

Uncertainty principle

XIYU HU

December 29, 2017

1 introduction

Problem 1 *Is there a Brunn-Minkowski inequality approach to the phenomenon charged by uncertainty principle? More precisely, is it possible to say some thing about the Gaussian distribution*

$$G(x) = e^{-|x|^2} \tag{1}$$

to be the best choice that $\|\hat{G} - G\|_2$ arrive minimum?

Remark 1 *Or some other suitable distance space on reasonable function (may be some gromov hausdorff distance? Any way, to say the gaussian distribution is the best function to defect the influence of uncertain principle.*

I do not know the answer of the problem 1, but this is a phenomenon of a universal phylosphy, aid, uncertainty principle, heuristic:

It is not possible for both function f and its Foriour transform \hat{f} to be localized on small set.

Now let me give some approach by intuition to explain why the phenomenon of "uncertainty principle" could happen.

The approach is based on:

1. level set decomposition.
2. area formula (or coarea formula), anyway, some kind of change variable formula.
3. integral by part.
4. Basic understanding on exponential sum.

Let our function $f \in S$ the Shwarz space, we begin with a intuition (not very rigorous) calculate:

$$\begin{aligned} \hat{f}(\xi) &= \int e^{2\pi i \langle \xi, x \rangle} f(x) dx \\ &\stackrel{\text{integral by part}}{=} \int \frac{1}{-2\pi i \xi} e^{-2\pi i \langle x, \xi \rangle} \cdot \nabla f(x) \\ &\stackrel{\text{Fubini}}{=} \int_{\inf|f|}^{\max|f|} \int_{\text{levelset}A(t)} \frac{-e^{2\pi i \langle \xi, x \rangle}}{-2\pi i \xi} \nabla f(x) dH^{n-1}(A) dt \end{aligned}$$

Now we try to understanding the result of the calculate, it is,

$$\hat{f}(\xi) = \int_{\inf|f|}^{\max|f|} \int_{\text{levelset}A(t)} \frac{-e^{2\pi i \langle \xi, x \rangle}}{-2\pi i \xi} \nabla f(x) dH^{n-1}(A) dt \quad (2)$$

$$\mathcal{L}(A(t), \xi) = \int_{\text{levelset}A(t)} \frac{-e^{2\pi i \langle \xi, x \rangle}}{-2\pi i \xi} \nabla f(x) dH^{n-1}(A) \quad (3)$$

The calculate is wrong, but not very far from the thing that is true, the key point is now the exponential sum involve. We could use the pole coordinate in the frequence space and get some very rough intuition of why the the uncertainty principle could occur.

Remark 2 *Why we consider the level set decomposition, due to the integral is a combination of linear sum of the integral on every level set, so shape of level set is the key point.*

The part of $\frac{-e^{2\pi i \langle \xi, x \rangle}}{-2\pi i \xi}$ in 2 is a rotation on the level set, a wave correlation of it and the christization function χ_{A_t} of level set A_t in the whole space, this is of course a exponential sum.

Now we can begin the final intuition explain of the phenomenon of uncertainty principle. If the density of function f is very focus on some small part of the physics space, then it is the case for level sets of f , but we could say some thing for the exponential sum $\mathcal{L}(A(t), \xi)$ 3 related to the level set, just by very simply argument with hardy litterwood circle method or Persaval identity? Any way, something similar to this argument will make sense, due to if the diameter of level set focus ois small, then we can not get a decay estimate for $\mathcal{L}(A(t), \xi)$ when $\xi \rightarrow \infty$ along one direction in frequency space, in fact we could say the inverse, i.e. it could not decay very fast.

2 Bernstein's bound and Heisenberg uncertainty principle

2.1 Motivation and Bernstein's bound

There is two different Bernstein's bound, we discuss the first with the motivation, and proof the second rigorously.

Form 1 A is a invertible affine map, then for a ball B , $A(B) = \epsilon$ is a ellipsoid.

$$\epsilon = \{x \in \mathbb{R}^d \mid \sum_{j=1}^d r_j^{-2} (x_j - y_j)^2 \leq 1\} \quad (4)$$

By a orthogonal transform we could make A to be a diagonal matrix, i.e. $A = \text{diag}(r_1, \dots, r_d)$. It is said, for $\forall f \in S$ or f is a smooth bump function, $f_A = f \circ A^{-1}$, so we have,

$$\hat{f}_A(\xi) = \int e^{2\pi i \langle x, \xi \rangle} \cdot f \circ A^{-1}(x) dx \quad (5)$$

We define dual of ϵ , $\epsilon^* := \{\xi \in \mathbb{R}^d \mid \sum_{j=1}^d \xi_j^2 r_j^2 \leq 1\}$.

Remark 3 Why there we use the metric $\xi_j^2 r_j^2 \leq 1$ but not the standard inner product $\langle \xi, x \rangle$? How to understand the choice?

Proposition 4 We have the following property:

1. $f_A \in L^\infty \implies \|\hat{f}_A\|_1 \leq +\infty$.
2. $|\hat{f}_A(\xi)| \leq c_N |\xi| (1 + |\xi|_{\epsilon^*}^2)^{-N}$

Remark 5

$$|\xi|_{\epsilon^*}^2 = \sum_{i=1}^d \xi_i^2 r_i^2$$

This is a norm of \mathbb{R}^d related to ϵ^* .

PROOF: Suffice to proof 2.

$$\begin{aligned} |\hat{f}_A(\xi)| &= \left| \int_{\mathbb{R}^d} e^{2\pi i \langle x, \xi \rangle} f_A(x) dx \right| \\ &\stackrel{\text{integral by part}}{\sim} \frac{1}{(1 + |\xi|)^N} \int |e^{2\pi i \langle x, \xi \rangle} \partial^N f_A(x) dx| \\ &= c_N |\xi| (1 + |\xi|_{\epsilon^*}^2)^{-N} \end{aligned}$$

□

More quantitative we have rigorous one:

Form 2 If $f \in L^2(\mathbb{R}^d)$, $\text{supp} f \in B(r, 0)$, then it is not possible for \hat{f} to be concentrate on a scale much less than R^{-1} .

Proposition 6 (Bernstein's bound) Suppose $f \in L^2(\mathbb{R}^d)$, $\text{supp} f \subset B_R(0)$. Then,

$$\|\partial^\alpha \hat{f}\|_2 \leq (2\pi r)^{|\alpha|} \|f\|_2, \forall \alpha. \quad (6)$$

PROOF: $\alpha = 0$ case is trivial by Paserval identity, which said on $L^2(\mathbb{R}^d)$, fourier transform is a isometry, $\|f\|_2 = \|\hat{f}\|_2$.

For general case, integral by part, and use trivial estimate,

$$\begin{aligned} \|\partial^\alpha \hat{f}\|_2 &\stackrel{\text{integral by part}}{=} \|x^\alpha f\|_2 \\ &\leq (2\pi r)^\alpha \|f\|_2 \end{aligned}$$

□

2.2 Heisenberg inequality

Theorem 7 (Heisenberg uncertain principle) $f \in L^2(\mathbb{R}^d)$, so $\hat{f} \in L^2(\mathbb{R}^d)$, $\|f\|_2 = \|\hat{f}\|_2$. then for any $x_0, \xi_0 \in \mathbb{R}^d$, every direction, we have

$$\|f\|_2^2 = \|f\|_2 \|\hat{f}\|_2 \leq \|(x - x_0)f\|_2 \|(\xi - \xi_0)\hat{f}\|_2 \quad (7)$$

Remark 8 We could understand the inequality by the following way. suffice to prove it with $f \in S$ and then by approximation argument. $f \otimes \hat{f} \in S(\mathbb{R}^d \times \mathbb{R}^d)$, define $\|f \otimes \hat{f}\|_2 := \|f\|_{L^2(\mathbb{R}^d)} \cdot \|\hat{f}\|_{L^2(\mathbb{R}^d)}$. then we have the following:

$$\|f \otimes \hat{f}\|_{L^2} \leq 4\pi \|xf \otimes \hat{x}\hat{f}\|_{L^2} \quad (8)$$

Remark 9 The inequality is shape, the extremizers being precisely given by the modulated Gaussians: arbitrary

$$f(x) = ce^{2\pi i \xi_0 x} e^{-\pi \delta (x - x_0)^2} \quad (9)$$

There are two proof strategies I have tried, I try them for several hour but not work out with a satisfied answer, the method more involve, I explain what happen in section 1, I have not tried, I will try it later. Both this two strategies i face some difficulties, I explain why I can not work out them with a proof:

Strategy 1 The first one is, we could work with $f \in S$ of course, by approximation, then we find, by Paserval, $\|f\|_2 = \|\hat{f}\|_2, \|\partial_x f\|_2 = \|\xi \hat{f}\|_2$ and are both true. then we use our favourite way to use Cauchy-Schwarz, the difficulty is we can not use a integral by part argument directly, even after restrict ourselves with monotonic radical symmetry inequality and by a rearrangement inequality argument, it seems reasonable due to rearrangement decreasing the kinetic energy as said in Lieb's book. But even work with monotonic one, then one involve with some complicated form, try to use Fubini theorem to rechange the order of integral try to say something, it is possible to work out by this way but I do not know how to do. There is some calculate under this way,

$$\begin{aligned} \|xf\|_2 \|\xi \hat{f}\|_2 &\stackrel{\text{Cauchy-Schwarz}}{\geq} \int xf \cdot \partial_x f \\ &\sim \int f^2 \end{aligned}$$

but you know, at a point we have $\partial_x(Xf) = f + x\partial_x f \neq f$, the reasonable calculate is following,

$$\partial_x(xf) = f + x\partial_x f \quad (10)$$

We want $\partial_x P(x, f) = f$, Then

$$\begin{aligned} \partial_x P(x, f) &= f \\ &= \partial_x(xf) - x\partial_x f \\ &= \partial_x(xf) - \partial_x\left(\frac{1}{2}x^2\partial_x f\right) + \frac{1}{2}x^2\partial_{x^2} f \\ &= \partial_x\left(\frac{1}{6}x^3\partial_{x^2} f\right) - \frac{1}{6}x^3\partial_{x^3} f \\ &\dots \\ &= \partial_x\left(\sum_{i=1}^{\infty} (-1)^{i+1} x^i \frac{1}{i!} \partial_{x^i} f\right) + (-1)^{i+1} x^i \frac{1}{i!} \partial_{x^{i+1}} f \end{aligned}$$

Seems to be $f = \partial_x(\ln(f))\dots$ I do not know.

Strategy 2 The second strategy is, in the quantity $\|xf\|_2 \|\xi \hat{f}\|_x$ we lose two cone very near x_0, ξ , we need use the extra thing to make up them. May be effective argument come from some geometric inequality.

3 The Amerein-Berthier theorem

Next we investigate following problem, the problem is following:

Problem 2 if $E, F \subset \mathbb{R}^d$ are of finite measure, can there be a nonzero $f \in L^2(\mathbb{R}^d)$ with $\text{supp}(f) \subset E$ and $\text{supp}(\hat{f}) \subset F$?

Some argument is following: Observe that:

$$\chi_F \hat{f} = \hat{f} \implies \chi_E(\chi_F \hat{f})^\vee = f. \quad (11)$$

Assume that: $Tf := \chi_E(\chi_F \hat{f})^\vee$ then $Tf = f$. So we have, at least $\|T\|_{2 \rightarrow 2} \geq 1$. Some dirty calculate show:

$$\begin{aligned} (Tf)(x) &= \int e^{2\pi i \xi x} \chi_F \hat{f}(\xi) \chi_E(x) d\xi \\ &= \int \int e^{2\pi i \xi(x-y)} f(y) \chi_F(\xi) \chi_E(x) dy d\xi \\ &\stackrel{\text{Fubini}}{=} \int_{\mathbb{R}^d} \chi_E(x) \chi_F^\vee(x-y) f(y) dy \end{aligned}$$

So we can define kernel of T ,

$$K(x, y) = \chi_E(x) \chi_F^\vee(x-y) \quad (12)$$

By Fubini, we calculate the Hilbert-Schmidt norm:

$$\int_{\mathbb{R}^{2d}} |K(x, y)|^2 dx dy = |E||F| = \sigma^2 < +\infty \quad (13)$$

So T is a compact operator and its L^2 operator norm satisfied $\|T\| = \min(\sigma, 1)$. So if $\sigma < 1$ then we can calculate we can not have $f \neq 0$ in the original question.

The story is in fact more interesting, the answer of the question is no even for $\sigma \geq 1$, so in all case. We have the following quantitative theorem:

Theorem 10 E, F finite measure in \mathbb{R}^d , then

$$\|f\|_{L^2(\mathbb{R}^d)} \leq c(\|f\|_{L^2(E^c)} + \|\hat{f}\|_{L^2(F^c)}) \quad (14)$$

for some constant $c = c(E, F, d)$.

Remark 11 There is a naive approach for this theorem: Area formula trick, the shape of level set. Obviously we have:

$$\|f\|_{L^2(\mathbb{R}^d)} \leq \|f\|_{L^2(E)} + \|f\|_{L^2(E^c)} \quad (15)$$

Key point is proof:

$$\|f\|_{L^2(E)} \leq c(E, F, d) \|\hat{f}\|_{L^2(F^c)} \quad (16)$$

Let us do some useless further calculate:

$$\begin{aligned}
\|f\|_{L^2(E)} &= \|\chi_E \cdot f\|_{L^2(\mathbb{R}^d)} \\
&= \|\widehat{\chi_E \cdot f}\|_{L^2(\mathbb{R}^d)} \\
&= \|\hat{\chi}_E \cdot \hat{f}^\vee\|_{L^2(\mathbb{R}^d)} \\
&= \|\chi_E^\vee * f^\vee\|_{L^2(\mathbb{R}^d)}
\end{aligned}$$

So suffice to have:

$$\|\chi_E^\vee * f^\vee\|_{L^2(\mathbb{R}^d)} \leq c(E, F, d) \|\hat{f}\|_{L^2(F^c)} \quad (17)$$

But there is counter example given by modified scaling Gaussian distribution... The point is form 15 to 16 is too loose.

Following I given a right approach, following by my sprite on level set and area formula argument and discretization.

PROOF: The story is the same for a discretization one. We need point out, change the space \mathbb{R}^d to \mathbb{Z}^d , then every thing become a discretization one, and the change could been argue as a approximation way. What happen then, we have a naive picture in mind which is:

$$\delta \rightarrow \text{wave}, \text{ wave} \rightarrow \delta$$

What is the case with L^2 norm, it become the standard inner product on \mathbb{Z}^d , and the scale involve, i.e. we have the following basic estimate:

$$\|\chi_E f\|_1^2 \leq \|\chi_E f\|_2 \cdot |E| \quad (18)$$

Now image if the density of f concentrate in a very small area, then by a cut off argument we consider the supp of f , $\text{supp} f = E$ is very small, then use the argument 18, we could conclute the density of \hat{f} could not very concentrate in the fraquence space. The constant $c(E, F, d)$ could be given presicely by this way, but I do not care about it. \square

4 Logvinenko-Sereda theorem

Next we formulate some result that provide further evidence of the non-concentration property of functions with Fourier support on B_1 .

4.1 A toy model

Theorem 12 *Let $\alpha > 1$ and suppose that $S \subset \mathbb{R}^d$ satisfies,*

$$|S \cap B| < \alpha|B|, \text{ for all balls } B \text{ of radius } 1. \quad (19)$$

If $f \in L^2(\mathbb{R}^d)$ satisfies $\text{supp}(\hat{f}) \subset B(0, 1)$ then

$$\|f\|_{L^2(S)} \leq \delta(\alpha)\|f\|_2 \quad (20)$$

Where $\delta(\alpha) \rightarrow 0$ as $\alpha \rightarrow 0$.

PROOF: This is a easy corollary of the argument I give in the proof of Amerein-Berthier theorem 10. \square

4.2 A refine version

Theorem 13 *Suppose that a measurable set $E \subset \mathbb{R}^d$ satisfies the following "thinkness" condition: there exists $\gamma \in (0, 1)$ such that*

$$|E \cap B| > \gamma|B| \text{ for all balls } B \text{ of radius } R^{-1}. \quad (21)$$

where $R > 0$ is arbitrary but fixed. Assume that $\text{supp}(\hat{f}) \subset B(0, R)$. Then

$$\|f\|_{L^2(\mathbb{R}^d)} \leq C\|f\|_{L^2(E)}. \quad (22)$$

where the constant C depends only on d and γ .

Remark 14 *This proof need some very good estimate come from several complex variables.*

5 The Malgrange-Ehrenpreis theorem

Theorem 15 *Let Ω be a bounded domain in \mathbb{R}^d and let $p \neq 0$ be a polynomial, Then, for all $g \in L^3(\Omega)$, there exists $f \in L^2(\Omega)$ such that $p(D)f = g$ in a distribution sence.*