\varphi \|f\|_{L^\infty(|x| \geq R)}^{1-\varphi} \\ &\lesssim \|f\|_{H^1}^\varphi [R^{-\frac{d-1}{2}} \|f\|_{H^1}]^{1-\varphi} && \text{(radial Sobolev with } q = 2) \end{aligned}$$

$$\lesssim \|f\|_{H^1} R^{-(1-\varphi)\frac{d-1}{2}}$$

This converges to 0 as  $R \rightarrow \infty$  on  $H_{rad}^1$  bounded sets.

□

**Lemma 7.9.** *The Gagliardo–Nirenberg inequality,  $2 < p < \frac{2d}{d-2}$ ,  $d \geq 2$  and  $1 < p < \infty$ ,  $d = 1$  admits optimizers.*

Example:  $H^1(R^3) \hookrightarrow L^6$  but  $\|f\|_{L^6} \lesssim \|f\|_{H^1}$  does not admit optimizer. However Sobolev embedding  $\|f\|_{L^6} \lesssim \|\nabla f\|_{L^2}$  does admit optimizers, say  $W$ , A ‘good’ optimizing sequence for  $\|f\|_{L^6} \lesssim \|f\|_{H^1}$  is  $f_n(x) = W(nx)e^{-|x|^2}$ .

**Corollary 7.10.** *{ $f^* : f \in H^1(\mathbb{R}^d)$  &  $\|f\|_{H^1} \leq 1$ } is a compact subset of  $L^p(\mathbb{R}^d)$  for  $2 < p \leq \infty$  if  $d = 1$ ; for  $2 < p < \infty$  if  $d = 2$  and for  $2 < p < \frac{2d}{d-2}$  if  $d \geq 3$ .*

*Proof.* For  $d \geq 2$ , this follows from the above since  $f^*$  is radial and the fact that

$$\|f^*\|_{H^1} \leq \|f\|_{H^1}. \quad (7.1)$$

For  $d = 1$ , we exploit monotonicity:

$$\int_{\frac{R}{2} \leq \|x\| \leq R} |f|^2 \leq 1. \quad (7.2)$$

Thus there is a Lebesgue point  $x_0$  with  $\frac{R}{2} \leq |x_0| \leq R$  so that

$$|f^*(x_0)| \lesssim \frac{2}{\sqrt{R}}. \quad (7.3)$$

But then since  $f^*$  is decreasing, we deduce that

$$|f^*(x)| \leq \frac{2}{\sqrt{|x|}}. \quad (7.4)$$

This tells us that

$$\{f^* : \|f\|_{L^2} \leq 1\} \quad (7.5)$$

is tight in  $L^p$ ,  $2 < p \leq \infty$ :  $\|f\|_{L^\infty(|x| \geq r)} \lesssim \frac{1}{\sqrt{r}}$ .

$$\|f\|_{L^p(|x| \geq r)} \lesssim \left(2 \int_r^\infty \left(\frac{1}{\sqrt{|x|}}\right)^p dx\right)^{\frac{1}{p}} \quad (7.6)$$

$$\lesssim |r|^{-\frac{p-2}{2p}}. \quad (7.7)$$

To finish we need boundedness:

$$|f(x)|^2 \leq \int_x^\infty |f'(y)f(y)| dy \quad (7.8)$$

$$\leq \|f'\|_{L^2} \|f\|_{L^2} \leq \|f\|_{H^1}^2, \quad (7.9)$$

as well as equicontinuity:

$$|f(x+h) - f(x)| \leq \int_x^{x+h} |f'(y)| dy \quad (7.10)$$

$$\leq \sqrt{h} \|f\|_{H^1}. \quad (7.11)$$

(baby Morrey inequality) □

**Lemma 7.11.** *The Gagliardo-Nirenberg inequality*

$$\|f\|_{L^p} \lesssim \|f\|_{L^2}^{\frac{2d-p(d-2)}{2p}} \|\nabla f\|_{L^2}^{\frac{d(p-2)}{2p}} \quad (7.12)$$

for  $2 < p < \frac{2d}{d-2}$ ,  $d \geq 2$  &  $2 < p < \infty$  for  $d = 1$  admits optimizers.

*Proof.* Clearly there is a sequence  $f_n \in H^1(\mathbb{R}^d) \setminus \{0\}$  so that

$$\frac{\|f_n\|_{L^p}}{\|f_n\|_{L^2}^\theta \|\nabla f_n\|_{L^2}^{1-\theta}} \rightarrow \sup_{g \in H^1 \setminus \{0\}} \frac{\|g\|_{L^p}}{\|g\|_{L^2}^\theta \|\nabla g\|_{L^2}^{1-\theta}}. \quad (7.13)$$

Rearrangement inequalities show  $\|f_n^*\|_{L^p} = \|f\|_{L^p}$  &  $\|f_n^*\|_{L^2} = \|f_n\|_{L^2}$  &  $\|\nabla f_n^*\|_{L^2} \leq \|\nabla f_n\|_{L^2}$ , and so  $J(f_n^*) \geq J(f_n)$ . Thus we may replace the optimizing sequence  $f_n$  by  $f_n^*$ , i.e., we may assume  $f_n = f_n^*$ .

Notice that if  $\tilde{f}$  is given by  $\tilde{f}(x) = \alpha f(\lambda x)$  with  $\lambda > 0$  &  $\alpha \neq 0$ , then

$$J(\tilde{f}) = J(f). \quad (7.14)$$

By exploiting this symmetry, we may replace the original sequence  $f_n$  by a sequence satisfying

$$\|f_n\|_{L^2} = \|\nabla f_n\|_{L^2} = 1. \quad (7.15)$$

Incidentally, if  $g$  is an optimizer, then likewise

$$x \rightarrow \alpha g(\lambda[x - y]) \quad (7.16)$$

is an optimizer.

By our corollary 7.10 above,  $\{f_n\}$  is precompact in  $L^p$  and so admits a convergent subsequence. Passing to a further subsequence, the Banach-Alaoglu Theorem allows us to assume that

$$f_n \rightharpoonup f \in H^1. \quad (7.17)$$

Notice that

- (a)  $f$  is the  $L^p$  limit
- (b)  $\|f\|_{L^2} \leq \liminf \|f_n\|_{L^2} = 1$ ,  $\|\nabla f\|_{L^2} \leq \liminf \|\nabla f_n\|_{L^2} = 1$ .

This is the weak low-semicontinuity of norms  $x_n \rightharpoonup x$  in  $X$ .

Then for all  $\ell \in X^*$

$$\langle \ell, x \rangle = \lim_{n \rightarrow \infty} \langle \ell, x_n \rangle \leq \|\ell\| \liminf_{n \rightarrow \infty} \|x_n\|. \quad (7.18)$$

So taking a sup over  $\ell \neq 0$

$$\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|. \quad (7.19)$$

Notice that

$$\|f\|_{L^p} = \lim_{n \rightarrow \infty} \|f_n\|_{L^p} = J_{\max} \neq 0 \quad (7.20)$$

and by (a), notice that

$$J_{\max} \geq J(f) = \frac{J_{\max}}{\|f\|_{L^2}^\theta \|\nabla f\|_{L^2}^{1-\theta}} \quad (7.21)$$

i.e.,

$$\|f\|_{L^2}^\theta \|\nabla f\|_{L^2}^{1-\theta} \geq 1. \quad (7.22)$$

So the possibility of strict inequality is averted. That is

$$1 = \|f_n\|_{L^2} \rightarrow \|f\|_{L^2} \quad \text{and} \quad 1 = \|\nabla f_n\|_{L^2} \rightarrow \|\nabla f\|_{L^2}. \quad (7.23)$$

By the Radon-Riesz Theorem these conclusions show that

$$f_n \rightarrow f \text{ in } H^1. \quad (7.24)$$

So not only is  $f$  an optimizer but our sequence  $f_n$  converges to it in  $H^1$ .

Radon-Riesz Theorem

$$\begin{cases} f_n \rightharpoonup f, \text{ in } X \\ \|f_n\|_X \rightarrow \|f\|_X \end{cases} \quad (7.25)$$

RRT is true if  $X$  is a Hilbert space or if  $X = L^p(\Omega, d\mu)$  with  $1 < p < \infty$ . The Radon-Riesz property holds in any uniformly convex Banach space. The subtext is that  $L^p$  is uniformly convex for  $1 < p < \infty$ .  $\square$

**Proposition 7.12.** *For  $d = 1$ ,  $2 < p < \infty$  or  $d \geq 2$ ,  $2 < p < \frac{2d}{2d-2}$ , the PDE*

$$-\Delta Q - Q^{p-1} = -Q \quad (7.26) \quad \boxed{\text{PDE}}$$

*has a unique positive radial  $H^1(\mathbb{R}^d)$  solution.*

*Proof. Case 1:  $d = 1$ .*

In this case, the PDE becomes an ODE, and we can solve it explicitly. Let  $Q = q(x)$ , then equation (7.26) becomes

$$-q'' - q^{p+1} = -q \quad (7.27) \quad \boxed{\text{ODE1}}$$

which can be written as

$$\frac{d}{dx} \left[ \frac{1}{2}(q')^2 + \frac{1}{p}q^p - \frac{1}{2}q^2 \right] = 0$$

Thus

$$\frac{1}{2}(q')^2 + \frac{1}{p}q^p - \frac{1}{2}q^2 = 0^3 \quad (7.28) \quad \boxed{\text{ODE2}}$$

Note that at its maximum say  $x_0$  we have  $q'(x_0) = 0$ , thus  $\frac{1}{p}q^p = \frac{1}{2}q^2$  at  $x_0$ . Solve for  $q$ , we get

$$q(x_0) = \left( \frac{p}{2} \right)^{\frac{1}{p-2}}$$

Moreover, equation (7.28) gives

$$q' = \pm \sqrt{q^2 - \frac{2}{p}q^p}$$

which is equivalent to

$$\frac{q'}{\sqrt{q^2 - \frac{2}{p}q^p}} = \pm 1$$

Integrate from  $x_0$  to  $x$ <sup>4</sup>, gives

$$\int_{x_0}^x \frac{q'}{\sqrt{q^2 - \frac{2}{p}q^p}} dx = -(x - x_0)$$

<sup>3</sup>The constant of integration has to be zero since  $q \in H^1$ , thus  $q \rightarrow 0$  as  $r \rightarrow \infty$ .

<sup>4</sup>Since  $x_0$  is the maximum,  $q$  is decreasing from  $x_0$  to  $x$ , so choose '-1'

Since  $q = q(x)$ ,  $dq = q' dx$ . Change variables in the integral, we get

$$\int_{q(x_0)}^{q(x)} \frac{dq}{\sqrt{q^2 - \frac{2}{p}q^p}} = -(x - x_0)$$

To evaluate this integral, we use the substitution

$$q = \left(\frac{p}{2}\right)^{\frac{1}{p-2}} \cosh(y)^{-\frac{2}{p-2}}$$

After some algebra, we get the solution

$$q(x) = \left(\frac{p}{2}\right)^{\frac{1}{p-2}} \cosh\left(\frac{p-2}{2}(x - x_0)\right)^{-\frac{2}{p-2}}$$

The requirement that  $q$  is radial forces  $x_0 = 0$ , and we get the unique positive radial  $H^1(\mathbb{R})$  solution to (7.27).

### *Case 2: $d \geq 2$*

We won't prove this directly. Existence follows from the existence of optimizers for Gagliardo-Nirenberg inequality (cf. below). Uniqueness is tricky. There is no known closed formula for the solution of (7.26).

We can compute it numerically to high accuracy using the shooting method. First, guess a value of  $Q$  at  $r = 0$ , call it  $Q(0)$ , then solve the ODE numerically. If the solution changes sign ("overshooting") at some radius  $r$ , this means our initial guess was too big. If the solution doesn't converge to zero ("undershooting"), our initial guess was too small. Once we find one solution that overshoots, and another solution that undershoots, we can do a bisection search repeatedly until we find the optimal guess  $Q(0)$  that gives the right solution within our expected numerical accuracy.

We can also solve the equation by power series expansion near the origin ( $r \ll 1$ ). For  $r$  large, the solution decays to zero, so  $Q^{p-1} \ll Q$ , and (7.26) becomes linear which we can solve explicitly. Then we can patch the two solutions together to get the full solution of the ODE.

The rigorous proof of uniqueness is precisely to make the shooting method rigorous, by proving the supremum of undershooting values is equal to the infimum of the overshooting values. By continuity arguments, the matching value of the two gives the correct guess  $Q(0)$ . As the ODE is uniquely determined by the value at the origin, the solution is unique.  $\square$

*Remark.*

- (1) We keep the ‘–’ sign in the PDE (7.26) to resemble the time independent Schrödinger equation

$$-\Delta\Psi + V\Psi = E\Psi$$

with wavefunction  $\Psi = Q$ , potential  $V = -Q^{p-2}$  and energy  $E = -1$ .

- (2) Elliptic regularity:  $Q$  is actually smooth. From the PDE, we see that  $\Delta Q$  is as smooth as  $Q$  and  $Q^{p-1}$ . Moreover, Calderon-Zygmund theory (247A) shows that  $\frac{\partial_j \partial_k}{\Delta}$  are bounded operators on  $L^p$ ,  $1 < p < \infty$ . Thus all second derivatives inherit the smoothness of  $\Delta$ .

(3) Pohozaev-type identities:

- (a) Multiply both sides of (7.26) by  $Q$  and integrate, we get the following Pohozaev-type identity:

$$\int |\nabla Q|^2 - \int Q^p = - \int Q^2 \quad (7.29) \quad \boxed{\text{poho1}}$$

- (b) Multiply both sides of (7.26) by  $x \cdot \nabla Q$  and integrate. Note that

$$\int Q^k \cdot x \cdot \nabla Q = \frac{1}{k+1} \int x \cdot \nabla (Q^{k+1}) = -d \int Q^{k+1} \quad (7.30)$$

and also<sup>5</sup>

$$\begin{aligned} \int x \cdot \nabla Q \cdot \Delta Q &= \int x_j Q_j Q_{kk} \\ &= \int x_j [(Q_j Q_k)_k - Q_{jk} Q_k] \\ &= \int x_j \left[ (Q_j Q_k)_k - \frac{1}{2} (Q_k^2)_j \right] \\ &= \int -\delta_{jk} Q_j Q_k dx + \int (x_j)_j \cdot \frac{1}{2} [Q_k]^2 dx \\ &= (-1 + \frac{d}{2}) \int |\nabla Q|^2 \end{aligned}$$

Combine the results above with  $k = p - 1$  and  $k = 1$ , we get another Pohozaev-type identity:

$$\frac{d-2}{2} \int |\nabla Q|^2 - \frac{d}{p} \int Q^p = -\frac{d}{2} \int Q^2 \quad (7.31) \quad \boxed{\text{poho2}}$$

### Lecture 25: Optimizers for Gagliardo-Nirenberg, continued

**Proposition 7.13.** Suppose  $d = 1$  and  $2 < p < \infty$ , or  $d = 2$  and  $2 < p < \frac{d}{d-2}$ . Let  $f$  be a nonzero optimizer for Gagliardo-Nirenberg. Then there exist  $\alpha \in \mathbb{C}, \lambda > 0, x_0 \in \mathbb{R}^d$  such that

$$f(x) = \alpha Q(\lambda(x - x_0)).$$

*Proof.* Under rearrangement we have  $\|f^*\|_p = \|f\|_p$ ,  $\|f^*\|_2 = \|f\|_2$ , and  $\|\nabla f^*\|_2 \leq \|\nabla f\|_2$ . If  $f$  is an optimizer, then so is  $f^*$ . Moreover,  $\|\nabla f^*\|_2 = \|\nabla f\|_2$ . Later we prove that  $\nabla f^*$  is nonzero a.e.

By the theorem of Brothers-Ziemer discussed earlier,<sup>6</sup> this guarantees that

$$f(x) = \alpha f^*(x - x_0)$$

for some  $x_0 \in \mathbb{R}^d, \alpha \in \mathbb{C}$  with  $|\alpha| = 1$ . Now we are reduced to studying radial nonnegative optimizers  $f \not\equiv 0$ .

As  $f \not\equiv 0$ , the quotient

$$J(f) = \frac{\|f\|_p}{\|f\|_2^\theta \|\nabla f\|_2^{1-\theta}}$$

---

<sup>5</sup>Here we use the Einstein notation, summation over repeated indices, where  $x_j$  is the  $j$ -th component of  $x$ , and  $Q_j$  denote the  $j$ -th component of  $\nabla Q$ .

<sup>6</sup>See also: J.E. Brothers, W.P. Ziemer, Minimal rearrangements of Sobolev functions, J. Reine Angew. Math. 384 (1988) 153–179.

is differentiable at  $f$ , in the sense that

$$J(f + \phi) - J(f) = J'(f)\phi + o(\|\phi\|)$$

as  $\|\phi\| \rightarrow 0$ .

In particular, given a Schwartz function  $\phi$ , we have that

$$\frac{d}{d\epsilon} \Big|_{\epsilon=0} J(f + \epsilon\phi)$$

exists and is equal to 0.

We have

$$\begin{aligned} \frac{d}{d\epsilon} \Big|_{\epsilon=0} \int |f + \epsilon\phi|^p dx &= \frac{d}{d\epsilon} \Big|_{\epsilon=0} \int (|f + \epsilon\phi|^2)^{p/2} dx \\ &= \int p|f|^{p-2} \operatorname{Re}(\bar{\phi}f) dx, \end{aligned}$$

where we have used that  $|f + \epsilon\phi|^2 = |f|^2 + \epsilon^2|\phi|^2 + 2\operatorname{Re}\epsilon\bar{\phi}f$ .

It follows that

$$\begin{aligned} \frac{d}{d\epsilon} \Big|_{\epsilon=0} \|f + \epsilon\phi\|_p &= \frac{d}{d\epsilon} \Big|_{\epsilon=0} (\|f + \epsilon\phi\|_p^p)^{1/p} \\ &= \frac{1}{p} \|f\|_p^{p(\frac{1}{p}-1)} \frac{d}{d\epsilon} \Big|_{\epsilon=0} \|f + \epsilon\phi\|_p^p \\ &= \frac{1}{p} \|f\|_p^{1-p} \int p|f|^{p-2} \operatorname{Re}(\bar{\phi}f) dx. \end{aligned}$$

Similarly, we have

$$\frac{d}{d\epsilon} \Big|_{\epsilon=0} \int |\nabla(f + \epsilon\phi)|^2 dx = -2\operatorname{Re} \int \Delta\bar{\phi}f dx.$$

Proceeding in this way, we find

$$A \int |f|^{p-2} f \bar{\phi} - B \int f(-\Delta\phi) - C \int f \bar{\phi} = 0$$

for every  $\phi$  and certain nonzero  $A, B, C$ . We dropped “Re” since the same identity holds for  $\phi$  replaced by  $i\phi$ . This says that  $f$  is an  $H^1$  distributional solution to

$$-B\Delta f - A|f|^p f = -Cf.$$

By choosing  $\lambda, \alpha > 0$  intelligently, we find that  $g(x) = (1/\alpha)f(x/\lambda)$  solves

$$-\Delta g - |g|^{p-2} g = -g.$$

(Indeed,  $\lambda$  only changes the constant in front of the Laplacian, and  $\alpha$  can be used to adjust the ratio between the other two coefficients.) Like  $f$ ,  $g$  is nonnegative and radial. Therefore by the preceding proposition,  $g = Q$ . Thus the only nonnegative optimizers are

$$f(x) = \alpha Q(\lambda x).$$

In order to apply the Brothers-Ziemer theorem to reduce to positive radial solutions, we need to check that  $\nabla Q \neq 0$  a.e. If this is not so, write  $Q(x) = q(|x|)$ , where

$$-q'' - \frac{d-1}{r}q' - |q|^{p-2}q = -q.$$

If  $r \neq 0$  is a Lebesgue point of  $\{x : q'(x) \neq 0\}$ , then since  $r$  is an accumulation point of the places where  $q' \neq 0$ , we have  $q''(r) = 0$ . This then shows that  $q(r) = 0$  or

$|q(r)| = 1$ . But by uniqueness for ODEs, these possibilities would imply  $q(\rho) \equiv q(r)$ , which contradicts that  $Q \in H^1 \setminus \{0\}$ .  $\square$

## 8. CONCENTRATION COMPACTNESS

Next, we outline an alternative proof of the existence of Gagliardo-Nirenberg optimizers when  $d = 1$  and  $2 < p < \infty$ . Our goal is to introduce some ideas of concentration compactness.

Let  $(f_n)$  be an optimizing sequence with  $\|f_n\|_2 = \|f'_n\|_2 = 1$ . First we observe that  $\liminf_{n \rightarrow \infty} \|f_n\|_\infty > 0$ . If not, then

$$\|f_{n_j}\|_p \leq \|f_{n_j}\|_2^{2/p} \|f_{n_j}\|_\infty^{1-2/p} \rightarrow 0$$

along a subsequence, so  $J(f_{n_j}) \rightarrow 0$ , a contradiction. As  $\liminf_{n \rightarrow \infty} \|f_n\|_\infty > \epsilon$ , there exists  $(x_n) \in \mathbb{R}$  such that  $\lim_{n \rightarrow \infty} |f_n(x_n)| > \epsilon$  along a subsequence. We replace our original sequence by  $(f_n(x + x_n))$ , so that  $\lim_{n \rightarrow \infty} |f_n(0)| > \epsilon$ .

By Rellich-Kondrachov, we may also ensure that  $f_n \rightarrow \phi$  in  $L^p(K)$  for any compact  $K$ . Passing to another subsequence, we may assume that  $f_n \rightarrow \phi$  a.e. As  $|f_n(0)| > \epsilon$  for  $n$  large and  $\|f_n\|_{C^{1/2}} \lesssim \|f_n\|_{H^1} \lesssim 1$ , for  $|x| \leq C\epsilon^2$  (with some  $C$ ), we have

$$|f(x)| \geq |f(0)| - \|f\|_{C^{1/2}} |x|^{1/2} \geq \frac{\epsilon}{2}.$$

Thus

$$\int |\phi|^p dx \geq \int_{|x| \leq C\epsilon^2} \left| \frac{\epsilon}{2} \right|^p dx > 0,$$

hence  $\phi \not\equiv 0$ . We will proceed with this argument in the next lecture.

*Alternate Proof of existence of optimizers for Gagliardo-Nirenberg.* This is a brief recap of last lecture. Given an optimizing sequence for Gagliardo-Nirenberg  $f_n$  normalized so that  $\|f_n\|_{L^2} = \|Df_n\|_{L^2} = \frac{1}{\sqrt{2}}$  (so that  $\|f_n\|_{H^1} = 1$ , we saw that there exists  $\varphi \neq 0 \in H^1$  and a sequence  $x_n \in \mathbb{R}$  so that  $f_n(x - x_n)$  converges to  $\varphi$  weakly in  $H^1$  for almost all  $x \in \mathbb{R}$  and in  $L^p(K)$  for almost all  $K$  compact. Thus, we may replace  $f_n(x)$  with  $f_n(x + x_n)$  so that  $f_n$  converges to  $\varphi$ . Let  $r_n = f_n - \varphi$ , so  $f_n = \varphi + r_n$ .

**Claim:**  $\|f_n\|_{L^p}^p = \|\varphi\|_{L^p}^p + \|r_n\|_{L^p}^p + o(1)$  as  $n \rightarrow \infty$  and  $\|f_n\|_{H^1}^2 = \|\varphi\|_{H^1}^2 + \|r_n\|_{H^1}^2 + o(1)$  as  $n \rightarrow \infty$ .

Notice that equality holds when  $\varphi$  and  $r_n$  have disjoint support. The second claim turns out to be a general property of Hilbert spaces: as  $f_n$  converges weakly to  $\varphi$ , it follows that

$$\|\varphi + r_n\|_{H^1}^2 = \|\varphi\|_{H^1}^2 + \|r_n\|_{H^1}^2 + 2\operatorname{Re} \langle \varphi, r_n \rangle_{H^1}$$

and the right-most term on the right hand side goes to zero since  $r_n$  converges weakly to 0. For the first claim, we need the Brezis-Leib lemma:

**Lemma 8.1.** *Let  $f_n \in L^p(\mathbb{R}^d)$  with  $1 \leq p < \infty$  with  $f_n \rightarrow f$  almost everywhere and*

$$\limsup_{n \rightarrow \infty} \|f_n\|_{L^p} < \infty.$$

*Then*

$$\int ||f_n|^p - |f_n - f|^p - |f|^p| dx \rightarrow 0.$$

*Remark.* This shows that  $\|f\|_{L^p}^p + \|f_n - f\|_{L^p}^p - \|f_n\|_{L^p}^p \rightarrow 0$ , which would show the first part of the claim. By comparison, Fatou's lemma shows that

$$\liminf \int |f_n|^p dx \geq \int \liminf_{n \rightarrow \infty} |f_n|^p dx$$

or equivalently,

$$\liminf_{n \rightarrow \infty} \|f_n\|_{L^p} \geq \|f\|_{L^p}.$$

Thus, in a sense, the Brezis-Leib lemma encompasses the “missing term” in Fatou’s lemma.

*Proof.* We claim that for any  $\epsilon > 0$ ,

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

If  $|a| \geq B^{-1}|b|$ , then  $|a + b|^p + |b|^p \lesssim_B |a|^p$  which is stronger than the inequality we are looking to prove. So we are looking to prove the inequality for  $|a| \ll |b|$ . By dividing by  $|b|$ , we only need to show that

$$||1 + z|^p - 1| \leq \epsilon + C_\epsilon |z|^p.$$

Since  $z \mapsto |1 + z|^p$  is smooth for sufficiently small  $p$ , it follows that  $||1 + z|^p - 1| \lesssim |z| \lesssim \epsilon + C_\epsilon |z|^p$ . This inequality shows that

$$||f_n - f|^p - |f_n|^p + |f|^p| \leq \epsilon |f_n - f|^p + (C_\epsilon + 1) |f|^p.$$

Thus

$$h_n := \max\{0, |f_n - f|^p - |f_n|^p + |f|^p - \epsilon |f_n - f|^p\}$$

satisfies  $|h_n(x)| \leq (C_\epsilon + 1) |f|^p \in L^1$ . In addition, since  $|f_n - f|^p - |f_n|^p + |f|^p \rightarrow 0$  almost everywhere and  $-\epsilon |f_n|^p \leq 0$ , it follows that  $h_n(x) \rightarrow 0$  almost everywhere. Thus,

$$\limsup_{n \rightarrow \infty} \int ||f_n|^p - |f - f_n|^p - |f|^p| dx \leq \limsup_{n \rightarrow \infty} \int h_n + \epsilon |f - f_n|^p dx \leq \epsilon \limsup_{n \rightarrow \infty} 2^p \|f_n\|_{L^p}^p$$

by applying the dominated convergence theorem to  $h_n$  and by Fatou’s lemma. This completes the proof of the lemma.  $\square$

Let  $C$  denote the optimal constant for the Gagliardo-Nirenberg inequality. Then

$$\begin{aligned} C &= \lim_{n \rightarrow \infty} \|f_n\|_{L^p} \\ &= \liminf_{n \rightarrow \infty} (\|\varphi_n\|_{L^p}^p + \|r_n\|_{L^p}^p)^{1/p} \\ &\leq \liminf_{n \rightarrow \infty} C(\|\varphi_n\|_{L^2}^{\theta p} \|D\varphi\|_{L^2}^{(1-\theta)p} + \|r_n\|_{L^2}^{\theta p} \|Dr_n\|_{L^2}^{\theta p})^{1/p} \end{aligned}$$

By the Young’s inequality, we have

$$\begin{aligned} &(\|\varphi\|_{L^2}^\theta \|D\varphi\|_{L^2}^{1-\theta})^2 + (\|r_n\|_{L^2} \|Dr_n\|_{L^2})^2 \leq \\ &\theta \|\varphi\|_{L^2}^2 + (1-\theta) \|D\varphi\|_{L^2}^2 + \theta \|r_n\|_{L^2}^2 + (1-\theta) \|Dr_n\|_{L^2}^2 \leq 1 + o(1) \end{aligned} \tag{8.1} \quad \boxed{\text{phibound}}$$

by our claims. Thus, we have the inequalities

$$\|\varphi_n\|_{L^2}^{\theta p} \|D\varphi\|_{L^2}^{(1-\theta)p} + \liminf_{n \rightarrow \infty} \|r_n\|_{L^2}^{\theta p} \|Dr_n\|_{L^2}^{\theta p} \geq 1$$

and

$$(\|\varphi\|_{L^2}^\theta \|D\varphi\|_{L^2}^{1-\theta})^2 + \left( \liminf_{n \rightarrow \infty} \|r_n\|_{L^2} \|Dr_n\|_{L^2} \right)^2 \leq 1.$$

But  $p > 2$  and the unit  $\ell^2$ -ball is inscribed in the unit  $\ell^p$  ball of two-dimensional Euclidean space, so the point  $(\|\varphi\|_{L^2}^\theta \|D\varphi\|_{L^2}^{1-\theta}, \liminf_{n \rightarrow \infty} \|r_n\|_{L^2} \|Dr_n\|_{L^2})$  lies on the point that the two balls touch, meaning that one of these must be zero. As  $\varphi$  is nonzero, this means that passing to a subsequence,  $\lim_{n \rightarrow \infty} \|r_n\|_{L^2} = 0$  or  $\lim_{n \rightarrow \infty} \|Dr_n\|_{L^2} = 0$ . Thus,  $r_n \rightarrow 0$  in  $H^1$  since if not, then 8.1 implies that  $\|\varphi\|_{L^2} \|D\varphi\|_{L^2} < 1$ , a contradiction to our assumption that  $C$  is the optimizing constant.  $\square$

We now return to our alternative proof of the existence of optimizers for the Gagliardo–Nirenberg inequality.

*Proof.* By taking repeated subsequences, we showed there is a sequence of optimizers  $f_n$  with

$$\|f_n\|_{L^2} = \|f'_n\|_{L^2} = 1.$$

We also know there is a  $\varphi \in H^1$  not identically zero so that

$$r_n := f_n - \varphi$$

converges to 0 both a.e. and weakly in  $H^1$ .

We showed also that this implies

$$\begin{aligned} 1 &= \|f_n\|_{L^2}^2 = \|\varphi\|_{L^2}^2 + \|r_n\|_{L^2}^2 + o(1) \\ 1 &= \|f'_n\|_{L^2}^2 = \|\varphi'\|_{L^2}^2 + \|r'_n\|_{L^2}^2 + o(1) \\ \|f_n\|_{L^p}^p &= \|\varphi\|_{L^p}^p + \|r_n\|_{L^p}^p + o(1). \end{aligned} \tag{8.2} \quad \boxed{27.\text{norms}}$$

Let  $C$  be the optimal constant for the Gagliardo–Nirenberg inequality. We compute

$$\begin{aligned} C &= \lim_{n \rightarrow \infty} \|f_n\|_{L^p} \\ &= \liminf_{n \rightarrow \infty} (\|\varphi\|_{L^p}^p + \|r_n\|_{L^p}^p)^{1/p} \\ &\leq \liminf_{n \rightarrow \infty} C \left( \left[ \|\varphi\|_{L^2}^\theta \|\varphi'\|_{L^2}^{1-\theta} \right]^p + \left[ \|r_n\|_{L^2}^\theta \|r'_n\|_{L^2}^{1-\theta} \right]^p \right). \end{aligned}$$

The last line comes from the Gagliardo–Nirenberg inequality applied to  $\varphi$  and  $r_n$ . By Young's inequality (or the weighted AM-GM inequality),

$$\begin{aligned} &\left[ \|\varphi\|_{L^2}^\theta \|\varphi'\|_{L^2}^{1-\theta} \right]^2 + \left[ \|r_n\|_{L^2}^\theta \|r'_n\|_{L^2}^{1-\theta} \right]^2 \\ &\leq \theta \|\varphi\|_{L^2}^2 + (1-\theta) \|\varphi'\|_{L^2}^2 + \theta \|r_n\|_{L^2}^2 + (1-\theta) \|r'_n\|_{L^2}^2 \\ &\leq 1 + o(1). \end{aligned}$$

Hence, the following two inequalities are true:

$$\begin{aligned} 1 &\leq \left[ \|\varphi\|_{L^2}^\theta \|\varphi'\|_{L^2}^{1-\theta} \right]^p + \left[ \liminf_{n \rightarrow \infty} \|r_n\|_{L^2}^\theta \|r'_n\|_{L^2}^{1-\theta} \right]^p \\ &\left[ \|\varphi\|_{L^2}^\theta \|\varphi'\|_{L^2}^{1-\theta} \right]^2 + \left[ \liminf_{n \rightarrow \infty} \|r_n\|_{L^2}^\theta \|r'_n\|_{L^2}^{1-\theta} \right]^2 \leq 1. \end{aligned}$$

This means that the point

$$\left( \|\varphi\|_{L^2}^\theta \|\varphi'\|_{L^2}^{1-\theta}, \liminf_{n \rightarrow \infty} \|r_n\|_{L^2}^\theta \|r'_n\|_{L^2}^{1-\theta} \right)$$

lies *inside* the closed  $\ell^2$  unit ball in  $\mathbb{R}^2$  but *outside* the open  $\ell^p$  unit ball in  $\mathbb{R}^2$  for a given  $p > 2$ . But the closed  $\ell^2$  ball lies entirely within the open  $\ell^p$  ball except at four points:

$$(\pm 1, 0), (0, \pm 1).$$

Since we know that  $\varphi$  is not identically 0, the first coordinate cannot be 0. This means that

$$\liminf_{n \rightarrow \infty} \|r_n\|_{L^2}^\theta \|r'_n\|_{L^2}^{1-\theta} = 0.$$

We can consider yet another subsequence to make this liminf an actual limit.

By the Gagliardo–Nirenberg inequality again, we know that

$$\|r_n\|_{L^p} \leq C \|r_n\|_{L^2}^\theta \|r'_n\|_{L^2}^{1-\theta},$$

so this goes to 0. Hence  $\|\varphi\|_{L^p} = C$ . By (8.2), we must also have that  $\|r_n\|_{H^1}$  goes to 0, otherwise one of  $\|\varphi\|_{L^2}$  or  $\|\varphi'\|_{L^2}$  would be strictly less than 1. This contradicts  $C$  as a bound.  $\square$

In our proof, we defeated two enemies:

- vanishing (i.e., for every choice of  $x_n$ ,  $f_n(x + x_n) \rightarrow 0$ , forcing  $\|f_n\|_{L^\infty}$  and the ratio to go to 0)
- dichotomy (i.e.,  $f_n = \varphi + r_n$  with  $r_n$  not going to 0)

to obtain the desired third outcome:

- compactness (i.e., the optimizing sequence converges in norm to  $\varphi$ , modulo a translation).

We will now prove that the Sobolev embedding inequality

$$\|f\|_{L^{\frac{2d}{d-2}}} \leq C \|\nabla f\|_{L^2} \quad d \geq 3$$

admits an optimizer. This has a particularly worrying scaling symmetry. If  $W$  is an optimizer, then so is

$$\lambda^{\frac{d}{2}-1} W(\lambda x).$$

Therefore even sequences of *radial* optimizers need not converge.

We will use the following lemma:

**Lemma 8.2** (Refined Sobolev embedding).

$$\|f\|_{L^{\frac{2d}{d-2}}} \lesssim \|\nabla f\|_{L^2}^{\frac{d-2}{d}} \sup_N \|f_N\|_{L^{\frac{2d}{d-2}}}^{\frac{2}{d}}$$

where  $f_N$  denotes the Littlewood–Paley projection of  $f$  to scale  $N \in 2^{\mathbb{Z}}$ .

*Proof.* Let's do  $d \geq 4$  first.

$$\begin{aligned} \|f\|_{L^{\frac{2d}{d-2}}}^{\frac{2d}{d-2}} &\lesssim \left\| \sqrt{\sum_N |f_N|^2} \right\|_{L^{\frac{2d}{d-2}}}^{\frac{2d}{d-2}} \\ &\lesssim \int \left[ \sum_N |f_N| \right]^{\frac{d}{2(d-2)}} \cdot \left[ \sum_M |f_M| \right]^{\frac{d}{2(d-2)}} dx \\ &\lesssim \int \sum_{N \leq M} |f_N|^{\frac{d}{d-2}} \cdot |f_M|^{\frac{d}{d-2}} dx. \end{aligned}$$

The last line follows due to the subadditivity of fractional powers. When  $d \geq 4$ , the exponent  $\frac{d}{2(d-2)}$  is no more than 1.  $\square$

**Lemma 8.3** (Refined Sobolev embedding).

$$\|f\|_{L^{\frac{2d}{d-2}}} \lesssim \|\nabla f\|_{L^2}^{\frac{d-2}{d}} \sup_N \|f_N\|_{L^{\frac{2d}{d-2}}}^{\frac{2}{d}},$$

where  $f_N$  denotes the Littlewood-Paley piece at scale  $N \in 2^{\mathbb{Z}}$ .

*Proof.* Lets do  $d \geq 4$ .

$$\begin{aligned} \|f\|_{L^{\frac{2d}{d-2}}}^{\frac{2d}{d-2}} &\lesssim \left\| \sqrt{\sum |f_N|^2} \right\|_{L^{\frac{2d}{d-2}}}^{\frac{2d}{d-2}} \\ &\lesssim \int \left[ \sum_N |f_N|^2 \right]^{\frac{2d}{2(d-2)}} \left[ \sum_M |f_M|^2 \right]^{\frac{2d}{2(d-2)}} dx \end{aligned}$$

By the subadditivity of fractional powers and the fact that  $d \geq 4$  implies  $\frac{d}{2(d-2)} \leq 1$ ,

$$\begin{aligned} &\lesssim \int \sum_{N \leq M} |f_N|^{\frac{d}{d-2}} |f_M|^{\frac{d}{d-2}} dx \\ &\quad \left( \frac{1}{2} \frac{d}{d-2} + \frac{1}{2} \frac{d-4}{d-2} + \frac{4}{d-2} \frac{1}{\infty} = 1 \right) \\ &\lesssim \sum_{N \leq M} \|f_M\|_{L^2}^{\frac{d}{d-2}} \|f_N\|_{L^2}^{\frac{d-4}{d-2}} \|f_N\|_{L^\infty}^{\frac{4}{d-2}} \end{aligned}$$

and so, using Bernstein,

$$\begin{aligned} &\lesssim \sum_{N \leq M} \|\nabla f_M\|_{L^2}^{\frac{d}{d-2}} M^{-\frac{d}{d-2}} \|\nabla f_N\|_{L^2}^{\frac{d-4}{d-2}} N^{-\frac{d-4}{d-2}} \|\nabla f_N\|_{L^{\frac{2d}{d-2}}}^{\frac{4}{d-2}} N^2 \\ &\lesssim \sup_k \|f_k\|_{L^{\frac{2d}{d-2}}}^{\frac{4}{d-2}} \sum_{N \leq M} \|\nabla f_M\|_{L^2}^{\frac{d}{d-2}} \|\nabla f_N\|_{L^2}^{\frac{d-4}{d-2}} \left( \frac{N}{M} \right)^{\frac{d}{d-2}} \end{aligned}$$

Consider the last part  $\sum_{N \leq M} \|\nabla f_M\|_{L^2}^{\frac{d}{d-2}} \|\nabla f_N\|_{L^2}^{\frac{d-4}{d-2}} \left( \frac{N}{M} \right)^{\frac{d}{d-2}}$  as vector · matrix vector. Then the part corresponding to  $\|\nabla f_M\|_{L^2}^{\frac{d}{d-2}}$  belongs to  $\ell^{\frac{2(d-2)}{d}}$  and the part corresponding to  $\|\nabla f_N\|_{L^2}^{\frac{d-4}{d-2}}$  belongs to  $\ell^{\frac{2(d-2)}{d-4}}$ . Note that these two values are dual that  $\frac{d}{2(d-2)} + \frac{d-4}{2(d-2)} = \frac{2d-4}{2(d-2)} = 1$ . By Schur's test the matrix with entries

$$\left( \frac{N}{M} \right)^{\frac{d}{d-2}} \cdot 1_{\{N \leq M\}}$$

is bounded from  $\ell^p \rightarrow \ell^{p'}$  for every  $1 \leq p \leq \infty$ . This shows

$$\begin{aligned} \|f\|_{L^{\frac{2d}{d-2}}}^{\frac{2d}{d-2}} &\lesssim \sup_k \|f_k\|_{L^{\frac{2d}{d-2}}}^{\frac{4}{d-2}} \left( \sum_N \|\nabla f_N\|_{L^2}^2 \right) \lesssim \sup_k \|f_k\|_{L^{\frac{2d}{d-2}}}^{\frac{4}{d-2}} \|\nabla f\|_{L^2}^2 \\ \implies \|f\|_{L^{\frac{2d}{d-2}}} &\lesssim \|\nabla f\|_{L^2}^{\frac{d-2}{d}} \sup_N \|f_N\|_{L^{\frac{2d}{d-2}}}^{\frac{2}{d}} \end{aligned}$$

*Remark.* For  $A : \ell^1 \rightarrow \ell^1$ ,  $\langle y, Ax \rangle \leq \|y\|_{\ell^\infty} \|x\|_{\ell^1}$ .  $\|A\|_{\ell^1 \rightarrow \ell^1} = \sup_j \sum_i |A_{ij}|$  and  $\|A\|_{\ell^\infty \rightarrow \ell^\infty} = \sup_i \sum_j |A_{ij}|$

This settles  $d \geq 4$ . When  $d = 3$ , we start with

$$\|f\|_{L^6} \lesssim \sum_{N_1 \leq N_2 \leq N_3} \int |f_{N_1}|^2 |f_{N_2}|^2 |f_{N_3}|^2 dx$$

put  $f_{N_1}$  and  $f_{N_2}$  in  $L^\infty$  and  $f_{N_3}$  in  $\dot{H}^1$ , then it all works out.  $\square$

**Theorem 8.4.** Fix  $d \geq 3$ . The optimal constant  $C$  Sobolev embedding,

$$\|f\|_{L^{\frac{2d}{d-2}}} \leq C \|\nabla f\|_{L^2},$$

is achieved.

*Proof.* Let us focus on  $d = 3$ . Suppose  $f_n$  is an optimizing sequence:

$$\frac{\|f_n\|_{L^6}}{\|\nabla f_n\|_{L^2}} \rightarrow \sup_{0 \neq f \in \dot{H}^1} \frac{\|f_n\|_{L^6}}{\|\nabla f_n\|_{L^2}} = C > 0$$

Without loss of generality, we may assume  $\|\nabla f_n\|_{L^2} = 1$ . (If not replace  $f_n$  with  $\alpha_n f_n$ ). So  $\|f_n\|_{L^6} \rightarrow C > 0$ . Thus by refined Sobolev embedding  $\exists N_n \in 2^{\mathbb{Z}}$  so that

$$\lim_{n \rightarrow \infty} \|P_{N_n} f_n\|_{L^6} > 0$$

Let's use this do rescale:

$$f_n^{(\text{new})}(x) = N_n^{-\frac{1}{2}} f_n^{(\text{old})}(x/N_n)$$

Note  $\|\nabla f_n\|_{L^2} = 1$ ,  $\lim_{n \rightarrow \infty} \|f_n\|_{L^6} \rightarrow C$  and  $\|P_1 f_n\|_{L^6} \gtrsim 1$ . According to the fact that  $\|P_1 f_n\|_{L^2} \leq 1$  and from Hölder's inequality,

$$\|P_1 f_n\|_{L^\infty}^{\frac{2}{3}} \|P_1 f_n\|_{L^2}^{\frac{1}{3}} \geq \|P_1 f_n\|_{L^6} \gtrsim 1 \implies \|P_1 f_n\|_{L^\infty} \gtrsim 1$$

So  $\exists x_n$  with  $(P_1 f_n)(x_n) \gtrsim 1$ .

Translating our sequence we may assume  $x_n \not\equiv 0$ .

**Claim:** A subsequence of our new sequence  $f_n$  converges in  $\dot{H}^1$  to an optimizer. Passing to a s.s.  $f_n \rightharpoonup \varphi$  in  $\dot{H}^1$ . This implies

$$\|\nabla f_n\|_{L^2}^2 = \|\nabla \varphi\|_{L^2}^2 + \|\nabla(f_n - \varphi)\|_{L^2}^2 + o(1), \text{ as } n \rightarrow \infty.$$

Writing  $\psi \in \mathbb{S}$  (Schwartz function) for the convolution kernel of  $P_1$ .

$$|(\psi * f_n)(0)| = |\psi(0 - y) f_n(y) dy| \gtrsim 1 \implies |\langle \psi, f_n \rangle| \gtrsim 1.$$

But  $f_n \rightharpoonup \varphi$  in  $\dot{H}^1 \hookrightarrow L^6$  and  $\psi \in L^{\frac{6}{5}} \hookrightarrow (\dot{H}^1)^*$ . Then

$$|\langle \psi, \varphi \rangle| \gtrsim 1 \implies \|\varphi\|_{L^6} \gtrsim \|\psi\|_{L^{\frac{6}{5}}}^{-1}.$$

In particular  $\varphi \not\equiv 0$ .

Next: passing to a subseq we may ensure that  $f_n \rightarrow \varphi$  a.e. Here's why:  $\chi_k f_n$  is  $L^2$ -precompact for every compact  $K \subset \mathbb{R}^3$ .

(Pf.  $\|\chi_k f_n\|_{L^2} \lesssim_k \|f_n\|_{L^6}$  so bounded. Tightness is trivial. Equicontinuity follows from  $\|\nabla f_n\|_{L^2} < \infty$ )

Thus we easily get subsequencial  $L^2$  convergence on compact sets and they are almost everywhere convergence.

The Brezis-Lieb lemma then guarantee that (along this subsequence)

$$\|f_n\|_{L^6}^6 = \|\varphi\|_{L^6}^6 + \|r_n\|_{L^6}^6 + o(1),$$

where  $r_n = f_n - \varphi$ .

$$\begin{aligned} C^6 &\leftarrow \|f_n\|_{L^6}^6 = \|\varphi\|_{L^6}^6 + \|r_n\|_{L^6}^6 + o(1) \\ &\leq C^6 [\|\nabla\varphi\|_{L^2}^6 + \|\nabla r_n\|_{L^6}^6] + o(1). \end{aligned}$$

Thus by strict [except at  $(0, \pm 1)$  and  $(\pm 1, 0)$ ] inclusion of the  $\ell^2 \subset \ell^6$ , we see that either  $\|\nabla\varphi\|_{L^2} = 0$  or  $\|\nabla r_n\|_{L^2} \rightarrow 0$ . The former is forbidden since  $\varphi \not\equiv 0$ . Thus  $\|\nabla r_n\|_{L^2} \rightarrow 0$ . Thus  $\|f_n - \varphi\|_{\dot{H}_1} \rightarrow 0$  and  $\|\nabla\varphi\|_{L^2} = 1$ . Thus  $f_n \rightarrow \varphi$  and  $\varphi$  is an optimizer.  $\square$

*Remark.* The concrete form of optimizers can be computed:

$$W(x) = \frac{[d(d-2)]^{\frac{d-2}{4}}}{(1+|x|^2)^{\frac{d-2}{2}}}.$$