Uncertainty Principles

Master's Thesis in Mathematics

by Jens Gerlach Christensen

May 28, 2003

Institute for Mathematical Sciences
Department of Mathematics
University of Copenhagen

Contents

Preface iii

I Preliminary Results **1**

1 Basic Analysis **2**
1.1 The Fourier Transform 2
1.2 Special Functions 3
1.3 Convex Functions 5

2 The Fourier Transform on $\mathbf{L_p}$-spaces **9**
2.1 Hausdorff-Young's Inequality 9
2.2 Fundamental Theorem of Calculus 16

II Uncertainty Principles in Fourier Analysis **18**

3 Quantitative Uncertainty Principles **19**
3.1 Heisenberg's Uncertainty Principle 19
3.2 Variation of Heisenberg's Uncertainty Principle 23

4 Qualitative Uncertainty Principles **30**
4.1 Hardy's Theorem 30
4.2 Variation of Hardy's Theorem 35
4.3 Functions With Support on Measurable Sets 40
4.4 Almost Vanishing Functions 42

5 Localised Uncertainty Principles **46**
5.1 An Uncertainty Principle in One Dimension 46
5.2 An Uncertainty Principle in Several Dimensions 49

III Uncertainty Principles for Operators **51**

6 Operators and Their Adjoints **52**
6.1 Self-adjoint and Skew-adjoint Operators 52
6.2 Symmetric and Normal Operators 55

7 Operators Generated by Lie-groups **59**
7.1 Invariant Measure on a Lie-group 59
7.2 Lie Theory 63
7.3 Smooth Vectors 64
7.4 The Gårding Subspace 65
7.5 Representation of a Lie-algebra 69
7.6 The Uncertainty Principle 76

IV Applications of Uncertainty Principles **80**

8 Signal Analysis **81**
8.1 Uncertainty Principles 81
8.2 Recovering Missing Segments of a Bandlimited Signal 82

9 Quantum Mechanics **84**
9.1 The Postulates of Quantum Mechanics in One Dimension 84
9.2 The Heisenberg Uncertainty Principle in One Dimension 86
9.3 The Uncertainty Principle for Angular Momentum and Angle 87

List of Symbols **88**

Bibliography **90**

Preface

Classical uncertainty principles give us information about a function and its Fourier transform. If we try to limit the behaviour of one we lose control of the other. Uncertainty principles have implications in two main areas: quantum physics and signal analysis. In quantum physics they tell us that a particles speed and position cannot both be measured with infinite precision. In signal analysis they tell us that if we observe a signal only for a finite period of time, we will lose information about the frequencies the signal consists of.

In this thesis I will first explore some classical uncertainty principles and then try to find generalisations.

The first part of this thesis presents some classical results in analysis. There is information about Fourier transforms, convex functions and tempered distributions.

The second part first investigates Heisenberg's uncertainty principle and then takes a look at variations thereof. Then it moves on to a classical uncertainty principle by Hardy. Hardy's Theorem is an instance of what will be called *qualitative uncertainty principles.* Three other principles of this type are explored, with the first a direct variation of Hardy's Theorem. The second part is concluded by a look at local versions of Heisenberg's uncertainty principle. The classical principle concerns a function whose Fourier transform is concentrated around a single point. But something similar can be established if the Fourier transform is concentrated around countably many evenly distributed points.

An uncertainty principle can be proved for certain operators, and this is the focus of the third part of the thesis. I first prove an uncertainty relation, from which a version of Heisenberg's uncertainty principle follows. This version holds for a smaller set of functions than the one derived in the second part. The domains for the operators involved are very important, and nice properties arise when we restrict ourselves to operators that are generated by Lie-groups. From these I again deduce Heisenberg's uncertainty principle, and this time it holds for the same set of functions as the original principle.

In the fourth and last part I look at applications of uncertainty principles to signal analysis and quantum mechanics.

The thesis was written while living in Galway in Ireland. I would like to thank the maths department at N.U.I. Galway for letting me use their library and a desk in the post-grad room. Especially Dr. Johnny Burns, Dr. Ray Ryan and Dr. Tony Christofides were of great help. Thanks to the people in the post-grad room for a great atmosphere and our trips to the pub.

Last but definately not least I would like to thank Associate Professor Henrik Schlichtkrull for his kind supervision. Thanks for pointing me in the direction of uncertainty principles and still agreeing to supervise me while I was abroad.

Jens Gerlach Christensen
May 2003

Part I

Preliminary Results

Chapter 1

Basic Analysis

Before I can state and prove uncertainty principles I will need a basic amount of analysis. Many results are stated without proof and are only included to make this thesis self-contained.

1.1 The Fourier Transform

In this section the main results about the Fourier transform are presented, but no proofs are made. A reference for this section is [Rud91, Chapter 7].

For vectors $x, y \in \mathbb{R}^n$ I will write the normal scalar product as

$$xy = x \cdot y = x_1 y_1 + \cdots + x_n y_n.$$

Then for functions $f : \mathbb{R}^n \to \mathbb{C}$ define

Definition 1.1. For $f \in L_1(\mathbb{R}^n)$ its Fourier transform can be defined by

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} f(x) e^{-ixy} dx, \qquad y \in \mathbb{R}.$$

The following theorem is taken from [Rud91, 7.9]

Theorem 1.2 (Plancherel). There is a linear isometry Ψ of $L_2(\mathbb{R}^n)$ onto $L_2(\mathbb{R}^n)$ which is uniquely determined by the requirement that

$$\Psi f = \hat{f} \qquad \text{for every } f \in \mathcal{S}.$$

Here $\mathcal{S}$ denotes the Schwartz space of rapidly decreasing functions from section 7.3 in [Rud91].

From the proof of Plancherel's Theorem in [Rud91], we get the following result

Theorem 1.3 (Parseval). For $f, g \in L_2(\mathbb{R}^n)$ it holds that

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx = \int_{\mathbb{R}^n} \hat{f}(y)\overline{\hat{g}(y)}dy.$$

From now on this isometry will be called the Fourier transform, and denoted $\hat{f}$ for $f \in L_2$. From [Rud87, Thm. 9.14] it is known that given $\hat{f}$ we can find f by the following inverse Fourier transform:

Theorem 1.4. If $f \in L_2(\mathbb{R}^n)$ and $\hat{f} \in L_1(\mathbb{R}^n)$ then

$$f(x) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} \hat{f}(y)e^{ixy}dy$$

almost everywhere.

Theorem 1.5. For $f \in \mathcal{S}$ it holds that $\widehat{\partial_j f} = iy_j\hat{f}$.

In [Rud91] the Fourier transform is generalized to the space of tempered distributions (the dual space to the Schwartz space). Let $\mathcal{F}$ denote the Fourier transform on the space $\mathcal{S}'$ of tempered distributions. For $f \in \mathcal{S}'$ the partial distributional derivative $\partial_j f$ of f can be defined as $\mathcal{F}^{-1}(iy_j\hat{f}(y))$ according to [Rud91, Thm. 7.15(b)]. Here y_j is the jth coordinate of y. Since $f \in L_r(\mathbb{R}^n)$ is a tempered distribution by [Rud91, Ex. 7.12], let me define the following

Definition 1.6. If f is in $L_r(\mathbb{R}^n)$ we define the partial distribution derivative $\partial_j f$ as $\mathcal{F}^{-1}(iy_j f(y))$ where y_j is the jth coordinate of y.

This is all the theory of Fourier analysis that will be needed to read this thesis.

1.2 Special Functions

In this section I will present some functions that will be needed later in connection with uncertainty principles. Before I introduce these functions I will note that

$$\int_{\mathbb{R}} e^{-x^2}dx = \sqrt{\pi} \tag{1.1}$$

This is shown by the following calculation using spherical coordinates

$$\begin{aligned}\Big(\int_{\mathbb{R}} e^{-x^2}dx\Big)^2 &= \Big(\int_{\mathbb{R}} e^{-x^2}dx\Big)\Big(\int_{\mathbb{R}} e^{-y^2}dy\Big)\\ &= \int_{\mathbb{R}^2} e^{-(x^2+y^2)}dxdy\\ &= \int_0^{2\pi}\int_0^{\infty} re^{-r^2}drd\theta\\ &= 2\pi\frac{1}{2}[-e^{-r}]_0^{\infty} = \pi.\end{aligned}$$

Now I am ready to present the first class of functions.

1.2.1 Hermite Functions

Definition 1.7. The Hermite functions are defined as

$$h_n(x) = \frac{(-1)^n}{n!} e^{x^2/2} D^n (e^{-x^2}) \tag{1.2}$$

where D is the differential operator.

Proposition 1.8. The Hermite functions satisfy the following recursion formula

$$h_n'(x) - x h_n(x) = -(n+1) h_{n+1}(x), \qquad x \geq 0 \tag{1.3}$$

Proof. This is shown by a simple differentiation:

$$\begin{aligned} h_n'(x) &= \frac{(-1)^n}{n!}(D(e^{x^2/2}) D^n(e^{-x^2}) + e^{x^2/2} D^{n+1}(e^{-x^2})) \\ &= \frac{(-1)^n}{n!} x e^{x^2/2} D^n(e^{-x^2}) - \frac{(-1)^{n+1}}{(n+1)!}(n+1) e^{x^2/2} D^{n+1}(e^{-x^2}) \\ &= x h_n(x) - (n+1) h_{n+1}(x). \end{aligned}$$

Rearranging yields the desired recursion formula. □

Proposition 1.9. The Hermite functions and their Fourier transforms are related in the following way

$$\hat{h}_n(x) = (-i)^n h_n(x)$$

Proof. First I will show that $\hat{h}_0 = h_0$. By (1.2), $h_0(x) = e^{-x^2/2}$ and

$$\begin{aligned} \hat{h}_0'(y) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2} e^{-iyx}(-ix) dx \\ &= \frac{i}{\sqrt{2\pi}} \int_{\mathbb{R}} [e^{-x^2}/2]' e^{-iyx} dx \\ &= \frac{i}{\sqrt{2\pi}} [e^{-x^2/2} e^{-iyx}]_{-\infty}^{\infty} - \frac{i}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2} e^{-iyx}(-iy) dx \\ &= \frac{-iy}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2} e^{-iyx} dx \\ &= -y \hat{h}_0(y). \end{aligned}$$

This shows that $\hat{h}_0$ satisfies the differential equation $u'(x) = -xu(x)$. By separation of variables the solution is $u(x) = u(0)e^{-x^2/2}$, which means that

$$\hat{h}_0(y) = e^{-x^2/2}$$

since

$$\hat{h}_0(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2} dx = 1$$

according to (1.1). I will now prove that h_n and $i^n\hat{h}_n$ satisfy the same recursion formula (1.3), by applying the Fourier transform to the formula:

$$\begin{aligned}-(n+1)i^{n+1}\hat{h}_{n+1} &= i^{n+1}\widehat{h_n'} - i^n\widehat{[ixh_n]}\\ &= i^{n+1}iy\hat{h}_n + (-i)^n(\hat{h}_n)'\\ &= -yi^n\hat{h}_h + [(-i)^n\hat{h}_n]'.\end{aligned}$$

Since $h_0 = \hat{h}_0$ and $i^n\hat{h}$ satisfy the same recursion formula I conclude that $\hat{h}_n = (-i)^n h_n$. $\square$

Remark 1.10. Proposition 1.9 states that the Hermite functions are eigenfunctions of the Fourier transform. The eigenvalues are $\sqrt{2}$ times a power of $-i$.

Remark 1.11. The Hermite function h_n is a polynomial of degree n multiplied by $e^{-x^2/2}$. This is seen by using induction. The polynomials for h_i for $i \in \{0,\dots,n\}$ form a basis for polynomials of degree n. I proved these facts in a project in first year of college. For a reference see [DM72, p. 98], where the claims are left as exercises with hints. Note that [DM72] defines the Hermite functions as $\frac{(-1)^n}{n!}\exp(\pi x^2)D^n\exp(-2\pi x^2)$.

1.2.2 Gaussian Functions

Here I will make a very short presentation of Gaussian functions.

A function is called Gaussian if it is of the form $\varphi(x) = e^{-ax^2}$ with $a > 0$. By calculations just like the ones for h_0 in the proof of Proposition 1.9 I get

Proposition 1.12. Let $\varphi(x) = e^{-ax^2}$ with $a > 0$ and $x \in \mathbb{R}$, then its Fourier transform is

$$\hat{\varphi}(y) = \frac{1}{\sqrt{a}}e^{-y^2/4a}$$

1.3 Convex Functions

In this section I will demonstrate Jensen's inequality. It holds for convex functions. Later I will need to show that certain functions are convex. Therefore I present a result that will help us decide if a function is convex and I include examples.

Definition 1.13. Let $-\infty \le a \le b \le \infty$. A function $f :]a,b[\to \mathbb{R}$ is called convex if the inequality

$$f((1-\gamma)x + \gamma y) \le (1-\gamma)f(x) + \gamma f(y) \tag{1.4}$$

holds whenever $x, y \in]a,b[$ and $\gamma \in [0,1]$.

The requirement is that for $a < s < t < u < b$ the point $(t, f(t))$ will be on or under the line ℓ_{su} that connects $(s, f(s))$ and $(u, f(u))$. This means that the line ℓ_{st} connecting $(s, f(s))$ and $(t, f(t))$ has slope less than or equal to the slope of ℓ_{su}. Also the slope of the line ℓ_{tu} connecting $(t, f(t))$ and $(u, f(u))$ has slope greater than or equal to ℓ_{su}, since ℓ_{tu} and ℓ_{su} both end in $(u, f(u))$. I have illustrated this on Figure 1.1. It means that, for $a < s < t < u < b$, it holds that

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(t)}{u - t}. \tag{1.5}$$

Let α_t be the supremum of the left hand side of (1.5). Then $f(s) \geq f(t) + \alpha_t(s - t)$ for $a < s < t$. Also α_t is smaller than the right hand side, which gives $f(u) \geq f(t) + \alpha_t(u - t)$ for $t < u < b$. Thus for all $s, t \in]a, b[$ there is a constant α_t such that

$$f(s) \geq f(t)\alpha_t(s - t). \tag{1.6}$$

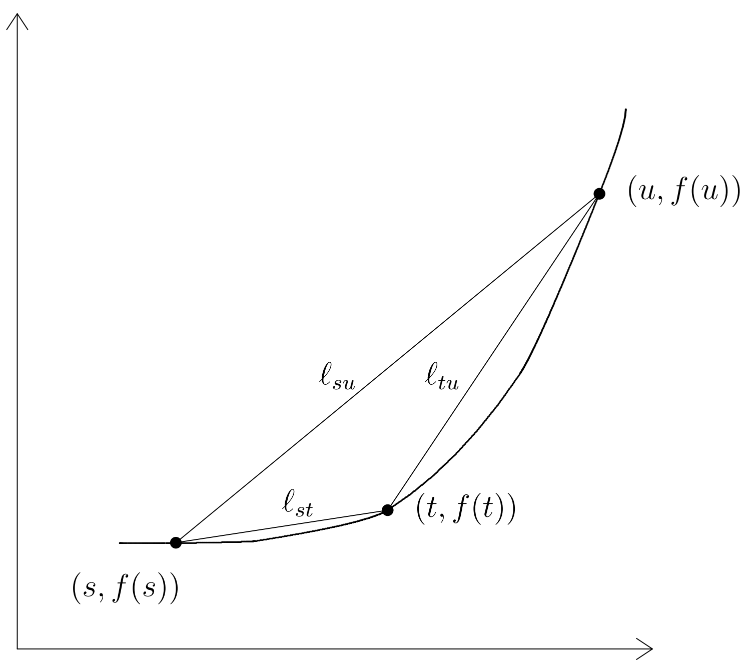

Figure 1.1: Convex function.

I note the following result taken from [Rud87, Thm. 3.2]

Lemma 1.14. If f is convex on $]a, b[$ then it is continuous on $]a, b[$.

The following theorem is Jensen's inequality and is the same as Theorem 3.3 in [Rud87], and I will skip the proof.

Theorem 1.15 (Jensen's Inequality). Let μ be a probability measure on $\mathbb{R}$ ($\mu(\mathbb{R}) = 1$). If f is a real function in $L_1(\mathbb{R})$ such that $a < f(x) < b$ for all $x \in \mathbb{R}$ and g is convex on $]a, b[$ then

$$g\Big(\int_{\mathbb{R}} f d\mu\Big) \leq \int_{\mathbb{R}} (g \circ f) d\mu.$$

I will need an easy way to show that certain functions are convex, so I state

Proposition 1.16. If f is two times differentiable on $]a,b[$ and $f'' \geq 0$ on this interval, then f is convex.

Proof. If f is differentiable, then the mean value theorem tells us that the inequality (1.5) is equivalent with $f'(s) \leq f'(t)$ for $s < t$. So f' is a non-decreasing function. If f' is also differentiable then this is equivalent with $f''(x) \geq 0$ for all $x \in]a,b[$. □

Example 1.17. (a) It is readily seen, that $t \mapsto t \log t$ is convex on $]0,\infty[$.

(b) Let $\log^+ x = \max\{0, \log x\}$ then $f : t \mapsto t^2 \log^+ t$ is convex on $[0,\infty[$ with $f(0) = 0$. f is 0 on $]0,e[$ and thus convex there. According to Proposition 1.16 it is also convex on $]e,\infty[$. Since $f(e) = 0$ and $0 = f(x) \leq f(y)$, when $x \in]0,e[$ and $y \in]e,\infty[$, it follows that f is convex on $[0,\infty[$.

(c) Let $\log^- x = \min\{0, \log x\}$ then $f : t \mapsto t^2(-\log^- t + 3/2)$ is convex. $t^2 \log^- t$ is 0 on $[e,\infty[$ and f is continuous in e. So as above I need only look at the interval $]0,e[$. There $f(t) = t^2(-\log t + 3/2)$ and $f''(t) = -2\log t \geq 0$. So f is convex on $[0,\infty[$.

Inspired by this I now extend to functions $f : \mathbb{R}^n \to \mathbb{R}$. To generalize I have to also introduce convex sets.

Definition 1.18. An open subset U of $\mathbb{R}^n$ is called convex if for any $x, y \in U$ the line $(1-\gamma)x + \gamma y$ is in U for all $\gamma \in [0,1]$.

A real function f on a convex set $U \subseteq \mathbb{R}^n$ is called convex if the inequality

$$f((1-\gamma)x + \gamma y) \leq (1-\gamma)f(x) + \gamma f(y)$$

holds whenever $x, y \in U$ and $\gamma \in [0,1]$.

Note that a convex function on $U \subseteq \mathbb{R}^n$ is convex in every variable (just pick x and y such that they only differ in one variable). Therefore if $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ then by recursively using (1.6) I get that there exists $c = (c_1, \ldots, c_n)$ such that

$$\begin{aligned} g(x_1, \ldots, x_n) &\geq g(y_1, x_2, \ldots, x_n) + c_1(x_1 - y_1) \\ &\geq \ldots \\ &\geq g(y_1, \ldots, y_n) + c_1(x_1 - y_1) + \cdots + c_n(x_n - y_n). \end{aligned}$$

This means that (1.6) also holds in $\mathbb{R}^n$: there exists $c \in \mathbb{R}^n$ such that for all x, y in a convex set $U \subseteq \mathbb{R}^n$ it holds

$$g(x) \geq g(y) + c(x - y), \tag{1.7}$$

when g is convex on U.

I am now ready to prove the extended form of Jensen's Inequality

Theorem 1.19 (Jensen's Inequality). Let μ be a probability measure on $\mathbb{R}$ ($\mu(\mathbb{R}) = 1$) and let g be a convex function on a set U. If $f : \mathbb{R} \to U$ is a function with each coordinate function in $L_1(\mathbb{R})$ then

$$g\Big(\int_{\mathbb{R}} f d\mu\Big) \leq \int_{\mathbb{R}} (g \circ f) d\mu.$$

Proof. Let $t = \int_{\mathbb{R}} f(x)d\mu(x)$. Then $t \in U$ since every coordinate function f_i is contained in an interval $]a_i, b_i[$, and thus $t \in]a_i, b_i[$. Now let z and set $x = f(z)$ and $y = t$ in (1.7):

$$g(f(z)) \geq g(t) + c(f(z) - t).$$

Integrating both sides and using the special value of t gives

$$\int_{\mathbb{R}} g(f(z))d\mu(z) \geq \int_{\mathbb{R}} g(t) + c(f(z) - t)d\mu(z) = g(t).$$

since μ is a probability measure. This is the desired result. □

A function on a convex subset of $\mathbb{C}$ is convex if it considered as a function of its real and imaginary parts is convex. This is equivalent to convexity in $\mathbb{R}^2$.

Example 1.20. The function $z \mapsto |z|$ is convex on $\mathbb{C}$. The function $t \mapsto t^2 \log t$ is growing on $]0, \infty[$ (since it is convex according to Example 1.17(b)) and therefore the function $z \mapsto |z|^2 \log^+ |z|$ is convex on $\mathbb{C}$.

The same is seen to hold for $z \mapsto |z|^2(-\log^- |z| + 3/2)$, since $t \mapsto t^2(-\log^- t + 3/2)$ is growing on $]0, \infty[$.

Chapter 2

The Fourier Transform on $\mathrm{L_p}$-spaces

In this chapter I will generalize the Fourier transform to a bigger class of functions, and show some results, which will be useful later in this thesis.

2.1 Hausdorff-Young's Inequality

Here I will generalize the Fourier-transform to L_p for $1 < p < 2$ and prove a classical inequality by Titchmarsh. The contents of this section are taken from chapter 4 p. 96ff in [Tit37].

I need this special version of Hölder's inequality

Theorem 2.1 (Hölder's inequality). Let a_i, b_i with $i = 1, \ldots, n$ be non-negative numbers. Let $p, q > 1$ be real numbers with $1/p + 1/q = 1$ then

$$\sum_{i=1}^{n} a_i b_i \leq \Big(\sum_{i=1}^{n} a_i^p\Big)^{1/p} \Big(\sum_{i=1}^{n} b_i^q\Big)^{1/q}.$$

Equality only holds if there is a constant c such that $a_i^p = c b_i^q$ for all $i = 1, \ldots, n$.

Proof. This is a special case of Theorem 3.5 p. 63 in [Rud87] using the discrete measure on $\mathbb{Z}$. The equality condition follows from the discussion on p. 65 immediatly after the proof. □

I will also need the following lemma (with a very long proof)

Lemma 2.2. Assume $1 < p < 2$ and $1/p + 1/q = 1$. For all $c_m \in \mathbb{C}$ with $m \in \{-n, \ldots, n\}$ it holds that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Big| \sum_{m=-n}^{n} c_m e^{imx} \Big|^q \leq \Big(\sum_{m=-n}^{n} |c_m|^p \Big)^{1/(p-1)}.$$

Proof. Let f be a function in $L_1(\mathbb{T})$, where $\mathbb{T} = [-\pi, \pi]$, and let c_m for $m \in \mathbb{Z}$ be the Fourier coefficients

$$c_m = \frac{1}{2\pi} \int_{\mathbb{T}} f(x) e^{-imx} dx.$$

Define

$$S_p(f) = \Big(\sum_{-\infty}^{\infty} |c_m e^{imx}|^q \Big)^{1/q}$$

and

$$J_p(f) = \Big(\frac{1}{2\pi} \int_{\mathbb{T}} |f(x)|^p dx \Big)^{1/p}.$$

The statement of the lemma is that $J_q(f) \leq S_p(f)$ for any trigonometric polynomial $f_n = \sum_{m=-n}^{n} c_m e^{imx}$. Assume that not all c_m are zero. I divide the proof into six steps.

1. For f define $f_n = \sum_{m=-n}^{n} c_m e^{imx}$. I will first show that $S_q(f_n)/J_p(f)$ has an upper bound for all f. At least one of $|c_m|^q$ is greater that or equal to $S_q^q(f_n)/(2n+1)$ or else

$$S_q^q(f_n) = \sum_{m=-n}^{n} |c_m|^q < \sum_{m=-n}^{n} S_q^q(f_n)/(2n+1) = S_q^q(f_n),$$

which cannot be true. So for this c_m holds

$$\begin{aligned} \frac{S_q(f_n)}{(2n+1)^{1/q}} \leq |c_m| &= \Big| \frac{1}{2\pi} \int_{\mathbb{T}} f(x) e^{-imx} dx \Big| \\ &\leq \frac{1}{2\pi} \int_{\mathbb{T}} |f(x)| dx \\ &\leq \Big(\frac{1}{2\pi} \int_{\mathbb{T}} |f(x)|^p \Big)^{1/p} \Big(\frac{1}{2\pi} \int_{\mathbb{T}} 1^q dx \Big)^{1/q} \\ &= J_p(f). \end{aligned}$$

The last inequality is Hölder's inequality. This shows that for all p and n there is a least upper bound $M = M_p$ on $S_q(f_n)/J_p(f)$ that does not depend on f. It also shows that

$$M_p \leq (2n+1)^{1/q}. \tag{2.1}$$

2. For given coefficients $\{c_m\}$ define $f_n = \sum_{m=-n}^{n} c_m e^{imx}$. I will show that there is an upper bound on $J_q(f_n)/S_p(f_n)$ for all possible sets $\{c_m\}$. Let

$$g(x) = |f_n(x)|^{q-1} \overline{\operatorname{sgn} f_n(x)}$$

where $\operatorname{sgn} z = z/|z|$. Also set

$$\gamma_m = \frac{1}{2\pi} \int_{\mathbb{T}} g(x) e^{-imx} dx$$

then

$$\sum_{m=-n}^{n} c_m\gamma_m = \frac{1}{2\pi}\sum_{m=-n}^{n} c_m \int_{\mathbb{T}} g(x)e^{-imx}dx$$
$$= \frac{1}{2\pi}\int_{\mathbb{T}} g(x)\sum_{m=-n}^{n} c_m e^{-imx}dx = \frac{1}{2\pi}\int_{\mathbb{T}} g(x)f(x)dx.$$

Take only the $2n+1$ first coefficients γ_m to define

$$g_n(x) = \sum_{m=-n}^{n} \gamma_m e^{imx}.$$

Then, using Hölder's inequality, the following calculations hold

$$\begin{aligned}
J_q^q(f_n) &= \frac{1}{2\pi}\int_{\mathbb{T}} f_n(x)g(x)dx = \sum_{m=-n}^{n} c_m\gamma_m \\
&\leq \sum_{m=-n}^{n} |c_m\gamma_m| \leq \Big(\sum_{m=-n}^{n}|c_m|^p\Big)^{1/p}\Big(\sum_{m=-n}^{n}|\gamma_m|^q\Big)^{1/q} \\
&= S_p(f_n)S_q(g_n) \leq MS_p(f_n)J_p(g) \\
&= MS_p(f_n)J_q^{1/(p-1)}(f_n) \qquad (2.2)
\end{aligned}$$

The last inequality follows since $S_q(f_n) \leq MJ_p(f)$ for all f by definition of M. The last equality follows by direct calculation using the definition of g in the following manner

$$J_p^p(g) = \frac{1}{2\pi}\int_{\mathbb{T}} |f_n(x)|^{p(q-1)}dx = \frac{1}{2\pi}\int_{\mathbb{T}} |f_n(x)|^q dx = J_q^q(f_n) \qquad (2.3)$$

and then using $q/p = 1/(p-q)$. Dividing by $J_q^{1/(p-1)}(f_n)$ on both sides we get

$$J_q(f_n) \leq MS_p(f_n)$$

which tells us that $J_q(f_n)/S_p(f_n)$ has a least upper bound M' and $M' \leq M$.

3. Now I show that $M = M' \geq 1$ by using the function

$$h_n(x) = \sum_{m=-n}^{n} |c_m|^{q-1}\overline{\operatorname{sgn} c_m}e^{-imx}.$$

Since terms of the form $c_m e^{imx}|c_l|^{q-1}\overline{\operatorname{sgn} c_l}e^{-ilx}$ integrate to zero unless $m = l$, it follows that

$$\begin{aligned}
S_q^q(f_n) &= \frac{1}{2\pi}\int_{\mathbb{T}} f_n(x)\bar{h}_n(x) \leq J_p(f_n)J_q(h_n) \\
&\leq M'J_p(f_n)S_p(h_n) = M'J_p(f_n)S_q^{q-1}(f_n)
\end{aligned}$$

Which shows that $M \leq M'$ and we then conclude that $M = M'$. Also $M \geq 1$, since if $f(x) = 1$ then $S_q(f_n) = J_p(f) = 1$.

4. I now investigate functions for which the maximum of $J_q(f_n)/S_p(f_n)$ is attained. I find the value of a coefficient d, that comes from the equality in Hölder's inequality. $J_q(f_n)/S_p(f_n)$ is continuous in the coefficients c_m, so there exist coefficients $\{c_m\}$ such that the supremum $M' = M$ is equal to $J_q(f_n)/S_p(f_n)$. Such an f_n is called a maximal polynomial. Inserting this for M in the last term of (2.2) tells us that equality must hold between all terms in (2.2). This gives us the following equality

$$S_q(g_n) = MJ_p(g). \tag{2.4}$$

The application of Hölder's inequality in (2.2) then tells us that for some constant d, $|c_m|^p = d|\gamma_m|^q$ for all m. So $S_p^p(f_n) = dS_q^q(g_n)$ and since f_n is maximal

$$S_p(f_n) = J_q(f_n)/M = J_p^{p-1}(g)/M = S_q^{p-1}(g_n)/M^p$$

The second equality follows from (2.3) and the last follows from (2.4). These two equations give

$$d = M_p^{-p^2} S_q^{p^2-p-q}(g_n) \tag{2.5}$$

which shows that d depends on p only.

5. I will now show that M_p is a growing sequence for $p \to \infty$. Define $r_1 = 2/(3-p)$ and $s_1 = 2q-2$, then $1/r_1 + 1/s_1 = 1$. Bessel's inequality then gives

$$S_2^2(g_n) \leq \frac{1}{2\pi}\int_{\mathbb{T}} |g(x)|^2 dx = J_{s_1}^{s_1}(f_n) \leq M_{r_1}^{s_1} S_{r_1}^{s_1}(f_n). \tag{2.6}$$

Using $|c_m|^p = d|\gamma_m|^q$ we get

$$S_{r_1}^{s_1} = \Big(\sum_{m=-n}^{n} d^{r_1/p}|\gamma_m|^{qr_1/p}\Big)^{s_1/r_1} = d^{s_1/p} S_{r_1/(p-1)}^{s_1/(p-1)}(g_n)$$

since $qr_1/p = r_1/(p-1)$ and $(s_1/r_1)(r_1/(p-1)) = s_1/(p-1)$. Inserting the value for d gives

$$S_{r_1}^{s_1} = M_p^{-s_1 p} S_q^{s_1(p-1-1/(p-1))}(g_n) S_{r_1/(p-1)}^{s_1/(p-1)}(g_n). \tag{2.7}$$

Hölder's inequality gives

$$\begin{aligned} S_{r_1/(p-1)}^{r_1/(p-1)}(g_n) &= \sum_{m=-n}^{n} |\gamma_m|^{r_1/(p-1)} = \sum_{m=-n}^{n} |\gamma_m|^{r_1(p-1)}|\gamma_m|^{qr_1(2-p)} \\ &\leq \Big(\sum_{m=-n}^{n} |\gamma_m|^2\Big)^{r_1(p-1)/2} \Big(\sum_{m=-n}^{n} |\gamma_m|^q\Big)^{r_1(2-p)} = S_2^{r_1(p-1)}(g_n) S_q^{qr_1(2-p)}(g_n) \end{aligned}$$

so

$$S_{r_1/(p-1)}^{s_1/(p-1)}(g_n) \leq S_2^{s_1(p-1)}(g_n) S_q^{qs_1(2-p)}(g_n) = S_2^2(g_n) S_q^{qs_1(2-p)}(g_n). \tag{2.8}$$

Now (2.6),(2.7) and (2.8) give

$$\begin{aligned} S_2^2(g_n) &\le M_{r_1}^{s_1} S_{r_1}^{s_1}(f_n) = M_{r_1}^{s_1} M_p^{-s_1 p} S_q^{s_1(p-1-1/(p-1))}(g_n) S_{r_1/(p-1)}^{s_1/(p-1)}(g_n) \\ &\le M_{r_1}^{s_1} M_p^{-s_1 p} S_q^{s_1(p-1-1/(p-1))}(g_n) S_2^2(g_n) S_q^{qs_1(2-p)}(g_n) = M_{r_1}^{s_1} M_p^{-s_1 p} S_2^2(g_n) \end{aligned}$$

where the last equality follows since $s_1(p-1-1/(p-1)) + qs_1(2-p) = 0$. Therefore $1 \le M_{r_1}^{s_1} M_p^{-s_1 p}$ and since $M_p \ge 1$

$$M_{r_1} \ge M_p^p \ge M_p.$$

6. In this last step I conclude that $M = 1$. If we set $r_2 = 2/(3-r_1)$ and $s_2 = 2s_1 - 2$ we can make the same calculations, as in the previous step, to get $M_{r_2} \ge M_{r_1} \ge M_p$. Setting $r_j = 2/(3-r_{j-1})$ and $s_j = 2s_{j-1} - 2$ we can make a non-decreasing sequence (M_{r_j}). It holds that $s_j - 2 = 2(s_{j-1}-2) = 4(s_{j-2}-2) = \cdots = 2^j(q-2)$. Since $1 < p < 2$, q is greater that 2 and therefore $s_j \to \infty$ for $j \to \infty$. And by (2.1), $M_p \le M_{r_j} \le (2n+1)^{1/s_j} \to 1$. This together with $M_p \ge 1$ shows that $M_p = 1$ and therefore $J_q^q(f_n) \le S_p^q(f_n)$ which is the statement of the lemma. □

Theorem 2.3 (Hausdorff-Young's inequality). Assume $f \in L_p$ and $1 < p < 2$. Let q satisfy $1/p + 1/q = 1$ then

$$F(y,a) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} f(x) e^{-ixy} dx$$

converges to a function, which we will call $\hat{f}$, in $L_q(\mathbb{R})$ for $a \to \infty$, and

$$\|\hat{f}\|_q \le (2\pi)^{1/q-1/2} \|f\|_p.$$

Proof. Given $s > 0$ and $n \in \mathbb{N}$ then for $l \in \mathbb{N}$ let

$$A_l = \int_{l/s}^{(l+1)/s} f(x) dx$$

and

$$B_n(x) = \sum_{l=-n}^{n} A_l e^{-ilx}.$$

1. First I will show that $B_n(x) \to F(x,a)$ for $n \to \infty$. For a certain a, let n be the biggest integer lower than sa, that is $n = \lfloor sa \rfloor$ then as $s \to \infty$

$$\begin{aligned} &|B_n(x) - F(x,a)| \\ &= \Big| \sum_{l=-n}^{n} \int_{l/s}^{(l+1)/s} f(y)(e^{-ilx/s} - e^{-ixy}) dy - \int_{(n+1)/s}^{a} f(y) e^{-ixy} dy - \int_{-a}^{-n/s} f(y) e^{-ixy} dy \Big| \\ &\le \sum_{l=-n}^{n} \int_{l/s}^{(l+1)/s} |f(y)||e^{-ilx/s} - e^{-ixy}| dy - \int_{(n+1)/s}^{a} |f(y)| dy - \int_{-a}^{-n/s} |f(y)| dy \\ &\le \sum_{l=-n}^{n} \frac{x}{s} \int_{-a}^{a} |f(y)| dy - \int_{(n+1)/s}^{a} |f(y)| dy - \int_{-a}^{-n/s} |f(y)| dy \to 0 \end{aligned}$$

since $n/s \to a$ and $(n+1)/s \to a$ for $s \to \infty$ and

$$|e^{-ilx/s} - e^{-ixy}| \leq x/s \tag{2.9}$$

when $y \in [l/s, (l+1)/s]$. The claim (2.9) is derives as follows: Let $y = (l+t)/s$ for $t \in [0,1]$

$$|e^{-ilx/s} - e^{-ixy}|^2 = |1 - e^{-ix(y-l/s)}|^2 = |1 - e^{-ixt/s}|^2 = 2 - 2\cos(tx/s).$$

The last term has maximum 4, so if $x/s \geq \pi$ the claim follows. If $x/s < \pi$ the maximum is attained when $t = 1$i, and therefore, using the Taylor expansion

$$\cos(x/s) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n)!}(x/s)^{2n} = 1 - (x/s)^2 + \frac{(x/s)^4}{4!} - \frac{(x/s)^6}{6!} + \cdots,$$

we see that $2 - 2\cos(x/s) - (x/s)^2 \leq 0$. Which proves (2.9).

2. Now I show that $F(x,a)$ is in L_q. Using Hölder's inequality for integrals (see [Rud87, Thm. 3.5]) the following holds for A_l

$$|A_l| \leq \int_{l/s}^{(l+1)/s} |f(y)| dy \leq \Big(\int_{l/s}^{(l+1)/s} |f(y)|^p dy\Big)^{1/p} \Big(\int_{l/s}^{(l+1)/s} dy\Big)^{1/q} \tag{2.10}$$

$$= ((l+1)/s - l/s)^{1/q} \Big(\int_{l/s}^{(l+1)/s} |f(y)|^p dy\Big)^{1/p} \tag{2.11}$$

so

$$|A_l|^p = s^{1-p} \int_{l/s}^{(l+1)/s} |f(y)|^p dy. \tag{2.12}$$

If $X \leq s\pi$ then

$$\int_{-X}^{X} |B_n(x)|^q dx \leq \int_{-s\pi}^{s\pi} |B_n(x)|^q dx = s\int_{-\pi}^{\pi} \Big|\sum_{l=-n}^{n} A_l e^{-ilx/s}\Big|^q dx \leq 2\pi s \Big(\sum_{l=-n}^{n} |A_l|^p\Big)^{1/(p-1)}$$

where the last inequality is due to Lemma 2.2. By (2.12) I get

$$\leq 2\pi s \Big(\sum_{l=-n}^{n} s^{1-p} \int_{l/s}^{(l+1)/s} |f(y)|^p dy\Big)^{1/(p-1)} \leq 2\pi \Big(\int_{-b}^{b} |f(y)|^p dy\Big)^{1/(p-1)}$$

Since this holds for all s and $B_n(x) \to F(x,a)$ for $s \to \infty$ then

$$\int_{-X}^{X} |F(x,a)|^q dx \leq 2\pi \Big(\frac{1}{\sqrt{2\pi}^p} \int_{-b}^{b} |f(y)|^p dy\Big)^{1/(p-1)} = (2\pi)^{1-p/2(p-1)} \Big(\int_{-b}^{b} |f(y)|^p dy\Big)^{1/(p-1)}$$

The right hand side does not depend on X so letting $X \to \infty$ shows that

$$\int_{-\infty}^{\infty} |F(x,a)|^q dx \leq (2\pi)^{1-p/2(p-1)} \Big(\int_{-b}^{b} |f(y)|^p dy\Big)^{1/(p-1)} \tag{2.13}$$

so $F(x,a)$ is in $L_q(\mathbb{R})$.

3. I want to show that as $a \to \infty$ $F(x,a)$ has a limit in $L_q(\mathbb{R})$. Since $L_q(\mathbb{R})$ is a complete space it is enough to show that for $a \in \mathbb{N}$, $F(x,a)$ is a Cauchy sequence. Write f as $f = f1_{[-a,a]} + f1_{\mathbb{R}\setminus[-a,a]}$. Then with $b > a$ it holds that

$$\begin{aligned}
&\int_{\mathbb{R}} |F(x,b) - F(x,a)|^q dx \\
&\quad = \int_{\mathbb{R}} \Big| \int_{-b}^{-a} f(y)e^{-ixy}dy + \int_a^b f(y)e^{-ixy}dy + \int_{-a}^{a} f(y)e^{-ixy}dy - F(x,a) \Big|^q dx \\
&\quad = \int_{\mathbb{R}} \Big| \int_{-b}^{-a} f(y)e^{-ixy}dy + \int_a^b f(y)e^{-ixy}dy \Big|^q dx.
\end{aligned}$$

This is the same as finding $F(x,b)$ for a function that is zero on $[-a,a]$. Therefore the inequality (2.13) can be used to get

$$\begin{aligned}
\int_{\mathbb{R}} |F(x,b) - F(x,a)|^q dx &\le 2\pi \Big(\frac{1}{\sqrt{2\pi}} \int_{-b}^{b} |f(y)|^p dy \Big)^{1/(p-1)} \\
&= (2\pi)^{1-p/2(p-1)} \Big(\int_{-b}^{-a} |f(y)|^p dy + \int_a^b |f(y)|^p dy \Big)^{1/(p-1)}
\end{aligned}$$

This tends to zero as $a \to \infty$ and $b \to \infty$, or otherwise f would not be in $L_p(\mathbb{R})$. This shows that for $a \in \mathbb{N}$, the sequence $F(x,a)$ is a Cauchy sequence. Therefore $F(x,a)$ has a limit in $L_q(\mathbb{R})$ as $a \to \infty$. Let us denote this limit $\hat{f}(x)$, then the inequality (2.13) shows that

$$\int_{\mathbb{R}} |\hat{f}(x)|^q dx \le (2\pi)^{1-p/2(p-1)} \Big(\int_{-\infty}^{\infty} |f(y)|^p dy \Big)^{1/(p-1)}.$$

This translates into

$$\|\hat{f}\|_q^q \le (2\pi)^{1-q/2} \|f\|_p^q$$

and taking the qth root we get the desired result. ☐

Remark 2.4. (a) For a Schwartz function f it holds that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(y) e^{ixy} dy$$

and since $\hat{f}$ is a Schwartz function it is in $L_p(\mathbb{R})$. We then have an similar inequality

$$\|f\|_q \le (2\pi)^{1/q-1/2} \|\hat{f}\|_p.$$

(b) The case $p = 1$. Since f is a Schwartz function and thus continuous $\|f\|_\infty = \sup\{|f(x)|\}$, and $|f(x)| \le (2\pi)^{-1/2}\|\hat{f}\|_1$. So in the case where $p = 1$ we can extend Hausdorff-Young's inequality to

$$\|f\|_\infty \le (2\pi)^{-1/2} \|\hat{f}\|_1.$$

(c) The case $p = 2$. For $p = 2$ Titchmarsh [Tit37] has shown that $\hat{f}$ can be defined as

$$\hat{f}(y) = \lim_{a\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} f(x)e^{-ixy}dx.$$

Since this definition is the same as our normal definition on $L_1 \cap L_2$, Plancherels theorem tells us that they are equivalent. In this special case we know that $\|f\|_2 = \|\hat{f}\|_2$.

2.2 Fundamental Theorem of Calculus

Now that I have defined the Fourier transform on general L_p I will also give a generalized version of a well known theorem

Theorem 2.5 (Fundamental Theorem of Calculus). If $f(x)$ and $xf(x)$ are in $L_p(\mathbb{R})$ for $1 < p \leq 2$ then, with $\partial\hat{f} = \widehat{-ixf}$, the following holds for all $a, y \in \mathbb{R}$

$$\hat{f}(y) - \hat{f}(a) = \int_{a}^{y} \partial\hat{f}(t)dt. \tag{2.14}$$

Proof. Let $1/p + 1/q = 1$. Since $-ixf(x)$ is in $L_p(\mathbb{R})$ by assumption then $\partial\hat{f}$ is in $L_q(\mathbb{R})$. It is therefore integrable on any interval $[a, x]$ and therefore the right hand side of (2.14) has meaning.

By Theorem 2.3 (see remark 2.4 for $p = 2$) it is known that for $f \in L_p$

$$\hat{f}(y) = \lim_{b\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} f(t)e^{-iyt}dt.$$

Using this for $-itf(t)$ the right hand side is

$$\int_{a}^{x} \widehat{-itf(t)}(y)dy = \int_{a}^{x} \lim_{b\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} -itf(t)e^{-iyt}dtdy$$

since we integrate over a finite interval we can take the limit outside the integral to get

$$= \lim_{b\to\infty} \int_{a}^{x} \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} -itf(t)e^{-iyt}dtdy$$

Using Fubini's theorem [Rud87, Thm. 8.8] I can swap the order of integration

$$\begin{aligned}
&= \lim_{b\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} -itf(t) \int_{a}^{x} e^{-iyt}dydt \\
&= \lim_{b\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-b}^{b} f(t)[e^{-iyt}]_{a}^{x}dt \\
&= \hat{f}(y) - \hat{f}(a).
\end{aligned}$$

This concludes the proof. $\square$

The proof could also be made using Schwartz functions. The right hand side of (2.14) is defined for all f such that $f, xf \in L_p$, since $\partial \hat{f}$ is in L_q. It is therefore enough to look at Schwartz functions. These functions are differentiable, so the fundamental theorem of calculus [Rud87, Thm. 7.21] holds.

Part II

Uncertainty Principles in Fourier Analysis

Chapter 3

Quantitative Uncertainty Principles

Quantitative uncertainty principles is just another name for some special inequalities. These inequalities give us information about how a function and its Fourier tranform relate. They are called uncertainty principles since they are similar to the classical Heisenberg Uncertainty Principle, which has had a big part to play in the development and understanding of quantum physics. In this chapter I will prove the classical uncertainty principle and a few variations. I will also include a generalisation of these results without proof.

3.1 Heisenberg's Uncertainty Principle

This section is devoted to the classical Heisenberg uncertainty principle.

Lemma 3.1. Let $r, s, t \in \mathbb{R}_+$ and $j \in \{1, 2, \ldots, n\}$. If $f \in L_r(\mathbb{R}^n)$ with partial distributional derivative $\partial_j f \in L_s(\mathbb{R}^n)$ and $x_j f \in L_t(\mathbb{R}^n)$ then there is a sequence of functions g_n in $C_c^\infty(\mathbb{R}^n)$ such that

$$\|g_n - f\|_r + \|\partial_j g_n - \partial_j f\|_s + \|x_j g_n - x_j f\|_t \to 0 \qquad \text{for } n \to 0. \tag{3.1}$$

Proof. The proof is divided into three steps, and I will start with a little sketch of the proof. First I will approximate f with a sequence f_p of functions in $L_r(\mathbb{R}^n)$ with compact support:

$$f_p(x) = k_p(x)f(x) = k(x/p)f(x)$$

where $k : \mathbb{R}^n \to [0, 1]$ is in $C_c^\infty(\mathbb{R}^n)$ and defined by

$$k(x) = \begin{cases} 1 & |x| \leq 1 \\ 0 \leq k(x) \leq 1 & 1 < |x| < 2 \\ 0 & |x| \geq 2. \end{cases}$$

Then for each p I approximate f_p with a sequence $g_{p,q}$ in $C_c^\infty(\mathbb{R}^n)$

$$g_{p,q}(x) = h_q * f_p(x)$$

where $h_q(x) = q^{-1}h(qx)$ with $h \in C_c^\infty(\mathbb{R}^n)$ and $\int_{\mathbb{R}} h(x)dx = 1$. Showing the convergences for each approximation yields the desired result.

Now I will start the actual proof.

1. $f_p \to f$ in $L_r(\mathbb{R}^n)$ since $f_p(x) \to f(x)$ pointwise and $|f_p(x)| \le |f(x)|$ for all $x \in \mathbb{R}^n$ (use [Rud87, 1.34]). The same argument applies to $x_j f_p(x) \to x_j f(x)$ in L_t. Since $\partial_j f$ is in $L_s(\mathbb{R}^n)$ it also holds that $k(x/p)(\partial_j f) \to \partial_j f$ in $L_s(\mathbb{R}^n)$. Also $\partial_j k_p = \frac{1}{p}\partial_j k$ so $f\partial_j k_p \to 0$ in $L_s(\mathbb{R}^n)$. Using Leibniz' rule (see [Rud91, (5),p.160]) and the triangle inequality then gives

$$\|\partial_j f_p - \partial_j f\|_s \le \|(\partial_j f)k_p - \partial_j f\|_s + \|f\partial_j k_p\|_s \to 0 \qquad \text{for } p \to \infty,$$

since $\partial_j k_p$ is zero when $|x| \le p$.

2. Now (3.1) has been shown for f_p, but the sequence f_p is not necessarily in $C_c^\infty(\mathbb{R}^n)$. As shown in [Rud91, 6.30(b)] the convolution $h_q * f_p$ is in $C^\infty(\mathbb{R}^n)$. The f_p's have compact support and their convolution with h_q thus has compact support so. We now have a sequence $g_{p,q} = h_q * f_p$ that according to [Rud91, 6.32] approximates f_p for $q \to \infty$. Since $\partial_j f_p$ is in $L_s(\mathbb{R}^n)$, [Rud91, 6.32] and [Rud91, 7.19(a)] give

$$\partial_j g_{p,q} = h_q * \partial_j f_p \to \partial_j f_p \qquad \text{in } L_s \text{ for } q \to \infty.$$

The convolution $h_q * f_p$ has compact support independent of q because

$$\operatorname{supp}(h_q * f_p) \subseteq \operatorname{supp}(h_q) + \operatorname{supp}(f_p) \subseteq \operatorname{supp}(h) + \operatorname{supp}(f_p).$$

The sets $\operatorname{supp}(h)$ and $\operatorname{supp}(f_p)$ are compact and therefore the sum is compact. On this set multiplication with x_j is a bounded operator on $L_r(\mathbb{R}^n)$ and is thus continuous. This gives the last convergence $x_j g_{p,q} \to x_j f_p$ in L_t for $q \to \infty$.

3. For each $k \in \mathbb{N}$ I can choose p and q such that

$$\begin{aligned} \|g_{p,q} - f_p\|_r + \|\partial_j g_{p,q} - \partial_j f_p\|_s + \|x_j g_{p,q} - x_j f_p\|_t &\le \frac{1}{2k} \\ \|f_p - f\|_r + \|\partial_j f_p - \partial_j f\|_s + \|x_j f_p - x_j f\|_t &\le \frac{1}{2k}. \end{aligned}$$

Setting $g_k = g_{p,q}$ and using Schwarz's inequality [Rud87, Thm. 3.5] proves that the sequence g_k in $C_c^\infty(\mathbb{R}^n)$ satisfies (3.1). □

Theorem 3.2 (Heisenberg). Let $f \in L_2(\mathbb{R}^n)$, then for all $j \in \{1, 2, \ldots, n\}$

$$\int_{\mathbb{R}} x_j^2 |f(x)|^2 dx \int_{\mathbb{R}} y_j^2 |\hat{f}(y)|^2 dy \ge \frac{1}{4}\Big(\int_{\mathbb{R}} |f(x)|^2\Big)^2. \tag{3.2}$$

Proof. The inequality is obvious if $f(x) = 0$ almost everywhere, so now I will assume that f is non-zero in $L_2(\mathbb{R}^n)$. Then neither $x_j f(x)$ nor $y_j \hat{f}(y)$ is non-zero and thus if either of them has infinite L_2-norm the inequality is obvious. So from now on I will assume that $x_j f(x)$ and $y_j \hat{f}(y)$ are in $L_2(\mathbb{R}^n)$.

I will start by showing the inequality for $f \in C_c^\infty(\mathbb{R}^n)$ and then use Lemma 3.1.

For $f \in C_c^\infty(\mathbb{R}^n)$ we have

$$\begin{aligned}
\int_{\mathbb{R}} x_j^2 |f(x)|^2 dx \int_{\mathbb{R}} y_j^2 |\hat{f}(y)|^2 dy &= \int_{\mathbb{R}} x_j^2 |f(x)|^2 dx \int_{\mathbb{R}} |(\partial_j f)(x)|^2 dx \\
&\geq \Big(\int_{\mathbb{R}} |x_j (\partial_j f)(x) \overline{f(x)}| dx \Big)^2 \\
&\geq \Big(\int_{\mathbb{R}} x_j \operatorname{Re}((\partial_j f)(x) \overline{f(x)}) dx \Big)^2 \\
&= \frac{1}{4} \Big(\int_{\mathbb{R}} x_j ((\partial_j f)(x) \overline{f(x)} + \overline{(\partial_j f)(x)} f(x)) dx \Big)^2 \\
&= \frac{1}{4} \Big(\int_{\mathbb{R}} x_j (\partial_j |f|^2)(x) dx \Big)^2 \\
&= \frac{1}{4} \Big([x_j |f(x)|^2]_{-\infty}^{\infty} - \int_{\mathbb{R}} |f(x)|^2 dx \Big)^2 \\
&= \frac{1}{4} \Big(\int_{\mathbb{R}} |f(x)|^2 \Big)^2
\end{aligned}$$

The first equality holds for any $f \in L_2(\mathbb{R}^n)$ with $\partial_j f \in L_2(\mathbb{R}^n)$ according to definition 1.6 (which is actually [Rud91, Thm. 7.15(b)]). Using the convergences of Lemma 3.1 with $r = s = t = 2$ on the following expression from the above calculations

$$\int_{\mathbb{R}} x_j^2 |f(x)|^2 dx \int_{\mathbb{R}} |(\partial_j f)(x)|^2 dx \geq \frac{1}{4} \Big(\int_{\mathbb{R}} |f(x)|^2 \Big)^2$$

proves Heisenberg's inequality in the general case (using [Rud91, Thm 7.15(b)]). $\square$

Remark 3.3 (The case of equality). If $n = 1$ and if both $f(x)$, $xf(x)$ and $y\hat{f}(y)$ are in $L_2(\mathbb{R})$ then equality only holds for Gaussian functions. First notice that if $xf(x)$ is in $L_2(\mathbb{R})$ then $\sqrt{|x|}f(x)$ is in $L_2(\mathbb{R})$ since g defined by

$$g(x) = \begin{cases} |f(x)| & |x| \leq 1 \\ x|f(x)| & |x| > 1 \end{cases}$$

is in $L_2(\mathbb{R})$ and $|x||f(x)|^2 \leq |g(x)|^2$ for all $x \in \mathbb{R}$. From this it follows that $(1 + |x|)f(x)$ is in $L_2(\mathbb{R})$. Since $(1 + |x|)^{-1}$ is also in $L_2(\mathbb{R})$ using Hölder's inequality tells us that $f(x)$ is in $L_1(\mathbb{R})$. The same holds for $\hat{f}$ and from [Rud91, 7.5] (used on the inverse Fourier transform) it follows that f is equivalent (in $L_2(\mathbb{R})$) to a continuous function.

Now assume that equality holds in (3.2). Then equality holds for the application of Schwarz's inequality [Rud87, Thm. 3.5] for the expression

$$\int_{\mathbb{R}} x_j^2 |f(x)|^2 dx \int_{\mathbb{R}} |(\partial_j f)(x)|^2 dx \geq \Big(\int_{\mathbb{R}} |x_j (\partial_j f)(x) \overline{f(x)}| dx \Big)^2,$$

which holds for any $f \in L_2(\mathbb{R})$. But this is only true if $kx\overline{f(x)} = \partial f(x)$ for some complex k. Then also $\partial f(x)$ is continuous. That ∂f is actually f' in the classical sense follows from

$$\int_0^x \partial f(t)dt = \lim_{n\to\infty} \int_0^x g_n'(t)dt = \lim_{n\to\infty} [g_n(x)]_0^x$$
$$= \lim_{n\to\infty} (g_n(x) - g_n(0)) = f(x) - f(0).$$

Here the sequence g_n is chosen as in Lemma 3.1 and the last equality is true because f is continuous. We now have an ordinary differential equation that looks like

$$f'(x) = kx\overline{f(x)}$$

which, solved by separation of the variables, shows that f is a Gaussian function.

Corollary 3.4. Let $f \in L_2(\mathbb{R})$ and let $a, b \in \mathbb{R}$, then

$$\int_{\mathbb{R}} (x-a)^2 |f(x)|^2 dx \int_{\mathbb{R}} (y-b)^2 |\hat{f}(y)|^2 dy \geq \frac{1}{4}\Big(\int_{\mathbb{R}} |f(x)|^2\Big)^2.$$

If $xf(x)$ and $y\hat{f}(y)$ are in $L_2(\mathbb{R})$ equality holds only for functions of type

$$f(x) = ce^{ibt}e^{-dt^2}, \qquad c, d \in \mathbb{R}.$$

Proof. We can again assume that $xf(x)$ and $y\hat{f}(y)$ are in $L_2(\mathbb{R})$. Define

$$f_{a,b}(x) = e^{-ibx} f(x+a).$$

Then $f_{a,b}$ is in $L_2(\mathbb{R})$ and so are $xf_{a,b}(x)$ and $y\hat{f}_{a,b}(y)$. For $f_{a,b}$ this follows from

$$\int_{\mathbb{R}} (x-a)^2 |f(x)|^2 dx = \int_{\mathbb{R}} (x^2 + a^2 - 2ax)|f(x)|^2 dx$$

and since

$$g(x) = \begin{cases} |f(x)| & |x| \leq 1 \\ x^2|f(x)| & |x| > 1 \end{cases}$$

is integrable and $|xf(x)| \leq g(x)$ for all $x \in \mathbb{R}$. By a change of variable it holds that

$$\int_{\mathbb{R}} (x-a)^2 |f(x)|^2 dx \int_{\mathbb{R}} (y-b)^2 |\hat{f}(y)|^2 dy = \int_{\mathbb{R}} x^2 |f_{a,b}(x)|^2 dx \int_{\mathbb{R}} y^2 |\hat{f}_{a,b}(y)|^2 dy$$
$$\geq \frac{1}{4}\Big(\int_{\mathbb{R}} |f_{a,b}(x)|^2\Big)^2 \geq \frac{1}{4}\Big(\int_{\mathbb{R}} |f(x)|^2\Big)^2$$

It follows that $f_{a,b}$ is Gaussian if equality holds which means f is as stated. $\square$

Remark 3.5. This can be extended to

$$\int_{\mathbb{R}} |x|^2 |f(x)|^2 dx \int_{\mathbb{R}} |y|^2 |\hat{f}(y)|^2 dy \geq \frac{n^2}{4}(\int_{\mathbb{R}} |f(x)|^2)^2,$$

since $x^2 = x_1^2 + \cdots + x_n^2$ has n factors.

3.2 Variation of Heisenberg's Uncertainty Principle

I will now concentrate on functions in one dimension. The first generalisation of Heisenberg's uncertainty principle is

Theorem 3.6. If $1 \leq p \leq 2$ and $f \in L_2(\mathbb{R})$ is nonzero, then

$$\|f\|_2^2 \leq 2(2\pi)^{1/2-1/p}\|xf\|_p\|y\hat{f}\|_p.$$

In the proof ot the original version of Heisenberg's uncertainty principle, I approximated f with functions in C_c^∞. This required only basic knowledge about Fourier transforms. Here I will use the fact that the Fourier transform of tempered distributions is defined using Schwartz functions, therefore I need only show the result for $f \in \mathcal{S}$. In other words I use the tools developed in chapter 7 of [Rud91].

Proof. Since the Fourier transform of tempered distributions is defined by use of Schwartz functions, I need only concentrate on $f \in \mathcal{S}$.

The case $p = 2$ has been dealt with, so assume $1 \leq p < 2$. But then

$$\|f\|_2^2 = \int_{\mathbb{R}} x(|f'|^2)dx \leq 2\int_{\mathbb{R}} |x\bar{f}f'|dx$$

like in the proof of Heisenberg's inequality. Then apply Hölder's inequality to get $\|f\|_2^2 \leq 2\|xf\|_p\|f'\|_q$ where $1/p+1/q = 1$. Then Remark 2.4 following Hausdorff-Young's inequality gives $\|f'\|_q \leq (2\pi)^{1/q-1/2)}\|\hat{f'}\|_p$ and [Rud91, Thm. 7.4(c)] gives $\|\hat{f'}\|_p = \|y\hat{f}\|_p$. Putting it all together we then get

$$\|f\|_2^2 \leq 2(2\pi)^{1/2-1/p}\|xf\|_p\|y\hat{f}\|_p.$$

$\square$

Note that if I define the Fourier transform by $\int f(x)e^{2\pi ixy}$, I would get the "nicer" result $4\pi\|xf\|_2\|y\hat{f}\|_2 \geq \|f\|_2^2$ (see p. 214 in [FS97]).

Later in this section I will need the gamma-function, so I define it here. To learn more about the gamma-function, the main reference is [Art64]

Definition 3.7. The gamma-function is defined for $x > 0$ by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Lemma 3.8.

$$x\Gamma(x) = \int_0^\infty \exp(-y^{1/x})dy$$

Proof.

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$$

Substituting by $y = t^x$ gives $dy/dt = xt^{x-1}$ and $y = t^{1/x}$ so

$$\Gamma(x) = \int_0^\infty \frac{\exp(-y^{1/x})}{x}dy.$$

□

I will now demonstrate one further variant of Heisenberg's uncertianty principle. It is due to Hirschman [Hir57, p. 153]. To do this it is useful to define the entropy

Definition 3.9. For a function $\phi \in L_1$ with $\|\phi\|_1 = 1$ define the entropy of ϕ to be

$$E(\phi) = \int_{\mathbb{R}} \phi(x)\log\phi(x)dx.$$

When $\phi(x) = 0$, set $\phi(x)\log\phi(x) = 0$ (by continuous extension of $t \mapsto t\log t$).

Remark 3.10. $E(\phi)$ can have values in $[-\infty, \infty]$. It can also be undefined if for example

$$\int_{\mathbb{R}} \phi(x)\log^+\phi(x)dx = \infty$$

and

$$\int_{\mathbb{R}} \phi(x)\log^-\phi(x)dx = -\infty$$

where $\log^+ x = \max\{0, \log x\}$ and $\log^- x = \min\{0, \log x\}$.

First a few lemmas

Lemma 3.11. If ϕ has $\|\phi\|_1 = 1$ and $M_a(\phi) = (\int |x|^a\phi(x)dx)^{1/a}$ is finite for some $a > 0$ then there is a constant $k_a > 0$ such that

$$-E(\phi) \leq \log(k_a) + \log M_a(\phi). \tag{3.3}$$

The smallest constant k_a for which the above always holds is given by $2e^{\frac{1}{a}}a^{\frac{1}{a}-1}\Gamma(\frac{1}{a})$.

Proof. If we dilate ϕ by c, that is replace $\phi(cx)$ by $\phi_c(x) = c\phi(cx)$ then both sides of the inequality (3.3) decrease by the amount $\log c$. This is seen by

$$\begin{aligned}\int_{\mathbb{R}} c\phi(cx)\log(c\phi(cx))dx &= \log c\int_{\mathbb{R}} c\phi(cx)dx + \int_{\mathbb{R}} c\phi(cx)\log(\phi(cx))dx \\ &= \log c + \int_{\mathbb{R}} \phi(x)\log(\phi(x))dx,\end{aligned}$$

and

$$M_a^a(\phi_c) = \int_{\mathbb{R}} |x|^a c\phi(cx)dx = \int_{\mathbb{R}} |y/c|^a \phi(y)dy = M_a^a(\phi)/c^a.$$

Also $\int \phi_c(x)dx = \int \phi(x)dx = 1$. So now I can assume that $M_a(\phi) = 1$ otherwise I can simply dilate by $M_a(\phi)$.

For a given $c > 0$ let k be given by $\int_{\mathbb{R}} \exp(-c|x|^a)dx$. Then $d\gamma(x) = k^{-1}\exp(-c|x|^a)dx$ is a probability measure (which means it has integral 1 over the real numbers). Defining $\psi(x) = k\exp(c|x|^a)\phi(x)$ we get $\int_{\mathbb{R}} \psi d\gamma = \int_{\mathbb{R}} \phi(x)dx = 1$. According to Remark 1.17 the mapping $\phi : t \mapsto t\log t$ is convex on $]0,\infty[$ so Jensen's inequality, Theorem 1.15, gives

$$\begin{aligned}
0 &= \Big(\int_{\mathbb{R}} \psi d\gamma\Big) \log\Big(\int_{\mathbb{R}} \psi d\gamma\Big) \leq \int_{\mathbb{R}} \psi \log \psi d\gamma \\
&= \int_{\mathbb{R}} k\exp(c|x|^a)\phi(x)\log(k\exp(c|x|^a)\phi(x))k^{-1}\exp(-c|x|^a)dx \\
&= \int_{\mathbb{R}} \phi(x)(\log k + c|x|^a + \log\phi(x))dx
\end{aligned}$$

Splitting the last integral and using $\int_{\mathbb{R}} \phi(x)dx = 1$ and $M_a(\phi) = 1$ gives

$$0 \leq \log(e^c k) + \int_{\mathbb{R}} \phi(x)\log\phi(x)dx.$$

Now I want to minimize the function $e^c k$ with respect to c. Since the integrand is even substitution with $y = c^{1/a}x$ gives

$$k = \int_{\mathbb{R}} \exp(-cx^a)dx = 2c^{-1/a}\int_0^\infty \exp(-y^a)dy = 2c^{-1/a}\Gamma(1/a)/a.$$

To minimize $e^c k$ is the same a minimising $e^c c^{-1/a}$. Since this is a growing function the minimum can be found in the following manner. Differentiation with respect to c gives $e^c c^{-1/a}(\frac{-1}{ac} + 1)$ which is zero if $c = 1/a$. Inserting this c gives the contant

$$k_a = 2e^{1/a}(1/a)^{-1/a}\Gamma(1/a)/a = 2e^{\frac{1}{a}}a^{\frac{1}{a}-1}\Gamma(1/a).$$

The dilation argument at the beginning of the proof gives the desired result. □

Lemma 3.12. Let f be a measurable function. Assume ψ is a convex function on $\mathbb{C}$ and $\psi \circ f$ is in L_1. If h_n is a sequence of non-negative functions in L_1 such that $\int_{\mathbb{R}} h_n(x)dx = 1$ and $h_n * f \to f$ pointwise then $\psi(h_n * f(x))$ is in L_1 and

$$\lim_{n\to\infty} \int_{\mathbb{R}} \psi(h_n * f(x))dx = \int_{\mathbb{R}} \psi(f(x))dx.$$

Proof. For a given x the following holds by Jensen's Inequality

$$\begin{aligned}\psi(h_n * f(x)) = \psi(f * h_n(x)) &= \psi\Big(\int_{\mathbb{R}} f(x-y)h_n(y)dy\Big) \\ &\leq \int_{\mathbb{R}} \psi(f(x-y))h_n(y)dy \\ &= \int_{\mathbb{R}} h_n(x-y)\psi(f(y))dy = h_n * (\psi \circ f)(x)\end{aligned}$$

It follows that $h_n * (\psi \circ f)$ is in L_1 since $\psi \circ f$ is and therefore $\psi(h_n * f(x))$ is also in L_1. Since $h_n * f \to f$ pointwise the Dominated convergence theorem [Rud87, Thm. 1.34] gives the desired result. □

Now follows the result originally proved by Hirschman in [Hir57].

Lemma 3.13. For $f \in L_2$ with $\|f\|_2 = 1$

$$E(|f|^2) + E(|\hat{f}|^2) \leq 0 \tag{3.4}$$

whenever the left hand side has meaning.

Proof. If one of $E(|f|^2)$ or $E(|\hat{f}|^2)$ is $-\infty$ then either the inequality is clear (the other is $-\infty$ or finite) or the left hand side is undefined (the other is ∞). So I need not consider these cases.

1. Start by assuming that $f \in L_1 \cap L_2$. Then f will be in L_p for $1 < p < 2$ since

$$g(x) = \begin{cases} |f(x)|^2 & \text{if } |f(x)| \geq 1 \\ |f(x)| & \text{if } |f(x)| < 1 \end{cases}$$

is an integrable majorant. With $1/p + 1/q = 1$ Hausdorff-Young's inequality tells us that $\hat{f}$ is in L_q. So we can define the functions

$$A(p) = \int_{\mathbb{R}} |f(x)|^p dx \qquad \text{and} \qquad B(q) = \int_{\mathbb{R}} |\hat{f}(y)|^q dy.$$

Since $y - 1 \geq \log y$ for all $y \in \mathbb{R}$ it holds that $(x^a - 1)/a \log x$ for $a > 0$ (insert $y = x^a$). But this shows that for each x and $p > 2$

$$\frac{|f(x)|^2 - |f(x)|^p}{2-p} \leq |f(x)|^2 \log |f(x)|.$$

For $p \to 2-$ it holds that $\frac{|f(x)|^2 - |f(x)|^p}{2-p} \to |f(x)|^2 \log |f(x)|$. The left hand side is integrable for each p and therefore $A'(2-)$ is always defined with values in $]-\infty, \infty]$ and it is given by

$$A'(2-) = \int_{\mathbb{R}} |f(x)|^2 \log |f(x)| dx.$$

Similarly for each x

$$\frac{|\hat{f}(x)|^2 - |\hat{f}(x)|^q}{2-q} \geq |f(x)|^2 \log |f(x)|$$

and for $q \to 2+$ it holds that $\frac{|f(x)|^2 - |f(x)|^q}{2-q} \to |f(x)|^2 \log |f(x)|$. The left hand side is integrable for each q and therefore $B'(2+)$ is always defined with values in $[-\infty, \infty[$ and it is given by

$$B'(2+) = \int_{\mathbb{R}} |\hat{f}(x)|^2 \log |\hat{f}(x)| dx.$$

Now define

$$C(p) = \log \|\hat{f}\|_q - \log \|f\|_p = q^{-1} \log B(q) - p^{-1} \log A(p).$$

By the remark at the beginning of the proof I can assume that $B'(2+) \neq -\infty$ so $C(p)$ has a derivative from the left in 2 (which can have values in $]-\infty, \infty]$). The derivate of q with respect to p is $q' = -(p-1)^2$, and so

$$C'(p) = -q^{-2} q' \log B(q) + q^{-1} q' B'(q)/B(q) + p^{-2} \log A(p) - p^{-1} A'(p)/A(p).$$

And for $p \to 2-$ we use $A(2) = B(2) = 1$ to get

$$C'(2-) = -(B'(2+) + A'(2-))/2. \tag{3.5}$$

which is defined since $B'(2+) \neq -\infty$. By Hausdorff-Young's inequality $C(p) \leq 0$ for $1 < p \leq 2$ and by Parsevals equality (Theorem 1.3) $C(2) = 0$ and therefore $C'(2-) \geq 2$. This together with (3.5) shows (3.4) inn this case.

2. Let us now assume that $f \in L_2$ and that $E(|f|^2)$ and $E(|\hat{f}|^2)$ are defined, and then approximate f by functions in $L_1 \cap L_2$. Let

$$\psi_n(x) = \begin{cases} 1 - |x|/n & \text{for } |x| \leq n \\ 0 & \text{for } |x| > n \end{cases}$$

then $\hat{\psi}_n = \sin^2(\pi n y)/(\pi^2 n y^2)$. Then the Fourier transform of $\psi_n f$ is $\hat{\psi}_n * \hat{f}$ and $\|\hat{\psi}_n * \hat{f}\|_2 = \|\psi_n f\|_2 \leq 1$. Pointwise $\psi_n f \to f$ and by the Lebesgue Dominated Convergence Theorem [Rud87, 1.34] it holds that $\psi_n f \to f$ in L_2. By the same theorem we see that

$$E(|\psi_n f|^2) \to E(|f|^2). \tag{3.6}$$

Now I want to show that the same convergence holds for $\hat{\psi}_n * \hat{f}$. The functions

$$\phi_1(t) = |z|^2 \log^+ |z| \qquad \text{and} \qquad \phi_2(t) = |z|^2(-\log^- |z| + 3/2)$$

are convex on $\mathbb{C}$ (see Remark 1.20). Since $\hat{\psi}_n \geq 0$ and $\int_{\mathbb{R}} \hat{\psi}_n = 1$ for $i = 1, 2$ we get

$$\lim_{n \to \infty} \int_{\mathbb{R}} \phi_i(\hat{\psi}_n * \hat{f}(x)) dx = \int_{\mathbb{R}} \phi_i(\hat{f}(x)) dx$$

by Lemma 3.12. It also holds that

$$E(|\hat{f}|^2) = 2\int_{\mathbb{R}} \phi_1(\hat{f}(x))dx - 2\int_{\mathbb{R}} \phi_2(\hat{f}(x))dx + 3\|\hat{f}\|_2^2$$

which then gives

$$\lim_{n\to\infty} E(|\hat{\psi}_n * \hat{f}|^2) = E(|\hat{f}|^2). \tag{3.7}$$

Notice that these limits also hold in the case where $E(|f|^2)$ and $E(\hat{f}|^2)$ are ∞ or $-\infty$.

If $\|f\|_2 = \|\hat{f}\|_2 \neq 1$ then by replacing f by $f/\|f\|_2$ the inequality (3.4) becomes

$$E(|f|^2) + E(|\hat{f}|^2) \leq 2\|f\|_2^2 \log\|f\|_2 \tag{3.8}$$

Since $\psi_n f$ is in $L_1 \cap L_2$ (it has compact support) and $\|\psi_n f\|_2 \leq 1$ the following holds by (3.8)

$$E(|\psi_n f|^2) + E(\hat{\psi}_n * \hat{f}) \leq 0.$$

And the limits (3.6) and (3.7) will complete the proof. □

Theorem 3.14. If $a, b > 0$ then there exists a constant $K > 0$ such that

$$\||x|^a f\|_2^{1/a} \||y|^b \hat{f}\|_2^{1/b} \geq K\|f\|_2^{(a+b)/ab}.$$

We can choose K to be $\frac{1}{(e^2 k_{2a}k_{2b})^2}$, where $k_t = 2e^{\frac{1}{t}} t^{\frac{1}{t}-1}\Gamma(\frac{1}{t})$.

Proof. Lemma 3.11 gives

$$\int_{\mathbb{R}} |f(x)|^2 \log|f(x)| \geq -\log(ek_{2a}) - \frac{1}{2a}\log\||x|^a f\|_2$$
$$\int_{\mathbb{R}} |f(x)|^2 \log|f(x)| \geq -\log(ek_{2b}) - \frac{1}{2b}\log\||x|^b \hat{f}\|_2.$$

The Lemma 3.13 then gives

$$-\log(e^2 k_{2a}k_{2b}) - \frac{1}{2a}\log\||x|^a f\|_2 - \frac{1}{2b}\log\||x|^b \hat{f}\|_2 \leq 0$$

Which is equivalent to

$$-2\log(e^2 k_{2a}k_{2b}) \leq \frac{1}{a}\log\||x|^a f\|_2 + \frac{1}{b}\log\||x|^a f\|_2$$

Taking the exponential function on both sides gives

$$\frac{1}{(e^2 k_{2a}k_{2b})^2} \leq \||x|^a f\|_2^{1/a}\||x|^b \hat{f}\|_2^{1/b}.$$

If $\|f\|_2 \neq 1$ then by substituting f by $f/\|f\|_2$ we arrive at the desired result. □

With $a = b$ in the above we get

Corollary 3.15. If $a > 0$ then, with $K = \frac{1}{(ek_{2a})^{4a}}$, where $k_t = 2e^{\frac{1}{t}} t^{\frac{1}{t}-1} \Gamma(\frac{1}{t})$, it holds that

$$\||x|^a f\|_2 \||y|^a \hat{f}\|_2 \geq K \|f\|_2^2,$$

when the left hand side exists.

After this we might want to extend to general inequalities of the form

$$\||x|^a f\|_p^{\gamma} \||y|^b \hat{f}\|_q^{1-\gamma} \geq K \|f\|_2$$

for $0 < \gamma < 1$, $a, b \in \mathbb{R}_+$ and $p, q \in [1, \infty]$.

The following lemma is Lemma 3.3 in [FS97]. I will skip the proof.

Lemma 3.16. Suppose that $p, q \in [1, \infty]$, $a, b \in]0, \infty[$, $\gamma \in]0, 1[$ and that

$$\gamma(a + 1/p - 1/2) = (1 - \gamma)(b + 1/q - 1/2) \tag{3.9}$$

holds. Then for $f \in L_2$ the following two inequalities are equivalent

$$\|f\|_2 \leq K \||x|^a f\|_p^{\gamma} \||y|^b \hat{f}\|_q^{1-\gamma} \tag{3.10}$$

$$\|f\|_2 \leq K(\gamma \||x|^a f\|_p + (1 - \gamma) \||y|^b \hat{f}\|_q) \tag{3.11}$$

(3.9) is a necessary condition for (3.10) to hold.

We can of course choose the constant K to be the same in both inequalities, but there might be a smaller constant for which one of the inequalities still hold. I have not investigated this.

For inequalities of the form (3.11) Cowling and Price [CP84] have proved the most general result

Theorem 3.17. Suppose $p, q \in [1, \infty]$ and $a, b \geq 0$. There exists a constant $K > 0$ such that

$$\||x|^a f\|_p + \||y|^b \hat{f}\|_q \geq K \|f\|_2$$

for all tempered distributions f with the property that f and $\hat{f}$ are locally integrable functions if and only if $a > 1/2 - 1/p$ and $b > 1/2 - 1/q$ or $(p, a) = (2, 0)$ or $(q, b) = (2, 0)$.

This is Theorem 5.1 in [CP84]. The proof is very long and I will therefore skip it.

Chapter 4

Qualitative Uncertainty Principles

Qualitative uncertainty principles are not inequalities, but are theorems that tell us how a function (and its Fourier transform) behave under certain circumstances. One such example can be the fact that a function and its Fourier transform cannot both have compact support. I will present a classical qualitative uncertainty principle called Hardy's theorem. I will then prove a generalisation of this principle and also look at questions about the support of a function and its Fourier transform. Finally I will investigate functions that are close to zero on measurable sets.

4.1 Hardy's Theorem

In this section I will explore an old theorem by Hardy about functions in $L_1(\mathbb{R})$.

The proof is inspired by [DM72, p. 156-158], but with a difference in normalization of the Fourier transform.

First I will need the following theorem

Theorem 4.1 (Phragmén-Lindelöf). Given $a \in]\frac{1}{2}, \infty[$ for $\phi < \frac{\pi}{2a}$ define

$$D = \{z \in \mathbb{C} | -\phi \leq arg(z) \leq \phi\}.$$

Assume the function f is holomorphic on the inner D° of D and continuous on the boundary ∂D of D and there exists constants b and C such that

$$|f(x)| \leq Ce^{b|z|^a}, \qquad z \in D.$$

If there exists a constant M such that $|f(z)| \leq M$ for $z \in \partial D$ then $|f(z)| \leq M$ for all $z \in D$.

Proof. Since $2a\phi < \pi$ we can choose a constants $s > a$ such that $2s\phi < \pi$. For any $A > 0$ define

$$h(z) = \frac{f(z)}{\exp(Az^s)}, \qquad z \in D.$$

The function $z \mapsto Az^s$ is holomorphic on D° and continuous in 0 and therefore the same holds for $h(z)$.

For $z = re^{i\phi}$ the following holds

$$|h(z)| = \frac{|f(re^{i\phi})|}{|\exp(Ar^s e^{is\phi})|} \leq \frac{M}{\exp(Ar^s cos(s\phi))} \leq M$$

because $s\phi > \pi/2$ ensures that $Ar^s cos(s\phi) \geq 0$. The same will hold for $z = e^{-s\phi}$. So $|h(z)| \leq M$ on ∂D.

Now define

$$D_r = \{z \in D||z| \leq r\}, \qquad r \geq 0.$$

I will show that there is a R_0 such that $|h(z)| \leq M$ on ∂D_r when $r \geq R_0$. It has already been shown for $z = re^{i\phi}$ and $z = re^{-i\phi}$ for any r. I now only need to show that there is a R_0 such that $|h(z)| \leq M$ for $z = re^{-i\theta}$ with $r \geq R_0$ and $-\phi < \theta < \phi$.

Let $m = \inf_{-\phi<\theta<\phi} cos(s\theta)$ then $m > 0$ since $s\theta < \pi/2$ and thus

$$|\exp(Ar^s e^{is\theta})| = \exp(Ar^s cos(s\theta)) \geq \exp(Ar^s m).$$

Given that $s > a$ for $|z| = r$ it then holds that

$$|h(z)| = \frac{|f(z)|}{|\exp(Az^s)|} \leq \frac{C\exp(br^a)}{\exp(Ar^s m)} = C\exp(br^a - Ar^s m) \to 0, \qquad \text{for } r \to \infty.$$

This shows there is a R_0 such that $|h(z)| \leq M$ for $z \in \partial D_r$ when $r \geq R_0$. The maximum modulus theorem [ST83, Thm. 10.14] then gives $|h(z)| \leq M$ for all $z \in D_r$ when $r \geq R_0$. It follows that $|h(z)| \leq M$ for all $z \in D$, which is the same as

$$|f(z)| \leq M\exp(Az^s), \qquad z \in D \text{ for every } A > 0. \tag{4.1}$$

Letting $A \to 0$ gives the desired result $|f(z)| \leq M$ for all $z \in D$. $\square$

Remark 4.2. If f is bounded on the lines of angles α and β with $0 < \beta - \alpha < \pi$ then we can define $f_1(z) = f(ze^{-i(\beta+\alpha)/2})$. Then f_1 is bounded and thus the same holds for f.

I will also need the following lemma

Lemma 4.3. Given a function f assume there exists constants $a > 0$ and $C \geq 0$ such that $|f(x)| \leq Ce^{-ax^2}$, then $\hat{f}$ defined by

$$\hat{f}(z) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-ixz}dx, \qquad \text{for all } z \in \mathbb{C} \tag{4.2}$$

is well defined and entire.

Proof. First I will show that $\hat{f}$ is well defined. Then that it is continous. Then I will integrate $\hat{f}$ over a loop, showing that the result is 0 for any loop. Applying Morera's Theorem [ST83, Thm. 10.4] finishes the proof.

1. The integral (4.2) is well defined for any $z \in \mathbb{C}$ since

$$\int_{\mathbb{R}} |f(x)||e^{-ixz}|dx = \int_{\mathbb{R}} |f(x)|e^{\mathrm{Im}(z)x}dx \leq C\int_{\mathbb{R}} e^{\mathrm{Im}(z)x-ax^2}dx < +\infty.$$

2. I will now show that $\hat{f}$ is continous. Given $z \in \mathbb{C}$ assume that $(z_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$ is a sequence such that $z_n \to z$. Then

$$\begin{aligned} |\hat{f}(z_n) - \hat{f}(z)| &= \frac{1}{\sqrt{2\pi}}\Big|\int_{\mathbb{R}} f(x)e^{-iz_nx}dx - \int_{\mathbb{R}} f(y)e^{-iz_ny}dy\Big| \\ &= \frac{1}{\sqrt{2\pi}}\Big|\int_{\mathbb{R}} f(x)(e^{-iz_nx} - e^{-iz_nx})dx\Big| \\ &\leq \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} |f(x)||e^{-iz_nx} - e^{-iz_nx}|dx \end{aligned} \tag{4.3}$$

Define $s = \sup\{\mathrm{Im}(z), \mathrm{Im}(z_1), \mathrm{Im}(z_2), \dots\}$ then for every $x \in \mathbb{R}$ and $n \in \mathbb{N}$ it holds that

$$|e^{-izx} - e^{-iz_nx}| \leq 2e^{s|x|}.$$

From that follows

$$|f(x)||e^{-izx} - e^{-iz_nx}| \leq 2Ce^{-Bx^2}e^{s|x|} = 2Ce^{s|x|-Bx^2}.$$

The function $2Ce^{s|x|-Bx^2}$ is integrable and the Dominated Convergence Theorem ([Rud87, 1.34]) with (4.3) then gives

$$|\hat{f}(z) - \hat{f}(z_n)| \leq \int_{\mathbb{R}} |f(x)||e^{-izx} - e^{-iz_nx}|dx \to 0, \qquad \text{for } n \to 0$$

since $z \mapsto e^{-izx}$ is continous for all $x \in \mathbb{R}$i and $z_n \to z$.

3. Given a loop $\gamma : [0,1] \mapsto \mathbb{C}$ I will now show that $\int_\gamma \hat{f}(z)dz = 0$. First notice that since e^{-izx} is entire for all $x \in \mathbb{R}$ it follows that $\int_\gamma e^{-izx}dz = 0$. Integrating f along γ gives

$$\begin{aligned} \int_\gamma \hat{f}(z)dz &= \int_\gamma \int_{\mathbb{R}} (f(x)e^{-izx}dxdz \\ &= \int_0^1 \int_{\mathbb{R}} f(x)e^{-i\gamma(s)x}dx\gamma'(s)ds \\ &= \int_{\mathbb{R}} f(x)\int_0^1 e^{-i\gamma(s)x}\gamma'(s)dsdx \\ &= \int_{\mathbb{R}} f(x)\int_\gamma e^{-izx}dzdx \\ &= 0 \end{aligned}$$

Applying Morera's Theorem [ST83, Thm. 10.4] tells us that $\hat{f}$ is an entire function. $\square$

Theorem 4.4 (Hardy). If $a, b, C, D \in \mathbb{R}_+$ exist such that

$$|f(x)| \leq Ce^{-ax^2}, \qquad \text{for all } x \in \mathbb{R} \tag{4.4}$$

$$|\hat{f}(y)| \leq De^{-by^2}, \qquad \text{for all } y \in \mathbb{R} \tag{4.5}$$

then there are the following three possibilities:

(i) if $ab > 1/4$ then $f = 0$

(ii) if $ab = 1/4$ then $f = ke^{-ax^2}$ for some $k \in \mathbb{R}$

(iii) if $ab < 1/4$ then there are infinitely many functions that satisfy (4.4) and (4.5).

Proof. Here is a small overview of the proof.

(i) is proved using (ii)

(ii) if f is even then using $h(y) = \hat{f}(\sqrt{y})$ show that $e^{by}h(y)$ is constant. If f uneven do the same for the even function $y^{-1}\hat{f}(y)$, but then the constant must be 0. Finally split f into even and uneven parts.

(iii) is proved using the Hermite functions

First I will prove that only the product ab is important and not the actual values of a and b. Assume that (4.4) and (4.5) hold. For any given $k \neq 0$ define $f_1(x) = f(kx)$. Then

$$\hat{f}_1(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(kx)e^{-ixy}dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} k^{-1}f(x)e^{-ik^{-1}xy}dx = k^{-1}\hat{f}(k^{-1}y)$$

For f_1 the inequalities (4.4) and (4.5) will take the form

$$|f_1(x)| = |f(kx)| \leq C'e^{-ak^2x^2} = C'e^{-a'x^2}$$
$$|\hat{f}_1(x)| = k^{-1}|f(k^{-1}x)| \leq D'e^{-bk^{-2}y^2} = C'e^{-b'y^2}$$

with $a' = ak^2$, $b' = bk^{-2}$ and $D' = k^{-1}D$. It is seen that $a'b' = ab$ so the values of a and b are not important as long as the product is the same.

(ii) I will now go on to prove the second case where $ab = 1/4$. I have just proved that I can assume $a = 1/4\pi$ and $b = \pi$.

If f is even then so is $\hat{f}$. $\hat{f}$ can thus be written as a powerseries of even powers $\hat{f}(y) = \sum_{j\in\mathbb{N}} c_j z^{2j}$, so I can define h by

$$h(y) = \hat{f}(\sqrt{y}) = \sum_{i\in\mathbb{N}} c_j z^j.$$

For $\hat{f}$ it holds

$$\begin{aligned}
|\hat{f}(y)| &= \frac{1}{\sqrt{2\pi}}\Big|\int_{\mathbb{R}} f(x)e^{-ixy}dx\Big| \\
&\leq \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} |f(x)|e^{\operatorname{Im}(y)x}dx \\
&\leq \frac{C}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-x^2/4\pi}e^{\operatorname{Im}(y)x}dx \\
&= \frac{C}{\sqrt{2\pi}}\int_{\mathbb{R}} \varphi(x)e^{-i(i\operatorname{Im}(y))x}dx \\
&= C\hat{\varphi}(i\operatorname{Im}(y)) \\
&= C'e^{\pi\operatorname{Im}^2(y)},
\end{aligned}$$

where $\varphi(x) = e^{-x^2/4\pi}$ is the Gaussian function with Fourier transform $\hat{\varphi}(y) = \frac{1}{\sqrt{4\pi}}e^{-\pi y^2}$ and $C' = C/\sqrt{4\pi}$. Setting $y = Re^{it}$ the following then holds for h

$$|h(y)| \leq C'e^{\pi\operatorname{Im}^2(\sqrt{y})} = C'e^{\pi R\sin^2(t/2)}. \tag{4.6}$$

If y is on the positive real line then by (4.5)

$$|h(y)| \leq De^{-\pi R}. \tag{4.7}$$

Choose $M = \max(C', D)$ so that both (4.6) and (4.7) hold with the constants chosen to be M.

Let $0 < \delta < \pi$ and define the plane D_δ and the function g_δ by

$$D_\delta = \{z = Re^{it} | 0 \leq t \leq \delta, R \geq 0\} \qquad \text{and} \qquad g_\delta(y) = \exp\Big(\frac{i\pi y e^{i\delta/2}}{\sin(\delta/2)}\Big).$$

Inserting $z = Re^{it}$ gives

$$g_\delta(Re^{it}) = \exp\left(\frac{-\pi R\sin(t-\delta/2)}{\sin(\delta/2)}\right). \tag{4.8}$$

For $t = 0$ and $t = \delta$ we then get

$$g_\delta(R) = e^{\pi R} \qquad \text{and} \qquad g_\delta(Re^{i\delta}) = e^{-\pi R},$$

from which it follows that

$$|g_\delta(R)h(R)| \leq M \qquad \text{and} \qquad |g_\delta(Re^{i\delta})h(Re^{i\delta})| \leq M.$$

So now the function $g_\delta h$ has been shown to be limited on the boundary of D_δ and since the function is analytic it is limited on the whole of D_δ by Phragmén-Lindelöf (a version of Theorem 4.1 rotated by $\delta/2$). Since

$$\frac{\sin(t-\delta/2)}{\sin(\delta/2)} \to -\cos t \qquad \text{for } \delta \to \pi,$$

(4.8) and $g_\delta h \leq M$ shows that

$$|h(y)| \leq M e^{-\pi R \cos t} \qquad \text{for } 0 \leq t \leq \pi.$$

A similar result can be established for $-\pi \leq t \leq 0$. For all $z = Re^{it} \in \mathbb{C}$ it therefore holds that

$$|e^{\pi z} h(z)| = |e^{\pi R(\cos t + i \sin t)} h(z)| = |e^{\pi R \cos t} h(z)| \leq M.$$

Liuville's theorem [ST83, Thm. 10.6] then states that $e^{\pi z} h(z)$ is constant for all $z \in \mathbb{C}$. This tells us that $\hat{f}$ is of the form $\hat{f}(y) = K e^{-\pi y^2}$. Using the inverse Fourier transform gives us that f is of the form $f(x) = K' e^{-x^2/4\pi}$.

If f is uneven then $\hat{f}$ is uneven. Since $\hat{f}$ is analytical it can be written as a power series $\hat{f}(y) = \sum_{n \in \mathbb{N}} c_n y^{2n+1}$. This means that $y^{-1}\hat{f}$ will be even and analytical. Applying the same arguments as for the even function we get

$$\hat{f}(y) = yKe^{-\pi y^2}.$$

But then $\hat{f}$ can only obey (4.5) if $K = 0$.

The last thing there is to do is to split f into even and uneven parts. This is done by setting $f_{even}(x) = (f(x) + f(-x))/2$ and $f_{uneven}(x) = (f(x) - f(-x))/2$. Then $f = f_{even} + f_{uneven}$. Since f obeys the inequalities (4.4) and (4.5) so do f_{even} and f_{uneven}. Which finishes the proof for $ab = 1/4$.

(i) Now assume that $ab > 1/4$. I can then assume that $a > 1/4\pi$ and $b > \pi$. Then f and $\hat{f}$ will surely satisfy

$$|f(x)| \leq Ce^{-ax^2} \leq Ce^{-x^2/4\pi}$$
$$|\hat{f}(y)| \leq De^{-by^2} \leq De^{-\pi y^2}$$

and using (ii) we get that $f(x) = C'e^{-x^2/4\pi}$ for some C'. But that means $|C'|e^{-x^2/4\pi} \leq Ce^{-ax^2}$ for all $x \in \mathbb{R}$. But this can only be true if $C' = 0$ since $a > 1/4\pi$.

(iii) If $ab < 1/4$ we can assume that $a = b < 1/2$. By remarks 1.10 and 1.11 we get that there is an M such that

$$|\hat{H}_n(x)| = \sqrt{2}|H_n(x)| \leq Me^{-ax^2}$$

since any polynomial is $O(e^{cx^2})$ for $c > 0$ (we can set $c = 1/2 - a > 0$). But then H_n is an infinite family of functions that satisfy (4.4) and (4.5). $\square$

4.2 Variation of Hardy's Theorem

In this section I will present a variation of Hardy's theorem. It extends the original result to the case where f goes to 0 faster than a polynomial times a Gaussian function. It is found in [Kaw72, p.278ff].

Definition 4.5. Assume that $f(z)$ with $z = re^{i\theta}$ is holomorphic in $\alpha \leq \theta \leq \beta$, $r \geq r_0$ for some $r_0 \geq 0$ and that

$$\limsup_{r\to\infty}(\log|f(re^{i\theta})|/r)$$

exists. Define

$$h_f(\theta) = \limsup_{r\to\infty}(\log|f(re^{i\theta}|/r)$$

Theorem 4.6. Let $\beta - \alpha < \pi$ and let $h_f(\alpha) \leq h_1$ and $h_f(\beta) \leq h_2$. Let $H(\theta)$ be the unique function of the form $a\cos\theta + b\sin\theta$ such that $H(\alpha) = h_1$ and $H(\beta) = h_2$, then $h_f(\theta) \leq H(\theta)$.

Proof. Let $\delta > 0$ and let $g_\delta(\theta) = a_\delta\cos\theta + b_\delta\sin\theta$ be the unique function of the form $a\cos\theta - b\sin\theta$ that has values $h_1 + \delta$ and $h_2 + \delta$ for α and β respectively. Let $f_\delta(z) = f(z)e^{-(a_\delta+ib_\delta)z}$. Then for $z = re^{i\theta}$ it holds that $|f_\delta(z)| = |f(z)|e^{-g_\delta(\theta)r)}$ is bounded on the lines $se^{i\alpha}$ and $se^{i\beta}$ with $s \geq 0$. Using theorem 4.1 we then get that $|f_\delta| \leq M$ for some positive constant M on $[\alpha, \beta]$. But then $|f(z)| \leq Me^{g_\delta(\theta)r)}$ and thus

$$h_f(\theta) = \limsup_{n\to\infty}(\log|f(r)|/r) \leq \limsup_{n\to\infty}(\log M/r) + g_\delta(\theta) = g_\delta(\theta).$$

When $\delta \to 0$ the function g_δ tends to H pointwise, so this ends the proof. □

Remark 4.7. The function $H(\theta)$

$$H(\theta) = \frac{h_1 \sin\rho(\beta-\theta) + h_2\sin\rho(\theta-\alpha)}{\sin\rho(\beta-\alpha)}$$

satisfies the theorem.

Lemma 4.8. Let $f(z)$ be analytic in $\{z \in \mathbb{C}|\operatorname{Im}(z) \geq 0\}$. Suppose that there are possitive constants A and B such that

- $|f(z)| \leq Ae^{B|z|}$ for $\operatorname{Im}(z) \geq 0$,
- $|f(r)| \leq 1$ for $r \in \mathbb{R}$, and
- $\limsup_{r\to\infty}(\log|f(ir)|/r) \leq 0$

Then $|f(z)| \leq$ for all z with $\operatorname{Im}(z) \geq 0$.

The proof is found in [PS72, Prob. 325]

Proof. The last condition tells us that for any $c > 0$ $\limsup_{r\to\infty}\log|f(ir)e^{-cr}|/r \leq -c$ so $|f(ir)e^{-cr}| \to 0$ for $r \to \infty$. Therefore $|f(ir)e^{-cr}|$ has a maximum F_c at some $r_0 \geq 0$. Also $|f(r)e^{-cr}| \leq 1$ on the real axis so using theorem 4.1 gives $|f(z)e^{-cz}| \leq \max\{1, F_c\}$ on $\operatorname{Im}(z) \geq 0$. Now I need only prove that $F_c \leq 1$. If the maximum of $|f(ir)e^{-cr}|$ is at $r_0 = 0$ then $F_c \leq 1$ because $|f(r)| \leq 1$ on the real axis. If $r_0 > 0$ then the maximum of $|f(ir)e^{-cr}|$ is obtained in an inner point ir_0 of $\operatorname{Im}(z) \geq 0$. But this conflicts with the maximum modulus theorem [ST83, Thm. 10.14]. Therefore $F_c \leq 1$. Since this holds for any $c > 0$ it follows that $|f(z)| \leq 1$. The proof is complete. □

Remark 4.9. This can be rotated by for example $\pi/2$ to give the same result for $\mathrm{Re}(z) \geq 0$ with f bounded on the imaginary axis.

Lemma 4.10. Let $f(z)$ be an entire function. If there exists constants C_1 and C_2 and $a > 0$ such that

$$f(z) \leq C_1|z|^n e^{a|z|} \qquad \text{for } z \in \mathbb{C}$$

and

$$f(x) \leq C_2|x|^m e^{-a|x|} \qquad \text{for } x \in \mathbb{R}$$

with $n, m \in \mathbb{N}$ then for all $z \in \mathbb{C}$

$$f(z) = e^{-az}P(z),$$

where $P(z)$ is a polynomial of degree less than m.

The proof has been inspired by the proof of [BG95, 1.3.19]

Proof. The assumptions mean that for some constant $C > 0$

$$\log|f(z)| \leq \log(Cr^n e^{ar}) = \log C + n\log r + ar.$$

So $h_f(\theta) \leq a$. Also

$$h_f(0) = \limsup_{r\to\infty} r^{-1}|f(r)| \leq -a.$$

For any $\phi \in]0, \pi[$ set

$$H(\theta) = -a\cos\theta + a\tan(\phi/2)\sin\theta.$$

Then $h_f(0) \leq -a = H(0)$ and $h(\pi - \phi) \leq a = H(\pi - \phi)$. Using theorem 4.6 we then get that $h(\theta) \leq H(\theta)$ for $0 \leq \theta \leq \pi - \phi$. Then $h(\pi/2) \leq H(\pi/2) = a\tan(\phi/2)$ and in the limit $\phi \to 0$ we get

$$h_f(\pi/2) \leq 0. \tag{4.9}$$

Now let $F(z) = e^{az}f(z)$ and $G(z) = F(z)/(z + i)^m$. Then $G(z)$ is analytic in $\mathrm{Im}(z) \geq 0$ and for some $c > 0$ $|G(z)| \leq e^{c|z|}$ since the exponential function grows faster than any polynomial of finite degree n. On the real axis $|G(x)|$ is bounded since $F(x) = O(|x|^m)$ by the assumption that $f(x) = O(|x|^m e^{a|x|})$. On the upper imaginary half-axis (4.9) gives

$$\begin{aligned}\limsup_{y\to\infty} \log|G(iy)|/y &= \limsup_{y\to\infty} \log(|F(iy)|/|iy + i|^m)/y \\ &= \limsup_{y\to\infty}(\log|f(iy)| - m\log(y + 1))/y \\ &= \limsup_{y\to\infty}(\log|f(iy)|)/y \leq 0.\end{aligned}$$

By Lemma 4.8 we then get that $G(z)$ is less than a certain constant. But then $F(z) = O(|z|^m)$ on $\mathrm{Im}(z) \geq 0$ and by a similar argument we can show the same for $\mathrm{Im}(z) \leq 0$. Therefore $F(z) = O(|z|^m)$ on all of $\mathbb{C}$. $F(z)/z^m$ is thus a constant which shows that $F(z)$ is a polynomial $P(z)$ of degree less than or equal to m (since F is also entire). But then $f(z) = P(z)e^{-az}$ and the proof is complete. □

Theorem 4.11. If $a, b \in \mathbb{R}_+$ with $ab \geq 1/4$ exist such that

$$|f(x)| = O(|x|^m e^{-ax^2}), \qquad \text{for } |x| \to \infty$$
$$|\hat{f}(y)| = O(|y|^m e^{-by^2}), \qquad \text{for } |y| \to \infty$$

for some $m \in \mathbb{N}$, then

$$\hat{f}(y) = P(y)e^{-by^2}, \qquad \text{for all } y \in \mathbb{R}.$$

Proof. The Fourier transform

$$\hat{f}(z) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-ixz}dx$$

is well defined and entire according to Lemma 4.3 (the exponential function decays faster than any polynomial, so the requirements of the lemma are satisfied). For some $u > 0$ and $C > 0$ it holds that $|f(x)| \leq C|x|^m e^{-ax^2}$ for $|x| \geq u$. Therefore

$$\begin{aligned} |\sqrt{2\pi}\hat{f}(z)| &\leq \int_{-u}^{u} |f(x)e^{-ixz}|dx + \int_{-\infty}^{-u} |f(x)e^{-ixz}|dx + \int_{u}^{\infty} |f(x)e^{-ixz}|dx \\ &\leq \int_{-u}^{u} |f(x)e^{-ixz}|dx + 2C\int_{u}^{\infty} x^m e^{-ax^2}e^{|xz|}dx \\ &\leq e^{|zu|}\int_{-u}^{u} |f(x)|dx + 2C\int_{0}^{\infty} x^m e^{-ax^2}e^{x|z|}dx. \end{aligned}$$

The last inequality is due to

$$|e^{-ixz}| = e^{x\,\mathrm{Im}(z)} \leq e^{|x\,\mathrm{Im}(z)|} \leq e^{|xz|} \leq e^{|uz|},$$

when $x \in [-u, u]$. The last integral exists and can be evaluated by

$$\int_0^\infty x^m e^{-ax^2}e^{x|z|}dx = \int_0^\infty \frac{\partial^m}{\partial |z|^m} e^{-ax^2+|z|x}dx = \frac{\partial^m}{\partial |z|^m}\int_0^\infty e^{-ax^2+|z|x}dx$$

and using $-a(x - |z|/2a)^2 = -ax^2 + x|z| - |z|^2/4a$ I get

$$= \frac{\partial^m}{\partial |z|^m} e^{|z|^2/4a}\int_0^\infty e^{-a(x-|z|/2a)^2}dx \leq \frac{\partial^m}{\partial |z|^m} e^{|z|^2/4a}\int_{-\infty}^\infty e^{-a(x-|z|/2a)^2}dx$$

and then a change of variable by $y = x - |z|/2a$ gives

$$= \frac{\partial^m}{\partial |z|^m} e^{|z|^2/4a}\int_{-\infty}^\infty e^{-ay^2}dy$$

The integral is finite by (1.1) and $\frac{\partial^m}{\partial |z|^m} e^{|z|^2/4a}$ is equal to a polynomial of degree m in $|z|$ times $e^{|z|^2/4a}$. So it follows that

$$|\hat{f}(z)| = O(|z|^m e^{|z|^2/4a}). \tag{4.10}$$

for large $|z|$.

Let $F_1(z) = \hat{f}(z) + \hat{f}(-z)$ then F_1 is an even and entire function and $|F_1(z)| = O(|z|^m e^{|z|^2/4a})$. Since F_1 is entire and even, its Taylor series contains only even powers of z and thus G defined by $G(z^2) = F_1(z)$ is entire. Assume that m is even and set $n = m/2$ then by (4.10)

$$G(w) = O(|w|^n e^{|w|/4a}) = O(|w|^n e^{b|w|})$$

where the last evaluation is due to the fact that $ab \geq 1/4$. Also by assumption on $\hat{f}(x)$ it follows that $F_1(x) = O(|x|^m e^{-bx^2})$ for $x \in \mathbb{R}$ and which tells us that

$$G(x) = O(|x|^n e^{-b|x|}).$$

Lemma 4.10 it follows that $G(w) = e^{-bw} Q(w)$ where Q is a polynomial of degree at most n. But then

$$F_1(z) = e^{-bz^2} P_1(z), \tag{4.11}$$

where P_1 is a polynomial with degree less than or equal to m. If m is odd then use $m+1$ in the above arguments to obtain (4.11) with $degr(P_1) \leq m+1$.

Now define the entire and odd function $F_2(z) = \hat{f}(z) - \hat{f}(-z)$ and let $\phi(z) = zF_2(z)$. Then $\phi(z)$ is even and entire and $\phi(z) = O(|z|^{m+1} e^{|z|^2/4a})$ and $\phi(x) = O(|x|^{m+1} e^{-b|x|^2})$ for real x. We can then use the argument we used for F_1 (and m) for ϕ and $m+1$ to get that

$$\phi(z) = e^{-bz^2} R(z)$$

where R is a polynomial of degree at most $m+2$. But since $\phi(0) = 0$ then R has no constant term. So

$$F_2(z) = e^{-bz^2} P_2(z) \tag{4.12}$$

where $P_2(z) = R(z)/z$ has degree less than or equal to $m+1$.

(4.11) and (4.12) together give

$$\hat{f}(z) = e^{-bz^2} P(z)$$

where $P(z)$ has degree at most $m+1$. But since $\hat{f}(x) = O(|x|^m e^{-bx^2})$ the polynomial P must have degree at most m. □

With $a = b = 1/2$ in the theorem above we get

Corollary 4.12. If $f(x)$ and $\hat{f}(x)$ are both $O(|x|^m e^{-x^2/2})$ for $|x| \to \infty$ then they are both of the form $P(x)e^{-x^2/2}$ where $P(x)$ is a polynomial of degree less than m.

Proof. According to Remark 1.11 the polynomial $P(y)$ from theorem 4.11 can be written as a linear combination of Hermite polynomials. Therefore $\hat{f}$ can be written as a linear combination of Hermite functions. By Remark 1.10 the Hermite functions are eigenfunctions of the Fourier transform (and so also the inverse transform). So f can be written as a linear combination of Hermite functions. □

4.3 Functions With Support on Measurable Sets

In this section I will state a theorem concerning the sizes of the sets where f and $\hat{f}$ are not zero. If both of these sets have finite measure then f is the zero function. This is an extension of the result that f and $\hat{f}$ cannot both have compact support.

In the following let $|A|$ be the normed Lebesgue measure of a measurable set $A \subseteq \mathbb{R}^n$.

Definition 4.13. For a function $f : \mathbb{R}^n \to \mathbb{R}$ consider the support of f

$$S(f) = \{x \in \mathbb{R}^n | f(x) \neq 0\}.$$

This is not the support $\mathrm{supp}(f)$ which is normally defined as $\overline{S(f)}$, so I have used a different notation to separate the two.

The following lemma is stated in [FS97]. The proof was left as an exercise.

Lemma 4.14. Let $\mathbb{T} = [0, 1]$. If $f \in L_1(\mathbb{R}^n)$ the series $\phi(x) = \sum_{k\in\mathbb{Z}^n} f(x + k)$ converges in $L_1(\mathbb{T}^n)$ and the Fourier series of ϕ is $\sum_{k\in\mathbb{Z}^n} \hat{f}(k)e^{2\pi ikx}$.

Proof. Assume $f \in L_1(\mathbb{R})$. First look at

$$\begin{aligned}\int_0^1 \Big|\sum_{|k|\leq N} f(x+k)\Big| dx &\leq \int_0^1 \sum_{|k|\leq N} |f(x+k)| dx = \sum_{|k|\leq N} \int_0^1 |f(x+k)| dx \\ &= \sum_{|k|\leq N} \int_k^{k+1} |f(x)| dx = \int_{-N}^{N+1} |f(x)| dx \leq \int_{\mathbb{R}} |f(x)| dx < \infty\end{aligned}$$

which shows that $\phi \in L_1(\mathbb{T})$.

The Fourier coefficients are

$$\begin{aligned}\sqrt{2\pi} c_n(\phi) &= \int_0^1 \phi(x) e^{-in2\pi x} dx \\ &= \int_0^1 \Big(\sum_{k\in\mathbb{Z}^n} f(x+k)\Big) e^{-in2\pi x} dx \\ &= \int_0^1 \Big(\sum_{k\in\mathbb{Z}^n} f(x+k) e^{-in2\pi x}\Big) dx \\ &= \int_0^1 \Big(\sum_{k\in\mathbb{Z}^n} f(x+k) e^{-in2\pi(x+k)}\Big) dx \qquad \text{since } e^{-in2\pi k} = 1 \\ &= \int f(x) e^{in2\pi y} dy = \sqrt{2\pi}\hat{f}(n).\end{aligned}$$

The Fourier series is therefore $\sum_{k\in\mathbb{Z}^n} \hat{f}(k)e^{2\pi ikx}$ as stated. $\square$

I will also need the following result

Lemma 4.15. Let $f : [0,1]^n \to \mathbb{R}$ be a positive measurable function and let dx be the Lebesgue-measure that gives $|[0,1]^n| = 1$. If $\int_{[0,1]^n} f(x)dx < \infty$ then there is a set $E \subseteq [0,1]^n$ with $|E| = 1$ such that $f(a) < \infty$ for $a \in E$.

Proof. Let $E = \{x \in [0,1]^n | f(x) < \infty\}$. I will show that $|E| = 1$. Let $F = [0,1]^n \setminus E$. If $|F| > 0$ then we cannot have $\int_F f(x)dx = \infty$ which cannot be true since $\int_{[0,1]^n} f(x)dx < \infty$. □

This next lemma about the zeros of an analytic function will also be needed to prove the main theorem.

Lemma 4.16. Assume f is an analytic function on $\mathbb{C}^n$ and let $Z(f) = \{x \in \mathbb{C}^n | f(x) = 0\}$. Then one of the following instances occur

(i) $Z(f)$ has measure zero in $\mathbb{C}^n$

(ii) $Z(f) = \mathbb{C}^n$

Proof. Let μ_n be the Lebesgue-measure on $\mathbb{C}^n$. For $n = 1$ the lemma follows from [Rud87, Thm. 10.18] using the fact that if $Z(f)$ has no limit point in $\mathbb{C}$ then the zeroes are isolated (and therefore $Z(f)$ is discrete or $Z(f) = \emptyset$). I will then use induction.

Let f be a function of n complex variables $x_1, \dots, x_n$ and assume that the lemma holds for a function of $n-1$ variables.

For $x_n \in \mathbb{C}$ define $f_n(x_1, \dots x_{n-1}) = f(x_1, \dots, x_n)$ and for $(x_1, \dots, x_{n-1}) \in \mathbb{C}^{n-1}$ define $f^n(x_n) = f(x_1, \dots, x_n)$. Then $Z(f) = Z(f_n) \otimes Z(f^n)$. It is then known that $\mu_n(Z(f)) = \mu_{n-1}(Z(f_n))\mu_1(Z(f^n))$.

If either $\mu_{n-1}(Z(f_n)) = 0$ or $\mu_1(Z(f^n)) = 0$ then $\mu_n(Z(f)) = 0$. If not then $Z(f_n) = \mathbb{C}^{n-1}$ and $Z(f^n) = \mathbb{C}$ such that $Z(f) = \mathbb{C}^n$. □

Now follows the main result stated and proved as Theorem 7.2 in [FS97].

Theorem 4.17. If $f \in L_1(\mathbb{R}^n)$ and $|S(f)||S(\hat{f})| < \infty$, then $f = 0$.

Proof. Since $|S(f)|$ and $|S(\hat{f})|$ are finite by dilating f by a constant k we make the support of $\hat{f}$ k times bigger. $|S(\hat{f})|$ is still finite and therefore we can assume that $|S(f)| < 1$ without loss of generality.

The following two calculations hold:

$$\int_{[0,1]^n} \sum_{k \in \mathbb{Z}^n} 1_{S(\hat{f})}(y+k)dy = \int_{\mathbb{R}^n} 1_{S(\hat{f})}(y)dy = |S(\hat{f})| < \infty,$$
$$\int_{[0,1]^n} \sum_{k \in \mathbb{Z}^n} 1_{S(f)}(x+k)dx = \int_{\mathbb{R}^n} 1_{S(f)}(x)dx = |S(f)| < 1.$$

According to Lemma 4.15 the first inequality tells us there exists an E with $|E| = 1$ such that $\sum_{k \in \mathbb{Z}^n} 1_{S(\hat{f})}(a+k) < \infty$ for $a \in E$. Therefore $\hat{f}(a+k) \neq 0$ only for finitely many k when $a \in E$ (otherwise we would have an infite sum of 1).

Let $F = \{x \in [0,1]^n | \sum_{k\in\mathbb{Z}^n} 1_{S(f)}(x+k) = 0\}$. Then $|F| > 0$. If not the integral above would be greater than or equal to 1 since $\sum_{k\in\mathbb{Z}^n} 1_{S(f)}(x+k) \geq 1$ if for some k $f(x+k) \neq 0$ for $x \notin F$. Therefore $f(x+k) = 0$ for all k if $x \in F$.

Given an $a \in E$ define

$$\phi_a(x) = \sum_{k\in\mathbb{Z}^n} f(x+k)e^{-2\pi i a(x+k)}.$$

By Lemma 4.14 $\phi_a \in L_1(\mathbb{T}^n)$ and the Fourier series of ϕ_a is $\phi_a(k) = \sum \hat{f}(a+k)e^{2\pi ikx}$. Since $a \in E$ the Fourier series converges uniformly and is equal to ϕ_a. So ϕ_a is a trigonometric polynomial and thus ϕ_a has analytic extension to $\mathbb{C}^n$. But then either $\phi_a = 0$ everywhere or $\phi_a \neq 0$ almost everywhere (Lemma 4.16). But $|\phi_a(x)| \leq \sum_{k\in\mathbb{Z}^n} |f(x+k)| = 0$ for $x \in F$ and F has measure greater than zero, so $\phi_a = 0$ everywhere for all $a \in E$. This means that its Fourier series is zero so $\hat{f}(a+k) = 0$ for all $a \in E$ and $k \in \mathbb{Z}^n$. But this tells us that $\hat{f} = 0$ almost everywhere. Since $\hat{f}$ is continuous $\hat{f} = 0$ and then $f = 0$. □

4.4 Almost Vanishing Functions

In the previous section I concentrated on the support of f and $\hat{f}$. Now I will investigate the case where f and $\hat{f}$ are close to zero outside measurable sets. The entire section is taken from section 3 in [DS89].

To formalise the notion of "close to zero" I start with a definition

Definition 4.18. We say that a function $f \in L_2(\mathbb{R})$ is ϵ-concentrated on a measurable set T if there is a measurable function $g(t)$ vanishing outside T such that $\|f - g\|_2 \leq \epsilon$.

This is only interesting if ϵ is small. Let T and W be measurable sets. As tools in the proof I will use two operators P_T and P_W. Define the operator

$$P_T f(x) = \begin{cases} f(x), & t \in T \\ 0, & t \notin T. \end{cases}$$

If f is ϵ_T-concentrated on T (g being the vanishing function) then

$$\|f - P_T f\|_2 = \int_{\mathbb{R}\setminus T} |f(x)|^2 dx \leq \int_{\mathbb{R}} |f(x) - g(x)|^2 dx \leq \epsilon_T$$

and therefore f is ϵ_T-concentrated on T if and only if $\|f - P_T f\|_2 \leq \epsilon_T$.

Since $\hat{f}$ is in $L_2(\mathbb{R})$ it is essentially bounded, and if $|W|$ has finite measure then the operator

$$P_W f(t) = \frac{1}{\sqrt{2\pi}} \int_W \hat{f}(y)e^{iyt} dy$$

is well-defined, since $\hat{f}1_W$ is in L_1. $P_W f$ is the inverse Fourier transform of $\hat{f}1_W$, so $\widehat{P_W f} = \hat{f}1_W$ almost everywhere, which tells us that it vanishes outside W. From this it follows as for P_T that $\hat{f}$ is ϵ_W-concentrated if and only if

$$\|\hat{f} - \widehat{P_W f}\|_2 = \|f - P_W f\|_2 \leq \epsilon_W.$$

The norm of an operator P is defined as

$$\|P\| = \sup_{\|f\|_2=1} \|Pf\|_2.$$

This is equivalent to the definition

$$\|P\| = \sup_{f\in L_2} \frac{\|Pf\|_2}{\|f\|_2}$$

from the bottom of page 909 in [DS89]. Using $f = \frac{1}{|W|}1_W$ we get $\|P_W f\|_2 = \|f\|_2 = 1$, so $\|P_W\| = 1$, since for other f, $\|P_W f\|_2 \leq \|f\|_2$.

For a bounded function $q(s,t)$ in $L_2(\mathbb{R}^2)$ that vanishes when t is outside a set of finite measure, define the operator Q by

$$(Qf)(t) = \int_{\mathbb{R}} q(s,t)f(s)$$

and its Hilbert-Schmidt norm

$$\|Q\|_{HS} = \|q\|_2 = \Big(\int_{\mathbb{R}}\int_{\mathbb{R}} |q(s,t)|^2 ds dt\Big)^{1/2}.$$

Then the following two lemmas hold

Lemma 4.19.

$$\|Q\|_{HS} \geq \|Q\|.$$

Proof. Let $f \in L_2$ have unit norm. q is bounded and vanished outside a measurable set so $q(s,t)f(s) \in L_1(\mathbb{R})$. Since $q, f \in L_2$ the following calculations hold

$$\begin{aligned}
\|Qf\|_2^2 &= \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} q(s,t)f(s)ds\Big|^2 dt \leq \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |q(s,t)f(s)|ds\Big)^2 dt \\
&\leq \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |q(s,t)|^2 ds\Big)\Big(\int_{\mathbb{R}} |f(s_1)|^2\Big) ds_1 dt = \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |q(s,t)|^2 ds\Big) dt \\
&= \|Q\|_{HS}^2.
\end{aligned}$$

The second inequality is Hölder's inequality. This proves the lemma. $\square$

Note that

$$P_W P_T f(t) = \int_{\mathbb{R}} q(s,t) f(s) ds$$

when

$$q(s,t) = \begin{cases} \frac{1}{\sqrt{2\pi}} \int_W e^{iw(s-t)} dw & t \in T \\ 0 & t \notin T. \end{cases}$$

This is true because

$$\begin{aligned} P_W P_T f(s) &= \frac{1}{2\pi} \int_W e^{iws} \widehat{P_T f}(w) dw \\ &= \frac{1}{2\pi} \int_W e^{iws} \int_T f(t) e^{-iwt} dt dw \\ &= \frac{1}{2\pi} \int_T f(t) \int_W e^{iw(s-t)} dt dw \end{aligned}$$

where Fubinis theorem [Rud87, Thm. 8.8]is used.

Lemma 4.20. If T and W are sets of finite measure, then

$$\|P_W P_T\|_{HS} = \sqrt{|T||W|}.$$

Proof. For $t \in T$ let $g_t(s) = q(s,t) = \frac{1}{\sqrt{2\pi}} \int_W e^{iw(s-t)} dw$. Then the inversion formula shows that $\hat{g}_t(w) = 1_W e^{-iwt}$. By Parsevals identity it then follows

$$\int_{\mathbb{R}} |q(s,t)|^2 ds = \int_{\mathbb{R}} |g_t(s)|^2 ds = \int_{\mathbb{R}} |\hat{g}_t(w)|^2 dw = |W|$$

And integrating over $t \in T$ gives

$$\|P_W P_T\|_{HS} = \sqrt{|T||W|}.$$

□

Theorem 4.21. Let T and W be measurable sets and suppose that f and $\hat{f}$ are of unit norm. Assume that $\epsilon_T + \epsilon_W < 1$, that f is ϵ_T-concentrated on T and $\hat{f}$ is ϵ_W-concentrated on W. Then

$$|W||T| \geq (1 - \epsilon_T - \epsilon_W)^2.$$

Proof. Since $\|f\|_2 = \|\hat{f}\|_2 = 1$ the measures of T and W must both be non-zero, since ϵ_T and ϵ_W are less than 1. If not $\|f - g\|_2 = \|f\|_2 = 1 > \epsilon_T$ and likewise for $\hat{f}$. So if at least one of $|T|, |W|$ is infinity the inequality is clear. It is therefore enough to consider the case where both T and W have finite positive measures. In this case the operators P_T and P_W are bounded.

Since $\|P_W\| = 1$ it follows

$$\begin{aligned}\|f - P_W P_T f\|_2 &\leq \|f - P_W f\|_2 + \|P_W f - P_W P_T f\|_2 \\ &\leq \epsilon_W + \|P_W\| \|f - P_T f\|_2 \\ &\leq \epsilon_W + \epsilon_T.\end{aligned}$$

The triangle inequality gives $\|g\|_2 \geq \|f\|_2 - \|f - g\|_2$ and thus

$$\|P_W P_T f\|_2 \geq \|f\|_2 - \|f - P_W P_T f\|_2 \geq 1 - \epsilon_T - \epsilon_W.$$

It then follows that $\|P_W P_T\| \geq 1 - \epsilon_T - \epsilon_W$. The Lemmas 4.19 and 4.20 then show that $\sqrt{|T||W|} \geq \|P_W P_T\| \geq 1 - \epsilon_T - \epsilon_W$. $\square$

Chapter 5

Localised Uncertainty Principles

The classical uncertainty principles tell us that if the variance of $\hat{f}$ is small then the variance of f cannot be small too. But assume that f is a function such that $|\hat{f}|$ has two (or more) points of concentration (see figure 5.1 and the discussion in chapter 8 section 8.1), then the variance of $|\hat{f}|$ is big, but the classical uncertainty principles do not give much information about f. Investigating this problem leads to uncertainty principles that can be called "local uncertainty principles".

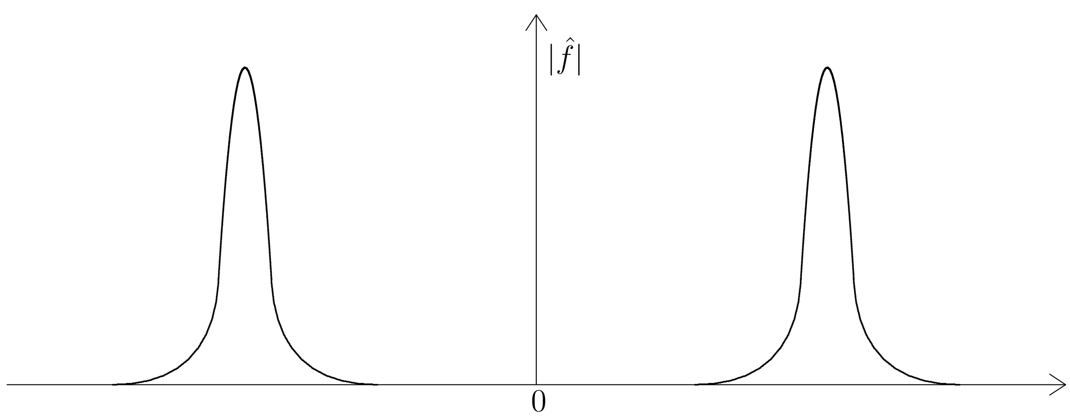

Figure 5.1: A function concentrated in several points.

5.1 An Uncertainty Principle in One Dimension

In this section I will present a result first proved by Strichartz [Str89]. It tells us that if the sum of $|\hat{f}|^2$ in countably many evenly distributed points is small, then the variance of f has a lower bound.

The following lemma is a corollary to the well known Fubini's Theorem [Rud87, Thm. 8.8]

Lemma 5.1. If f is an integrable function and $y > 0$ then we can swap the order of integration in the following sense

$$\int_0^y \int_0^x f(t,x)dtdx = \int_0^y \int_t^y f(t,x)dxdt$$

Proof. Define

$$F(t,x) = \begin{cases} f(t,x) & t \leq x \\ 0 & t > x \text{ and } t \leq y \end{cases}$$

then the following calculations hold according to Fubini's Theorem [Rud87, Thm. 8.8]

$$\begin{aligned} \int_0^y \int_0^x f(t,x)dtdx &= \int_0^y \int_0^y F(t,x)dtdx \\ &= \int_0^y \int_0^y F(t,x)dxdt \\ &= \int_0^y \int_t^y f(t,x)dtdx. \end{aligned}$$

□

Soon we will need the following special instance of the above Lemma:

$$\int_0^y x \int_0^x f(t)dtdx = \int_0^y f(t) \int_t^y xdxdt = \frac{1}{2} \int_0^y f(t)(y^2 - t^2)dt. \tag{5.1}$$

And now the main theorem proved in [Str89, p. 98f]

Theorem 5.2. Assume that f is a normed function in $L_2(\mathbb{R})$ and that $xf(x) \in L_2(\mathbb{R})$. Let a_j be an increasing sequence of numbers such that $a_j \to \infty$ for $j \to \infty$ and $a_j \to -\infty$ for $j \to -\infty$ with $a_{j+1} - a_j \leq b$. If

$$\Big(\sum_{j\in\mathbb{Z}} |\hat{f}(a_j)|^2\Big)^{1/2} \leq \frac{1-\epsilon}{\sqrt{b}} \tag{5.2}$$

then

$$\int_{\mathbb{R}} x^2 |f(x)|^2 dx \geq \frac{8\epsilon^2}{b^2} \tag{5.3}$$

Proof. Write $\hat{f}'$ instead of $\partial \hat{f}$. Let m_j be the midpoint of the interval $[a_j, a_{j+1}]$. Then by the fundamental theorem of calculus

$$\hat{f}(y) = \hat{f}(a_j) + \int_{a_j}^y \hat{f}'(t)dt \qquad \text{for } y \in]m_{j-1}, m_j[. \tag{5.4}$$

This gives the following estimate

$$\Big(\int_{m_{j-1}}^{m_j} |\hat{f}(y)|^2 dy\Big)^{1/2} = \Big(\int_{m_{j-1}}^{m_j} \Big|\hat{f}(a_j) + \int_{a_j}^y \hat{f}'(t)dt\Big|^2 dy\Big)^{1/2}$$

Schwarz' inequality [Rud87, Thm. 3.5] then gives

$$\leq \Big(\int_{m_{j-1}}^{m_j} |\hat{f}(a_j)|^2 dy + \Big| \int_{a_j}^{y} \hat{f}'(t)dt \Big|^2 dy \Big)^{1/2}$$

Minkowski's inequality [Rud87, Thm. 3.5(2)] and the fact that $m_j - m_{j-1} \leq b$ then give

$$\leq \sqrt{b}|\hat{f}(a_j)| + \Big(\int_{m_{j-1}}^{m_j} \Big| \int_{a_j}^{y} \hat{f}'(t)dt \Big|^2 dy \Big)^{1/2} \qquad \text{Minkowski}$$
$$\leq \sqrt{b}|\hat{f}(a_j)| + \Big(\int_{m_{j-1}}^{m_j} \Big(\int_{a_j}^{y} |\hat{f}'(t)|dt \Big)^2 dy \Big)^{1/2}$$

and a final application of Hölder's inequality yields

$$\leq \sqrt{b}|\hat{f}(a_j)| + \Big(\int_{m_{j-1}}^{m_j} \Big(\Big(\int_{a_j}^{y} |\hat{f}'(t)|^2 dt \Big)^{1/2} \Big(\int_{a_j}^{y} 1 dt \Big)^{1/2} \Big)^2 dy \Big)^{1/2}$$
$$= \sqrt{b}|\hat{f}(a_j)| + \Big(\int_{m_{j-1}}^{m_j} \int_{a_j}^{y} |\hat{f}'(t)|^2 dt |y - a_j| dy \Big)^{1/2}$$

Now look at the last integral. By first splitting the integral into two intervals, then chaning variables and at last using (5.1) the integral can be rewritten as

$$\begin{aligned}
&\int_{m_{j-1}}^{m_j} \int_{a_j}^{y} |\hat{f}'(t)|^2 dt |y - a_j| dy \\
&\quad = \int_{m_{j-1}}^{a_j} \int_{a_j}^{y} |\hat{f}'(t)|^2 dt |y - a_j| dy + \int_{a_j}^{m_j} \int_{a_j}^{y} |\hat{f}'(t)|^2 dt |y - a_j| dy \\
&\quad = \int_{0}^{a_j - m_{j-1}} \int_{0}^{y} |\hat{f}'(-t + a_j)|^2 dt y dy + \int_{0}^{m_j - a_j} \int_{0}^{y} |\hat{f}'(t + a_j)|^2 dt y dy \\
&\quad = \int_{0}^{a_j - m_{j-1}} \frac{1}{2} |\hat{f}'(-t + a_j)|^2 ((a_j - m_{j-1}) - t^2) dt \\
&\qquad + \int_{0}^{m_j - a_j} \frac{1}{2} |\hat{f}'(t + a_j)|^2 ((m_j - a_j)^2 - t^2) dt
\end{aligned}$$

and then, since $a_j - m_{j-1}$ and $m_j - a_j$ are less than $b/2$, estimated by

$$\leq \frac{b^2}{8} \int_{0}^{a_j - m_{j-1}} |\hat{f}'(-t + a_j)|^2 dt + \frac{b^2}{8} \int_{0}^{m_j - a_j} |\hat{f}'(t + a_j)|^2 dt,$$

then changing variable again and joining the integral, we obtain

$$= \frac{b^2}{8} \int_{m_{j-1}}^{m_j} |\hat{f}'(t)|^2 dt.$$

For every interval $[m_{j-1}, m_j]$ we then have

$$\Big(\int_{m_{j-1}}^{m_j} |\hat{f}(y)|^2 dy\Big)^{1/2} \le \sqrt{b}|\hat{f}(a_j) + \frac{b}{\sqrt{8}}\Big(\int_{m_{j-1}}^{m_j} |\hat{f}(t)|^2 dt\Big)^{1/2} \tag{5.5}$$

Using (5.4) and (5.5) we then get

$$\begin{aligned}
\Big(\int_{\mathbb{R}} |\hat{f}(y)|^2 dy\Big)^{1/2} &= \Big(\sum_{j\in\mathbb{Z}} \int_{m_{j-1}}^{m_j} \Big|\hat{f}(a_j) + \int_{a_j}^{y} \hat{f}'(t)dt\Big|^2 dy\Big)^{1/2} \\
&\le \Big(\sum_{j\in\mathbb{Z}} \Big(\sqrt{b}|\hat{f}(a_j)| + \frac{b}{\sqrt{8}}\Big(\int_{m_{j-1}}^{m_j} |\hat{f}'(t)dt|^2\Big)^{1/2}\Big)^2 dy\Big)^{1/2} \\
&\le \Big(\Big(\Big(\sum_{j} (\sqrt{b}|f(a_j)|)^2\Big)^{1/2} + \Big(\sum_{j\in\mathbb{Z}} \frac{b^2}{8} \int_{m_{j-1}}^{m_j} |\hat{f}'(t)|^2 dt\Big)^{1/2}\Big)^2\Big)^{1/2} \\
&= \sqrt{b}\Big(\sum_{j\in\mathbb{Z}} |\hat{f}(a_j)|^2\Big)^{1/2} + \frac{b}{\sqrt{8}}\Big(\sum_{j\in\mathbb{Z}} \int_{m_{j-1}}^{m_j} |\hat{f}'(t)|^2 dt\Big)^{1/2} \\
&= \sqrt{b}\Big(\sum_{j\in\mathbb{Z}} |\hat{f}(a_j)|^2\Big)^{1/2} + \frac{b}{\sqrt{8}}\Big(\int_{\mathbb{R}} |\hat{f}'(t)|^2 dt\Big)^{1/2}.
\end{aligned} \tag{5.6}$$

Rearranging this and using (5.2) and Parsevals theorem gives us

$$\int |xf(x)|^2 dx = \int |\hat{f}'(t)|^2 dt \ge \frac{8\epsilon^2}{b^2}.$$

This is the statement of the theorem. $\square$

I will interpret this uncertainty principle in section 8.1.

5.2 An Uncertainty Principle in Several Dimensions

I will now extend the result of the previous section to functions on $\mathbb{R}^n$ while simplifying the distribution between the "points of observation" by stating

Theorem 5.3. Let $b > 0$ and E be the union of $E_k = \{x | x_n = 2kb\}, k \in \mathbb{Z}$. Let σ be the canonical meassure of the Lebesgue meassure on $\mathbb{R}^n$ to E. Then for every $0 \le c_1 \le 1/2$ there exists c such that

$$\int_E |\hat{f}|^2 d\sigma \le c_1/b$$

implies

$$\int |x|^2 |f(x)|^2 dx \ge c/b^2.$$

Specifically it holds for $c = 2(1 - (2c_1)^{1/2})^2$.

First a short note on the measure σ. Since the measure on $\mathbb{R}^n$ is a product measure the measure σ on an E_k corresponds to the Lebesgue measure on $\mathbb{R}^{n-1}$. Let $T_k = [(2k-1)b, (2k+1)b] \times E_k$ such that the T_k's overlap only on a set of measure zero. The for a function $f(x) = f(y, x_n)$ the integral with respect to x can be written as

$$\int_{T_k} f(x)dx = \int_{[}(2k-1)b, (2k+1)b] \int_{E_k} f(y, x_n) d\sigma(y) dx_n.$$

The sum of the integrals $\int_{E_k} f d\sigma$ is then $\int_E f d\sigma$.

The proof is taken from p. 100-101 in [Str89].

Proof. Let $x = (y, x_n)$ where $y \in \mathbb{R}^{n-1}$. For $k \in \mathbb{Z}$ let $T_k = \{x \in \mathbb{R}^n | (2k-1)b \leq x_n \leq (2k+1)b\}$. By the fundamental theorem of calculus

$$\begin{aligned}\hat{f}(x) = \hat{f}(y, x_n) &= \hat{f}(y, 2kb) + \int_{2kb}^{x_n} \partial_n \hat{f}(y, s) ds \\ &= \hat{f}(y, 2kb) + \int_{2kb}^{x_n} 1_n(x) \nabla \hat{f}(y, s) ds\end{aligned}$$

for $x \in T_k$, where $1_n(x) = (0, 0, \dots, 0, 1)$ and ∇f is the n-tuple $(\partial_1 f, \dots, \partial_n f)$. The Schwarz inequality then gives

$$|\hat{f}(y, x_n)| \leq |\hat{f}(y, 2kb)| + \Big(|x_n - 2kb| \int_{2kb}^{x_n} 1_n(x) |\nabla \hat{f}(y, s)|^2 ds \Big)^{1/2}.$$

Using the note before the proof, the following evaluation is equivalent to (5.6)

$$\begin{aligned}\Big(\int_{T_k} |\hat{f}|^2 dx \Big)^{1/2} &= \Big(\int_{(2k-1)b}^{(2k+1)b} \int |\hat{f}(y, x_n)|^2 dy dx_n \Big)^{1/2} \\ &\leq \Big(2b \int_{E_k} |\hat{f}|^2 d\sigma \Big)^{1/2} + \Big(\frac{b^2}{2} \int_{(2k-1)b}^{(2k+1)b} \int |\nabla \hat{f}(y, s)|^2 dy ds \Big)^{1/2} \\ &= \Big(2b \int_{E_k} |\hat{f}|^2 d\sigma \Big)^{1/2} + \Big(\frac{b^2}{2} \int_{T_k} \int |\nabla \hat{f}(x)|^2 dx \Big)^{1/2}\end{aligned}$$

Taking the sum over all $k \in \mathbb{Z}$ and using Minkowski's inequality for the sum gives

$$\|\hat{f}\|_2 \leq \sqrt{2b} \Big(\int_E |\hat{f}|^2 d\sigma \Big)^{1/2} + \frac{b}{\sqrt{2}} \|\nabla \hat{f}\|_2$$

Rearranging this gives

$$\int |x|^2 |f(x)|^2 dx = \|\nabla \hat{f}\|_2^2 \geq 2(1 - (2c_1)^{1/2})^2 / b^2$$

as desired. $\square$

Strichartz also extended this to an uncertainty principle on the sphere (see [Str89, Thm 2.2,p. 102]).

Part III

Uncertainty Principles for Operators

Chapter 6

Operators and Their Adjoints

In this section I will investigate operators that in some way relate to their adjoints. I will demonstrate uncertainty principles for these operators. First I will prove the uncertainty principle for self-adjoint and skew-adjoint operators. Then I will extend the result to normal or symmetric operators. I will assume the reader knows the definition of these types of operators (see [Ped89]).

6.1 Self-adjoint and Skew-adjoint Operators

In this section I will present the uncertainty principle for self-adjoint and skew-adjoint operators. I also include some results about operators of this kind that will be used later.

Proposition 6.1. If A and B are self-adjoint or skew-adjoint operators then

$$\|Au\|\|Bu\| \geq \frac{1}{2}|\langle [A,B]u,u\rangle| \qquad \text{for all } u \in \mathcal{D}([A,B])$$

Proof. Let A and B be skew-adjoint (the self-adjoint case follows from the same argument). If $u \in \mathcal{D}([A,B])$ then $u \in \mathcal{D}(A) \cap \mathcal{D}(B)$ so the following has meaning

$$\begin{aligned} |\langle [A,B]u,u\rangle| &= |\langle ABu,u\rangle - \langle BAu,u\rangle| = |\langle Bu,-Au\rangle - \langle Au,-Bu\rangle| \\ &= |\langle Au,Bu\rangle - \langle Bu,Au\rangle| = 2|\operatorname{Re}\langle Au,Bu\rangle| \leq 2\|Au\|\|Bu\|. \end{aligned}$$

□

Example 6.2 (Heisenberg). Assume $f \in L_2(\mathbb{R}^n)$. Let $A : f(x) \mapsto x_i f(x)$ for functions for which $x_i f(x)$ is also in L_2. Then A is self-adjoint. Let $B : f(x) \mapsto \frac{\partial f}{\partial x_i}(x)$ for functions for which $\frac{\partial f}{\partial x_i}(x)$ is in L_2. Assume f is in $\mathcal{D}([A,B]) = \{f \in L_2(\mathbb{R}^n) | x_i f, x_i \frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_i} \in L_2(\mathbb{R})\}$. The following calculations hold

$$\begin{aligned} [A,B]f(x) &= AB(f(x)) - BA(f(x)) \\ &= x_i \frac{\partial}{\partial x_i} f(x) - \frac{\partial}{\partial x_i}(x_i f(x)) \\ &= x_i \frac{\partial}{\partial x_i} f(x) - f(x) - x_i \frac{\partial}{\partial x_i} f(x) = f(x). \end{aligned}$$

This gives the following inequality

$$\frac{1}{2}\|f\|_2 \leq \left\|\frac{\partial f}{\partial x_i}\right\|_2 \|x_i f\|_2.$$

There is a slight extension of Proposition 6.1

Proposition 6.3. If A and B are self-adjoint or skew-adjoint operators on a complex (or real) Hilbert space and a, b are complex (or real) numbers then for all $u \in \mathcal{D}([A, B])$

$$\|(A + aI)u\|\|(B + bI)u\| \geq \frac{1}{2}|\langle [A, B]u, u\rangle|.$$

Proof. Since aI and bI commute with all operators we get that $[A + aI, B + bI] = [A, B]$. Using first the triangle inequality and then Cauchy-Schwarz's inequality

$$\begin{aligned}
|\langle [A, B]u, u\rangle| &= |\langle [A + a, B + b]u, u\rangle| \\
&\leq |\langle (B + b)u, (A^* + \bar{a})u\rangle| + |\langle (A + a)u, (B^* + \bar{b})u\rangle| \\
&\leq \|(A^* + \bar{a})u\|\|(B + b)u\| + \|(A + a)u\|\|(B^* + \bar{b})u\|.
\end{aligned}$$

The rest follows from the fact that $\|A + aI\| = \|A^* + \bar{a}I\|$ since

$$\begin{aligned}
\|(A^* + \bar{a}I)u\|^2 &= \|A^*u\|^2 + |a|^2\|u\|^2 - \langle \bar{a}u, A^*u\rangle - \langle A^*u, \bar{a}u\rangle \\
&= \|Au\|^2 + |a|^2\|u\|^2 - \bar{a}\langle Au, u\rangle - a\langle A^*u, u\rangle \\
&= \|Au\|^2 + |a|^2\|u\|^2 - \langle Au, au\rangle - \langle au, Au\rangle \\
&= \|(A + aI)u\|^2.
\end{aligned}$$

The same holds for B. $\square$

Remark 6.4. Most of the calculations used to show that $\|(A+aI)u\| = \|(A^*+\bar{a}I)u\|$ use the definition of an adjoint operator. The only question is whether $\|Au\| = \|A^*u\|$, which holds for self-adjoint and skew-adjoint operators. It is true for other operators as well, a fact I return to in the next section.

Assume that A and B are closed self-adjoint operators. It does not necessarily hold that $[A, B]$ is closed. But the commutator is densely defined and skew-adjoint so it is closable. Let $C = \overline{[A, B]}$, then it is interesting to see if the uncertainty principle

$$\|Au\|\|Bu\| \geq \frac{1}{2}|\langle Cu, u\rangle|$$

still holds for $u \in \mathcal{D}(A) \cap \mathcal{D}(B) \cap \mathcal{D}(C)$. As the following example will demonstrate, this is not always the case. In chapter 7 section 7.6, I will deal with a specific algebra of operators for which it always holds.

Example 6.5 (Uncertainty principle fails). The example is taken from [FS97, p. 211]. Consider the Hilbert space $L_2([0,1])$. Let $Af = f'$ with

$$\mathcal{D}(A) = \{f \in L_2([0,1]) | f \text{ is differentiable and } f(0) = f(1)\},$$

which is dense in $L_2([0,1])$. Let $Bf(x) = xf(x)$ with $\mathcal{D}(B) = L_2([0,1])$. Then $Bf \in \mathcal{D}(A)$ only if $f(0) = f(1) = 0$, so $\mathcal{D}([A,B]) = \{f \in \mathcal{D}(A) | f(0) = f(1) = 0\}$ and $[A,B] = I$. The domain of $[A,B]$ is still dense in $L_2([0,1])$, though smaller than $\mathcal{D}(A)$, and since $[A,B]$ is bounded, its closure is $C = \overline{[A,B]} = I$ on the whole of $L_2([0,1])$. But if $f(x) = 1$ for all $x \in [0,1]$, then $Af = 0$, $Bf(x) = x$ and $Cf = f$, such that the uncertainty principle

$$\|Af\|_2 \|Bf\|_2 \geq \frac{1}{2} \|f\|_2$$

does not hold.

The problem is that if a sequence $f_n \in \mathcal{D}(A)$ tends to $f(x) = 1$ in $L_2([0,1])$ then $Af_n = f_n'$ does not tend to $Af = 0$ in $L_2([0,1])$, a fact that I will now show.

Let f_n be differentiable and tend to $f(x) = 1$ in $L_2([0,1])$. Assume that $f_n' \to 0$ in $L_2([0,1])$ and write $\tilde{f} = \lim_{n\to\infty} f_n'$ is in $L_2([0,1])$. Setting $g(x) = x$ I get the inner product is

$$\int_0^1 x f_n'(x) dx = [x f_n(x)]_0^1 - \int_0^1 f_n(x) dx = -\int_0^1 f_n(x) dx.$$

When $n \to \infty$ this tends to -1 since $f_n \to f$ in $L_2([0,1])$. This is in contradiction with the assumption $f_n' \to 0$.

I will return to this example in section 9.3.

Later I will need a variant of Stone's theorem, so I state it here. Since it was part of my curriculum in fourth year, I will not include the proofs.

Definition 6.6. A strongly continuous one-parameter unitray group is a strongly continuous function U from $\mathbb{R}$ to the unitary operators on at Hilbert space such that $U_{s+t} = U_s U_t$ for all $s, t \in \mathbb{R}$.

The first result is Proposition 5.3.13 in [Ped89].

Theorem 6.7. If S is a self-adjoint operator in the separable Hilbert space H, let $U_t = \exp(itS)$ for $t \in \mathbb{R}$. Then U_t is a strongly continuous one-parameter unitary group and for each $x \in \mathcal{D}(S)$

$$Sx = \lim_{t\to 0} \frac{U_t x - x}{it}.$$

If for $x \in H$ the limit above exists then $x \in \mathcal{D}(S)$ and the limit is Sx.

Remark 6.8. For a skew-adjoint operator S the same statement holds with $U_t = \exp(tS)$ and the limit

$$Sx = \lim_{t\to 0} \frac{U_t x - x}{t}.$$

This is seen by replacing S with the self-adjoint operator $-iS$.

Stone's theorem is Theorem 5.3.15 in [Ped89]

Theorem 6.9 (Stone). If $(U_t)_{t\in\mathbb{R}}$ is a strongly continuous one-parameter unitary group, there is a self-adjoint operator S on H, such that $U_t = \exp(itS)$ for all t.

6.2 Symmetric and Normal Operators

In physical systems we do not always have self-adjoint operators. An example is the Laplace operator, which I will present as an example at the end of this section. Another example is the Dunkl operator, which has been treated in [Sel02, Section 5.1]. So here I extend the uncertainty principle to also hold for symmetric and normal operators. The results are similar to the results for self-adjoint and skew-adjoint operators, since these operators are normal. Here I will also investigate the case of equality in the uncertainty principle. To do this I will need

To extend the results for self-adjoint operators slightly I will need the following

Lemma 6.10 (Cauchy-Schwarz's inequality). For x, y in a Hilbert space the following always holds

$$|\langle x, y\rangle| \le \|x\|\|y\|,$$

and equality only holds if there are constants $a, b \in \mathbb{C}$ with $|a| + |b| > 0$ such that $ax = by$.

Proof. The inequality is a well known result. If there are constants as described above the equality follows by straight forward calculations. Assume now that equality holds. If $x = 0$ then we can choose $a = 1$ and $b = 0$. Similarly for $y = 0$. So now we can assume that $x \neq 0$ and $y \neq 0$. Let $x' = x/\|x\|$ and $y' = y/\|y\|$, then, since we consider the case of equality, $\langle x', y'\rangle = e^{it}$ for some $t \in [0, 2\pi]$. But then $\langle e^{-it}x', y'\rangle = 1$ and

$$\|e^{-it}x' - y'\|^2 = \|x'\|^2 + \|y'\|^2 - \langle e^{-it}x', y\rangle - \langle y', e^{-it}x'\rangle = 0.$$

So we can choose $b = 1/\|y\|$ and $a = e^{-it}/\|x\|$. $\square$

The following result is found in [Sel02].

Theorem 6.11. If $A, B : H \to H$ are symmetric or normal operators on a Hilbert space H, and $a, b \in \mathbb{C}$ are given, then

$$\|(A-a)u\|\|(B-b)u\| \ge \frac{1}{2}|\langle [A, B]u, u\rangle| \tag{6.1}$$

for all $u \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$. For given $u \neq 0$ and a, b equality holds if and only if there are constants $c_1, c_2, d_1, d_2 \in \mathbb{C}$ with $(c_1 + d_1)(c_2 + d_2) > 0$ such that

$$c_1(A^* - \bar{a})u = d_1(B - b)u \qquad \text{and} \qquad c_2(A - a)u = d_2(B^* - \bar{b})u.$$

Either at least one of these constants equal to zero or $d_1/c_1 = -\bar{d}_2/\bar{c}_2$.

Proof. Let u be in $\mathcal{D}(AB) \cap \mathcal{D}(BA)$. Using first the triangle inequality and then Cauchy-Schwarz's inequality (as for skew-adjoint operators) we get

$$\begin{aligned} |\langle [A,B]u, u\rangle| &= |\langle [A-a, B-b]u, u\rangle| \\ &\le |\langle (B-b)u, (A^*-\bar{a})u\rangle| + |\langle (A-a)u, (B^*-\bar{b})u\rangle| \\ &\le \|(A^*-\bar{a})u\|\|(B-b)u\| + \|(A-a)u\|\|(B^*-\bar{b})u\|. \end{aligned} \tag{6.2}$$

If A is normal then $\|Au\| = \|A^*u\|$ (see [Rud91, Thm. 12.12]). If A is symmetric then for $u \in \mathcal{D}(A)$ it holds that $A^*u = Au$ and therefore $\|A^*u\| = \|Au\|$. As in the proof of Proposition 6.3 we get

$$\|(A^*-\bar{a})u\| = \|(A-a)u\| \quad \text{and} \quad \|(B^*-\bar{b})u\| = \|(B-b)u\| \tag{6.3}$$

if A and B are normal or symmetric.

Given a, b and u, then equality in Cauchy-Schwarz's inequality holds if and only if there are constants $c_1, c_2, d_1, d_2 \in \mathbb{C}$ with $|c_j| + |d_j| > 0$ such that

$$c_1(A^*-\bar{a})u = d_1(B-b)u \qquad \text{and} \qquad c_2(A-a)u = d_2(B^*-\bar{b})u. \tag{6.4}$$

If $c_1 = 0$ then $d_1 \neq 0$ and thus $(B-b)u = 0$ which gives $(B^*-\bar{b})u = 0$ by (6.3). The same argument holds if $d_1 = 0$, so the equality in Cauchy-Schwarz's inequality holds. If none of the constants are zero, then (6.4) and (6.3) give

$$\begin{aligned} \|(A-a)u\| &= \|(A^*-\bar{a})u\| = |d_1/c_1|\|(B-b)u\| \\ &= |d_1/c_1|\|(B^*-\bar{b})u\| = |d_1/c_1||c_2/d_2|\|(A-a)u\| \end{aligned}$$

so that

$$|d_1/c_1| = |d_2/c_2|. \tag{6.5}$$

Equality in the triangle inequality (6.2) gives

$$\begin{aligned} &|c_1/d_1|\|(A^*-\bar{a})u\|^2 + |c_2/d_2|\|(A-a)u\|^2 \\ &\qquad = |\langle (B-b)u, (A^*-\bar{a})u\rangle - \langle (A-a)u, (B^*-\bar{b})u\rangle| \\ &\qquad = \left|\frac{c_1}{d_1}\right| \|(A^*-\bar{a})u\|^2 - \frac{\bar{c}_2}{\bar{d}_2}\|(A-a)u\|^2. \end{aligned}$$

From this we deduce

$$|c_1/d_1| + |c_2/d_2| = |c_1/d_1 - \bar{c}_2/\bar{d}_2|. \tag{6.6}$$

(6.6) together with (6.5) tells us that

$$2|c_1/d_1| = |c_1/d_1 - \bar{c}_2/\bar{d}_2|,$$

which can only be fulfilled if $c_1/d_1 = -\bar{c}_2/\bar{d}_2$. □

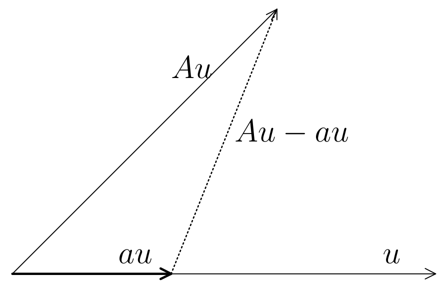

Figure 6.1: $\|Au - au\|$ is minimized when a is the orthogonal projection of Au on u.

Theorem 6.11 holds for arbitrary a and b, but we might want to ask when the left hand side in (6.1) is minimized. The minimum of $\|(A-a)u\|$ is attained when au is the projection of Au on u (see Figure 6.1).

On this background define the variance

$$\sigma_A(u) = \min_a \|Au - au\| = \left\| Au - \frac{\langle Au, u\rangle}{\|u\|^2} u \right\|. \tag{6.7}$$

Then from the theorem above we get that

Corollary 6.12. If A, B are symmetric or normal operators on a Hilbert spcae H, then

$$\sigma_A(u)\sigma_B(u) \geq \frac{1}{2}|\langle [A,B]u, u\rangle|$$

for all nonzero $u \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$.

The following example is found in section 5.2 in [Sel02], I have extracted the most important information.

Example 6.13 (The Laplace operator). Let ω be a probability function and define the inner product

$$\langle f, g\rangle = \int f(x)\overline{g(x)}\omega(x)dx.$$

Consider the Hilbert space $L_2([0,\pi], \omega_\alpha)$ with

$$\omega_\alpha(t) = c_\alpha(\sin x)^{2\alpha+1}, \qquad c_\alpha = \frac{\Gamma(2\alpha+2)}{\Gamma(\alpha+1)^2 2^{2\alpha+1}}$$

where Γ is the gamma-function (see Definition 3.7) and $\alpha \geq -1/2$.

Define the Laplace operator

$$L_\alpha f(x) = -\Big(f''(x) + (2\alpha+1)\frac{\cos x}{\sin x} f'(x)\Big),$$

with $\mathcal{D}(L_\alpha) = \{f \in C^2([0,\pi]) | f'(0) = f'(\pi)0\}$. Then L_α is symmetric (to show this [Sel02, p. 170] refers to [RV97, Lemma 3.1]). Let $Hf = hf$ where $h \in \mathcal{D}(L_\alpha)$ then $H^*f = \bar{h}f$, so H is normal. The operator $-L_\alpha H$ is given by

$$\begin{aligned}
-L_\alpha(hf)(x) &= (hf)''(x) + (2\alpha+1)\frac{\cos x}{\sin x}(hf)'(x) \\
&= h''(x)f(x) + h(x)f''(x) + 2h'(x)f'(x) \\
&\qquad + (2\alpha+1)\frac{\cos x}{\sin x}(h'(x)f(x) + h(x)f'(x)). \\
&= 2h'(x)f'(x) - L_\alpha h(x) - L_\alpha f(x).
\end{aligned}$$

Therefore the commutator is

$$[H, L_\alpha]f = hL_\alpha f - L_\alpha(hf) = 2h'f' - L_\alpha h.$$

The uncertainty principle (Corollary 6.12) then gives

$$\Big(\|hf\| - \frac{|\langle hf, f\rangle|^2}{\|f\|^2}\Big)\Big(\|L_\alpha f\| - \frac{|\langle L_\alpha f, f\rangle|^2}{\|f\|^2}\Big) \geq \frac{1}{4}|\langle 2f'h' - fL_\alpha h, f\rangle|^2.$$

To get the result above, I have used that

$$\Big\|Af - \frac{\langle Af, f\rangle}{\|f\|^2}f\Big\|^2 = \|Af\|^2 - \frac{|\langle Af, f\rangle|^2}{\|f\|^2},$$

since $Af - \frac{\langle Af, f\rangle}{\|f\|^2}f$ is the orthogonal projection of Af on f.

Chapter 7

Operators Generated by Lie-groups

For operators A,B and C such that $C = \overline{[A, B]}$ the uncertainty principle

$$\|Au\|\|Bu\| \geq \frac{1}{2}|\langle Cu, u\rangle|,$$

does not always hold. This was shown in Example 6.5 in the previous chapter. Now I will go on to investigate a class of operators for which the uncertainty principle above holds. A lot of theory is needed before I can state and prove the main result: Theorem 7.33. I assume the reader is familiar with the first four chapters of [War71].

7.1 Invariant Measure on a Lie-group

In this section I will investigate the Haar-meassure on a Lie group. In [CR97] we proved that there exists a right-invariant Haar-measure on any locally compact group. We also proved that it is unique up to multiplication by a positive constant. See [Hal50, p. 250ff] for an english reference. Since a Lie group is related to a Lie algebra which is a vectorspace and thus has a Haar-measure, I will link these two measures. The steps that are presented here actually show the existence of a right invariant Haar-measure on Lie groups. As inspiration I have used some lecture notes by Henrik Stetkær. The notes are in danish and not publicly available, but the main result, equation (7.2), is in fact just a consequence of Theorem 1.14 in [Hel84].

For a Lie group G let $\mathfrak{g}$ be the corresponding Lie algebra. Now $(\mathfrak{g}, +)$ can also be regarded as a Lie group. Let $T_L\mathfrak{g}$ be the tangent space for $L \in \mathfrak{g}$. Since $\mathfrak{g}$ is a vectorspace, $T_L\mathfrak{g}$ can be identified with $\mathfrak{g}$ by associating to $v \in \mathfrak{g}$ the tangent vector of the curve $\alpha_{L,v} : t \mapsto L + tv$. The inverse of this mapping $\Phi_L : T_L\mathfrak{g} \to \mathfrak{g}$ is called the canonical identification. Notice that $\Phi_L^{-1}(v) = \alpha'_{L,v}(0)$.

Let r_a be right translation of a function by a.

Definition 7.1. For any $L \in \mathfrak{g}$ define the map $J(L) : \mathfrak{g} \to \mathfrak{g}$ by

$$J(L) = (dr_{\exp(-L)})_{\exp L} \circ (d\exp)_L \circ \Phi_L^{-1}.$$

Using $\alpha_{L,Y}$ as defined above, we see that $J(L)(Y)$ is the tangent vector to the curve $\gamma(t) = r_{\exp(-L)} \circ \exp \circ \alpha_{L,Y}(t)$, since $\gamma'(0) = dr_{\exp(-L)} \circ d\exp_L \circ \Phi_L^{-1}(Y)$.

Lemma 7.2. The map $L \mapsto J(L)$ is a smooth map from $\mathfrak{g}$ to $End(\mathfrak{g})$.

Proof. Since $J(L)$ is linear (and thus can be represented by a matrix) it is enough to prove that $L \mapsto J(L)L_0$ is smooth for any chosen $L_0 \in \mathfrak{g}$ (it is actually enough to only look at a basis for $\mathfrak{g}$). It is well known that if (ϕ, U) is a map and $e \in U$, then $x_i : a \mapsto \phi(a)_i$ is differentiable and a vector field can be written $L_a = L_a(x_i)\frac{\partial}{\partial x_i}|_a$ for $a \in G$ (see Proposition 1.43 and Remark 1.20(b) in [War71]). Also if α is a curve, such that $\alpha(0) = a$ and $\alpha'(0) = L$, then $L_a(f) = (f \circ \alpha)'(0)$. Using this to find $L_a(x_i)$ for the vector field $J(L)L_0$ gives

$$J(L)L_0(x_i) = \frac{d}{dt}\Big|_{t=0} x_i \circ \alpha(t) = \frac{d}{dt}\Big|_{t=0} x_i(\exp(L + tL_0)\exp(-L))$$

From this is seen that $L \mapsto J(L)L_0(x_i)$ is smooth for all i and thus $L \mapsto J(L)L_0$ is smooth. □

Lemma 7.3. Define the map

$$\rho : L \mapsto |\det(J(L))|.$$

This map is smooth where $J(L)$ is invertible.

Proof. If $J(L)$ is invertible then $\det(J(L))$ is not zero. The determinant is an analytical function (it is a polynomial) and J is smooth so ρ is continuous. Therefore there is a neighbourhood of L such that $\rho(L)$ is not zero (this means that it is either negative or positive on the neighbourhood). Here the function $x \mapsto |x|$ is also smooth. So ρ is smooth where $J(L)$ is invertible. □

Let N_0 be an open neighbourhood of 0 in $\mathfrak{g}$ and N_e an open neighbourhood of e in G such that $\exp|_{N_0} : N_0 \to N_e$ is a diffeomorphism (this can be done according to [War71, Thm. 3.31(d)]). Then ρ is smooth on N_0, because $d\exp$ is close to the identity in this neighbourhood. Fix a Haar-measure dL on $\mathfrak{g}$ and for $f \in C_c(N_e)$ define

$$I_1(f) = \int_{N_0} f(\exp L)\rho(L)dL.$$

Lemma 7.4. If $a \in G$ and both f and $r(a)f$ (right translation of f by a) are in $C_c(\exp N_0)$ then $I_1(f) = I_1(r(a)f)$.

Proof. $\mathrm{supp}(f)$ is compact so there exists an open neighborhood U_1 such that $\mathrm{supp}(f) \subseteq U_1 \subseteq N_e$ (this is a consequence of [CR97, Lem. 3.9], but for an english reference use Theorem 4.10 p. 20 in [HR63]). Since $\mathrm{supp}(r(a)f)$ is compact there is an open neighborhood U_2 such that $\mathrm{supp}(r(a)f) \subseteq U_2 \subseteq N_e$. Then $\mathrm{supp}(f) \subseteq U_2 a$ and $U = U_1 \cap U_2 a$ is open, and while U covers $\mathrm{supp}(f)$, Ua^{-1} covers $\mathrm{supp}(r(a)f)$. Both U and Ua^{-1} are inside N_e. This shows that we can choose an open neighborhood $\exp V$ of $\mathrm{supp}(f)$ so that $V \subseteq N_0$ and $(\exp V)a^{-1} \subseteq \exp N_0$.

Since $\operatorname{supp}(r(a)f) = \operatorname{supp}(f)a^{-1} \subseteq \exp Va^{-1}$ we can make the change of variable $\exp Y = \exp La^{-1}$ to get

$$\begin{aligned}
I_1(r(a)f) &= \int_{N_0} r(a)f(L)(\exp L)\rho(L)dL = \int_{\log(\exp Va)} f(\exp La^{-1})\rho(L)dL \\
&= \int_V f(\exp Y)\rho(\phi^{-1}(Y))|J_{\phi^{-1}}(Y)|dY
\end{aligned}$$

where $\phi(L) = \log(\exp La^{-1})$ is a diffeomorphism from $\log(\exp Va)$ to V. The last equality is due to a change of variable on a real vector space (see [Rud87, Thm 7.26]). The support $\operatorname{supp}(f) \subseteq \exp V$ so to show the desired we only need to show that for all $Y \in V$

$$\rho(\phi^{-1}(Y))|J_{\phi^{-1}}(Y)| = \rho(Y). \tag{7.1}$$

First I will rewrite

$$\begin{aligned}
r_{\exp(-\phi^{-1}(Y))} \circ \exp \circ \phi^{-1}(L) &= r_{\exp(-\phi^{-1}(Y))} \circ \exp(\log(\exp La)) \\
&= (\exp La)(\exp(-\log(\exp Ya))) \\
&= (\exp La)(\exp Ya)^{-1} \\
&= \exp L \exp(-Y) \\
&= r_{\exp(-Y)} \circ \exp(L).
\end{aligned}$$

Using this in the last equality, and the chain rule in the second last equality, in the following calculations, gives us

$$\begin{aligned}
&\rho(\phi^{-1}(Y))|J_{\phi^{-1}}(Y)| \\
&= |\det(dr_{\exp(-\phi^{-1}(Y))} \circ d\exp \circ \Phi^{-1}_{\phi^{-1}(Y)})||\det(\Phi_{\phi^{-1}(Y)} \circ d\phi^{-1} \circ \Phi_Y^{-1})| \\
&= |\det(dr_{\exp(-\phi^{-1}(Y))} \circ d\exp \circ \Phi^{-1}_{\phi^{-1}(Y)} \circ \Phi_{\phi^{-1}(Y)} \circ d\phi^{-1} \circ \Phi_Y^{-1})| \\
&= |\det(dr_{\exp(-\phi^{-1}(Y))} \circ d\exp \circ d\phi^{-1} \circ \Phi_Y^{-1})| \\
&= |\det d(r_{\exp(-\phi^{-1}(Y))} \circ \exp \circ \phi^{-1} \circ \Phi_Y^{-1})| \\
&= |\det d(r_{\exp(-Y)} \circ \exp \circ \Phi_Y^{-1})| = \rho(Y).
\end{aligned}$$

This finishes the proof of the lemma according to (7.1). □

If $f \in C_c(G)$ and there exists a $a \in G$ such that $\operatorname{supp}(f) \subseteq N_e a^{-1}$ then define

$$I_2(f) = I_1(r(a)f).$$

The preceding lemma shows that the definition is independent on the choice of a.

The group G can be covered by translations $N_e a$ of N_e. Now let $\{\psi_a | a \in A\}$ be a countable smooth partition of unity (see Definition 1.8 in [War71]), which exists according to [War71, Thm. 1.11]. Then $\psi_a f$ has support inside a translation of N_e and is different from zero for finitely many a. Then $f(a) = \sum_{a\in A} \psi_a f(a)$ and we can define the integral of f to be

$$I(f) = \sum_{a\in A} I_2(\psi_a f).$$

This is a right invariant Haar-measure on G. We note that if f has support inside N_e then $I(f) = I_1(f)$.

Using the topology of G we can fix a right invariant Haar-measure da on G (see [Hal50]). Since we now have two right-invariant Haar-measures on G they are proportional, and therefore the integral of $f \in C_c(N_e)$ can be written

$$\int_G f(a)da = \int_{N_0} f(\exp L)k(L)dL \tag{7.2}$$

where k is a smooth non-negative function with $k(0) \neq 0$. The function k is just a constant multiplication of ρ. This is the main result of the section. Its equivalent is Theorem 1.14 in [Hel84].

We define the function $\Delta : G \to \mathbb{R}_+$ by

$$\int_G f(ba)da = \Delta(b) \int_G f(a)da, \quad b \in G\text{and } f \in C_c(G).$$

Proposition 7.5. The modular function $\Delta : G \to (\mathbb{R}_+, \cdot)$ is a Lie group homomorphism. It is given by

$$\Delta(a) = |\det(Ad(a))|. \tag{7.3}$$

Proof. The following calculation with $b, c \in G$ shows that Δ is a group homomorphism

$$\Delta(cb) \int_G f(a)da = \int_G f(cba)da = \Delta(c) \int_G f(ba)da = \Delta(c)\Delta(b) \int_G f(a)da.$$

To show that the modular function if smooth I will first show that it is in fact given by (7.3). Let N_e be neighborhood of e and N_0 neighborhood of 0 such that $\exp : N_0 \to N_e$ is a diffeomorphism. Let $a \in G$ be given and choose $f \in C_c(G)$ with $\operatorname{supp}(f)$ so small that both $f(\exp L)$ and $f(\exp Ad(a)L)$ have support inside N_0. Then

$$\begin{aligned}
\Delta(a) \int_{N_0} f(\exp L)\rho(L)dL &= \Delta(a) \int_G f(b)db \\
&= \int_G f(ab)db \\
&= \int_G f(aba^{-1})db
\end{aligned}$$

equation (7.2) and $\exp(Ad(a)L) = a\exp(L)a^{-1}$ give

$$= \int_{N_0} f(\exp(Ad(a)L))\rho(L)dL$$

letting $Y = Ad(a^{-1})L$ and using the rule for change of variable [Rud87, Thm. 7.26] we get

$$= \int_{N_0} f(\exp Y)\rho(Ad(a)Y)|\det(Ad(a))|dY$$
$$= |\det(Ad(a))| \int_{N_0} f(\exp Y)\rho(Ad(a)Y)dY$$

Since $\mathfrak{g}$ is a vector space we can define a δ-distribution on $\mathfrak{g}$. Let $f \circ \exp$ tend to the δ-distribution at 0. Doing so we get $\rho(Ad(a)Y) \to \rho(0) = 1$, yielding the desired equality.

$Ad(a)$ is invertible so $\det(Ad(a))$ is either positive or negative in a neighborhood of a. Since $a \mapsto Ad(a)$ is smooth it also follows that Δ is a smooth function. $\square$

7.2 Lie Theory

Representations can be defined very differently so here I state the definition used in this text.

Definition 7.6. A representation π of a Lie-group G on a Hilbert-space H is a homomorphism $\pi : G \to Aut(H)$ such that $(a, v) \mapsto \pi(a)v$ is continuous. It will be written (H, π). A representation is unitary if $\pi(a)$ are unitary for all $a \in G$.

The following result will not be used later, but I needed it to understand the proof of Theorem 2.4 in [FS97].

Proposition 7.7. If $\mathfrak{I}$ is an ideal in a finite dimensional Lie-algebra $\mathfrak{g}$ then for $L_1 \in \mathfrak{I}$ and $L_2 \in \mathfrak{g}$ it holds that $\exp(ad(L_2))L_1$ is in $\mathfrak{I}$.

Proof. The Lie-bracket is bilinear and the ideal $\mathfrak{I}$ is therefore a subspace of $\mathfrak{g}$. Since $\mathfrak{g}$ is of finite dimension it follows that $\mathfrak{I}$ is closed. Thus any converging series in $\mathfrak{I}$ has its limit in $\mathfrak{I}$.

Since $\mathfrak{g}$ has dimension $n < \infty$ the linear mappings from $\mathfrak{g}$ to $\mathfrak{g}$ (denoted $End(\mathfrak{g})$) can be represented by all $n \times n$ matrices on $\mathbb{R}$ which is the Lie-algebra $\mathfrak{gl}(n, \mathbb{R})$ of $GL(n, \mathbb{R})$. The exponential mapping on $\mathfrak{gl}(n, \mathbb{R})$ is known to be the mapping

$$x \mapsto \sum_{j=0}^{\infty} \frac{x^j}{j!} \in GL(n, \mathbb{R}).$$

Also $Aut(\mathfrak{g})$ are the bijective linear mappings from $\mathfrak{g}$ onto $\mathfrak{g}$ and can therefore be represented by the matices $GL(n, \mathbb{R})$.

$ad(L_2)$ is in $End(\mathfrak{g})$ and $ad(L_2)Z = [L_2, L_1] \in \mathfrak{I}$ so

$$s_N L_1 = \Big(\sum_{i=0}^{N} \frac{(ad(L_2))^j}{j!}\Big)L_1 \in \mathfrak{I}.$$

This expression converges to $\exp(ad(L_2))L_1$ in $GL(n, \mathbb{R})$ which is the same as $Aut(\mathfrak{g})$. Since $\mathfrak{I}$ is a closed subspace it follows that $\exp(ad(L_2))L_1 \in \mathfrak{I}$. $\square$

Definition 7.8. Let L be a left-invariant vectorfield on G. We say that a right-invariant vectorfield R corresponds to L if $R_a = dT \circ L_{a^{-1}}$ where $T(a) = a^{-1}$.

First let me show that if L is a left invariant vector field, then R as defined is a right invariant vector field. Let $a, b \in G$ and $f \in C_c^\infty(G)$ and let l_b denote left translation by b. Since $f \circ R \circ r_b(x) = f \circ l_{b^{-1}} \circ T(x)$ and L is left invariant it follows

$$\begin{aligned} dr_b \circ R_a(f) &= dr_b \circ dT \circ L_{a^{-1}}(f) = dr_b \circ L_{a^{-1}}(f \circ T) \\ &= L_{a^{-1}}(f \circ T \circ r_b) = L_{a^{-1}}(f \circ l_{b^{-1}} \circ T) \\ &= dT \circ L_{a^{-1}}(f \circ l_{b^{-1}}) = dT \circ L_{b^{-1}a^{-1}}(f) = R_{ab}(f). \end{aligned}$$

So to a left invariant vector field there always exists a corresponding right invariant vector field. It is uniquely defined as above.

Proposition 7.9. Let $L \in \mathfrak{g}$ and let R correspond to L. Then

$$\begin{aligned} L_a(f) &= \frac{d}{dt}|_{t=0} f(a \exp(tL)) \qquad \text{and} \\ R_a(f) &= -\frac{d}{dt}|_{t=0} f(\exp(tL)a). \end{aligned}$$

Proof. By the definition of tangent vectors and integral curves we know that $L_a(f) = \frac{d}{dt}|_{t=0} f(a \exp(tL))$, see [War71, Thm. 3.31(e)]. Using the definition $R_a = dT \circ L_{a^{-1}}$ we get

$$\begin{aligned} R_a(f) &= dT \circ L_{a^{-1}}(f) \\ &= L_{a^{-1}}(f \circ T) \\ &= \frac{d}{dt}|_{t=0} f(T(a^{-1} \exp(tL))) \\ &= \frac{d}{dt}|_{t=0} f(\exp(-tL)a) \\ &= -\frac{d}{dt}|_{t=0} f(\exp(tL)a). \end{aligned}$$

This shows the desired equation. □

From this proposition it follows

Corollary 7.10. If $f \in C_c^\infty(G)$, then $L(f)$ and $R(f)$ defined as $L(f)(a) = L_a(f)$ and $R(f)(a) = R_a(f)$ are in $C_c^\infty(G)$. The supports of $L(f)$ and $R(f)$ are contained in the support of f.

7.3 Smooth Vectors

This section is equivalent to Chapter 3 section 3 (p. 51) in [Kna86]. Where Knapp uses right invariant vectorfield I use left invariant vectorfields.

Recall that a function f from $\mathbb{R}$ or $\mathbb{C}$ to a vector space V is differentiable in x_0 if $(f(x) - f(x_0))/(x - x_0)$ has a limit for $x \to x_0$. Then I can consider differentiability of a function from G to V.

Definition 7.11. Let (H, π) be a representation for the Lie-group G. A vector $x \in H$ is called smooth for π if the map $a \mapsto \pi(a)x$ is $C^\infty(G)$. The collection of smooth vectors for π will be denoted H_π^∞.

It is seen that H_π^∞ is a subspace of H and therefore also a Hilbert-space with the inner product inherited from H.

Proposition 7.12. If (H, π) is a representation of the Lie-group G and if $S \subseteq H$ is a subspace such that for $x \in S$ the limit

$$\varphi(L)x = \lim_{t \to 0} \frac{\pi(exptL)x - x}{t} \tag{7.4}$$

exists and is in S for all $L \in \mathfrak{g}$, then $S \subseteq H_\pi^\infty$.

Until now $\varphi(L)$ is just short for the limit operation, but later I will show that it is a densely defined operator, defined on the smooth vectors.

Proof. This will be proved by induction.

For $x \in S$ define $f_x : G \to H$ by $f_x(a) = \pi(a)x$. By (7.4) all partial derivatives $L(f_x)$ exists in $a = e$ and are given by $\varphi(L)x$. Since $\pi(a)$ is continuous we can apply it to both sides in (7.4) to translate the limit to any point $a \in G$, showing that f_x has partial derivatives equal to $a \mapsto \pi(a)\varphi(L)x$. This is a continuous mapping and therefore f_x is C^1.

Now assume that f_x is a C^k-mapping for all $x \in S$. The assumption is that $\varphi(L)x$ is in S which then again leads to the mapping $a \mapsto \pi(a)\varphi(L)x$ being C^k. But this mapping is the same as the partial derivative $L(f_x)$. Therefore f_x has partial derivatives of order $k+1$ in all directions. These are continuous and f_x is therefore C^{k+1}.

By induction it follows that f_x is C^∞. □

7.4 The Gårding Subspace

Still following the development in [Kna86] I move on to Chapter 3, Section 4. Knapp uses a left invariant Haar-measure, but I use a right invariant measure. Apart from this the claims and proofs correspond to those in [Kna86].

Let π be a representation of a Lie-group G on a Hilbert space H. Let da denote a right invariant Haar-measure on G. For $f \in C_c^\infty(G)$ the mapping $a \mapsto f(a)\pi(a)x$ is continuous from G to H, since f and π are continuous ($a \mapsto \pi(a)x$ is continuous for all $x \in H$). The mapping has compact support since f has compact support, and is therefore integrable. Remember that the integral, $\int_G f(a)da$, of a vector valued function f is defined as the vector v for which

$$\langle v, g \rangle = \int_G \langle f(a), g(a) \rangle da,$$

for all vector valued functions g. According to [Rud91, Thm. 3.29] the resulting integral is in H. This makes the following definition plausible.

Definition 7.13. Let (H, π) be a representation of a Lie-group G. Let da denote a right-invariant Haar-measure on G. The Gårding subspace of H for π is the linear subspace of all vectors of the form

$$\pi(f)x = \int_G f(a)\pi(a^{-1})x da \tag{7.5}$$

for $x \in H$ and $f \in C_c^\infty(G)$.

Since da is unique up to multiplication with a constant the subspace defined above does not depend on the chosen right-invariant Haar-measure.

If $\pi(f)$ is regarded as an operator on H then since $x \mapsto \pi(f)x$ is composed by continuous functions $\pi(f)$ is a bounded operator.

The following proposition is composed by result from [Kna86] and [Seg51, Lemma 3.1.6].

Proposition 7.14. If (H, π) is a representation of G then the Gårding subspace is stable under the limit (7.4), which for $x \in H$ and $f \in C_c^\infty(G)$ is found to be

$$\varphi(L)(\pi(f)x) = \pi(L(f))x.$$

Also for any $L \in \mathfrak{g}$ define

$$\gamma = \frac{d}{dt}\Big|_{t=0} \Delta(\exp tL).$$

then for all y in the Gårding subspace

$$\pi(f)\varphi(L)y = -\pi(R(f))y - \gamma\pi(f)y,$$

where R is the right-invariant vectorfield corresponding to L.

Proof. Let $t \neq 0$ then

$$\frac{\pi(\exp tL) - I}{t}\pi(f)x = t^{-1}\int_G f(a)\pi(\exp tL)\pi(a^{-1})x da - t^{-1}\int_G f(a)\pi(a^{-1})x da$$

substituting a with $a \exp tL$ and using the right-invariance of da, we get

$$\begin{aligned}
&= t^{-1}\int_G f(a\exp tL)\pi(\exp tL)\pi((\exp tL)^{-1}a^{-1})x da \\
&\quad - t^{-1}\int_G f(a)\pi(a^{-1})x da \\
&= t^{-1}\int_G f(a\exp tL)\pi(a^{-1})x da - t^{-1}\int_G f(a)\pi(a^{-1})x da \\
&= \int_G \frac{f(a\exp tL) - f(a)}{t}\pi(a^{-1})x da
\end{aligned} \tag{7.6}$$

I need to show that the last integral converges for $t \to 0$. Construct the function $h : \mathbb{R} \times G \to H$ by $h(t, a) = f(a\exp tL)$. This function is continuous since it is constructed from continuous functions. Therefore $D = h([-1,1], \mathrm{supp}(f))$ is compact because f has compact support. Since f is differentiable with differential $L(f)(a)$ in a, it holds that $|(f(a \exp tL) - f(a))/t - L(f)(a)| + \epsilon$ for $t \leq \delta$. For all $|t| \leq \delta$ it thus holds that

$$\Big|\frac{f(a \exp tL) - f(a)}{t}\Big|| \leq L(f)(a)| + 1_D \epsilon.$$

Since $L(f)$ has compact support, the right hand side is integrable and independent of t, so the expression (7.6) has dominated convergence with limit $\int_G L(f)(a)\pi(a^{-1})xda = \pi(L(f))x$ for $t \to 0$.

Now to the second part of the proposition. Let $y = \pi(g)x$ be in the Gårding space, with $g \in C_c^\infty(G)$. Then

$$\pi(f)\varphi(L)\pi(g)x = \lim_{t \to 0} t^{-1} \int_G \int_G f(a)\pi(a^{-1})(g(b \exp tL) - g(b))\pi(b^{-1})xdbda$$

Splitting up the integral

$$t^{-1} \iint f(a)g(b \exp tL)\pi(a^{-1})\pi(b^{-1})xdbda - t^{-1} \iint f(a)g(b)\pi(a^{-1})\pi(b^{-1})xdbda$$

and substituting b with $b \exp(-tL)$ in the first term and a with $\exp(-tL)a$ in the second yields

$$\begin{aligned} &= t^{-1} \iint f(a)g(b)\pi(a^{-1})\pi(\exp tL)\pi(b^{-1})xdbda \\ &\quad - t^{-1} \iint \Delta(\exp tL) f(\exp(-tL)a)g(b)\pi(a^{-1})\pi(\exp tL)\pi(b^{-1})xdbda \end{aligned}$$

thus by subtracting $t^{-1} \iint f(\exp(-tL)a)g(b)\pi(a^{-1})\pi(\exp tL)\pi(b^{-1})xdbda$ in both terms I get

$$\begin{aligned} &= \iint t^{-1}(f(a) - f(\exp(-tL)))g(b)\pi(a^{-1})\pi(\exp tL)\pi(b^{-1})xdbda \\ &\quad - t^{-1}(\Delta(\exp tL) - 1) \iint f(\exp(-tL)a)g(b)\pi(a^{-1})\pi(\exp tL)\pi(b^{-1})xdbda. \end{aligned}$$

For $t \to 0$, $t^{-1}(\Delta(\exp tL) - 1) \to \gamma$ and $t^{-1}(f(a) - f(\exp(-tL))) \to -R_a(f)$, so the integrals above converge to

$$-\pi(R(f))\pi(g)x - \gamma\pi(f)\pi(g)x$$

This shows that for $y = \pi(g)x$ in the Gårding space

$$\pi(f)\varphi(L)y = -\pi(R(f))y - \gamma\pi(f)y.$$

This proves the last part of the proposition. □

From this and Proposition 7.12 follows that

Corollary 7.15. The Gårding-subspace is contained in H_π^∞.

Proposition 7.16. Let (H, π) be a representation of G. Let U_n be a sequence of neighborhoods of $e \in G$ such that $\cap_n U_n = \{e\}$. Let f_n be a sequence of positive smooth functions such that f_n has support inside U_n and $\lim_{n\to\infty} \int_G f_n(a)da = M$ exists. Let $g : G \to H$ be continuous and bounded. Then $\lim_{n\to\infty} g(a)f_n(a)da$ exists and is equal to $Mg(e)$.

Proof. f_n has compact support so $a \mapsto g(a)f_n(a)$ is continuous with compact support. Therefore the integral exists by [Rud91, Thm. 3.29]. The same theorem gives

$$\begin{aligned}
&\| \int_G g(a)f_n(a)da - Mg(e)\| \\
&\qquad \le \| \int_G g(a)f_n(a)da - \int_G g(e)f_n(a)da\| + \| \int_G g(e)f_n(a)da - Mg(e)\| \\
&\qquad \le \int_G \|g(a)f_n(a) - g(e)f_n(a)\|da + \| \int_G g(e)f_n(a)da - Mg(e)\| \\
&\qquad \le \sup_{a\in \mathrm{supp}(f_n)} (\|g(a) - g(e)\|) \int_G \|f_n(a)da + \|g(e)\| \Big| \int_G f_n(a)da - M\Big|.
\end{aligned}$$

The support $\mathrm{supp}(f_n)$ tends to e and g is continuous, so the last term above tends to 0 for $n \to \infty$. □

Proposition 7.17. Let (H, π) be a representation of G and let f_n be a sequence as in Proposition 7.16 with $M = 1$. Then for all $x \in H$ it holds that $\pi(f_n)x \to x$.

Proof. The Gårding vector $\pi(f_n)x$ is defined as

$$\pi(f_n)x = \int_G f_n(a)\pi(a^{-1})xda.$$

$a \mapsto \pi(a^{-1})x$ is continous and bounded on $\mathrm{supp}(f_n)$ so the proposition gives $\lim \pi(f_n)x = M\pi(e)x = x$. □

Remark 7.18. I will now show that a sequence of functions as described in Proposition 7.17 exists. They are suggested in the proof of Lemma 3.1.8 in [Seg51]. Let ϕ be the compactly supported smooth function

$$\phi(x) = \begin{cases} e^{-(1-x^2)^{-1}} & \|x\| < 1 \\ 0 & \|x\| \ge 1. \end{cases}$$

$A = \int_{\mathbb{R}} \phi(x)dx$ is positive and bounded. Now let

$$\phi_n(x) = A^{-1}n\phi(nx),$$

then $\int_{\mathbb{R}} \phi_n(x)dx = 1$ for all n, so $\|\phi_n\|_1$ is bounded by 1.

If x_i are canonical coordinates in a neighbourhood N_e around e let $f_n(a) = \prod_i \phi_n(x_i(a))$ for $a \in N_e$. From a certain n, f_n will have support inside N_e so only consider those n. Since f_n is continuous and has compact support it is integrable. The support of f_n is inside N_e, so (7.2) gives

$$\int_{\mathrm{supp}(f_n)} f_n(a)da = \int_{N_0} k(L) \prod_i \phi_n(x_i)dL,$$

where k is smooth and $k(0) \neq 0$.

Since $k(L)$ is smooth and bounded on the compact set $\exp^{-1}(\mathrm{supp}(f_n))$, the integral above converges to

$$\lim_{n\to\infty} \int_{\mathrm{supp}(f_n)} f_n(a)da = \lim_{n\to\infty} k(0) \int_{N_0} \prod_i \phi_n(x_i)dL = k(0).$$

The last equality can be shown as in the proof of Proposition 7.16. Thus if we set

$$f_n(a) = k(0)^{-1} \prod_i \phi_n(x_i(a)), \tag{7.7}$$

we get a function as described in Proposition 7.17.

From this remark the following theorem by Gårding follows

Theorem 7.19. If (H, π) is a representation of the Lie-group G then the Gårding-subspace is dense in H.

Corollary 7.20. If (H, π) is a representation of the Lie-group G then H_π^∞ is dense in H.

7.5 Representation of a Lie-algebra

In this section I will show that $\varphi(L)$ is a densely defined operator and that it is essentially skew-adjoint (its closure is skew-adjoint).

Definition 7.21. Let V be a vector space and $\mathfrak{g}$ a Lie-algebra. A representation (φ, V) of $\mathfrak{g}$ is a linear mapping $\varphi : \mathfrak{g} \to End(H)$ such that $\varphi([A, B]) = [\varphi(A), \varphi(B)]$.

Let $x \in H_\pi^\infty$ and $L \in \mathfrak{g}$, then the mapping $L \mapsto \pi(\exp L)x$ is differentiable since the exponential mapping is differentiable. The mapping $t \mapsto \pi(\exp tL)x$ is then also differentiable which leads us to formulate the following

Proposition 7.22. Let (H, π) be a representation of the Lie-group G. For L in the Lie-algebra $\mathfrak{g}$ the mapping $\varphi(L) : H_\pi^\infty \to H_\pi^\infty$ defined by

$$\varphi(L)x = \lim_{t\to 0} \frac{\pi(\exp tL)x - x}{t} \tag{7.8}$$

is an automorphism of H_π^∞. (H_π^∞, φ) is a representation of $\mathfrak{g}$.

The proof is equivalent to the proof of Proposition 3.9 in [Kna86].

Proof. Since $t \mapsto \pi(\exp tL)x$ is differentiable it is also differentiable in $t = 0$ so the limit exists in H. If $x, y \in H_\pi^\infty$ and c_1, c_2 are constants then it follows that

$$\varphi(L)(c_1x + c_2y) = \lim_{t\to 0} \frac{\pi(\exp tL)(c_1x + c_2y) - (c_1x + c_2y)}{t} \tag{7.9}$$

$$= c_1 \lim_{t\to 0} \frac{\pi(\exp tL)x - x}{t} + c_2 \lim_{t\to 0} \frac{\pi(exptL)y - y}{t} \tag{7.10}$$

$$= c_1\varphi(L)x + c_2\varphi(L)y \tag{7.11}$$

proving that $\varphi(L)$ is linear from H_π^∞ to H. Let $x \in H_\pi^\infty$ and define $f(g) = \pi(g)x$. Applying $\pi(g)$ to (7.8) gives

$$\pi(g)\varphi(L)x = \lim_{t\to 0} \frac{\pi(g \exp tL)x - \pi(g)x}{t} = L(f)(g). \tag{7.12}$$

Since f is smooth so is $L(f) : g \to L(f)(g)$ and therefore $g \mapsto \pi(g)\varphi(L)x$ is smooth, such that $\varphi(L)x$ is in H_π^∞. This shows that $\varphi(L)$ is an automorphism of H_π^∞ for each $L \in \mathfrak{g}$.

To show that φ is a representation we need to show that $L \mapsto \varphi(L)x$ is linear and $\varphi([L_1, L_2])x = \varphi(L_1)\varphi(L_2)x - \varphi(L_2)\varphi(L_1)x$ for all $x \in H_\pi^\infty$.

To show that $L \mapsto \varphi(L)x$ is linear we make the following calculations

$$\varphi(L_1 + L_2)x = \lim_{t\to 0} \frac{\pi(\exp t(L_1 + L_2))x - x}{t} \tag{7.13}$$

$$= \lim_{t\to 0} \frac{\pi(\exp t(L_1 + L_2)x - \pi(\exp tL_2)x}{t} + \lim_{t\to 0} \frac{\pi(\exp t(L_2)x - x}{t} \tag{7.14}$$

$$= \lim_{t\to 0} \frac{\pi(\exp tL_1)\pi(\exp tL_2)x - \pi(\exp tL_2)x}{t} + \lim_{t\to 0} \frac{\pi(\exp t(L_2)x - x}{t} \tag{7.15}$$

$$= \varphi(L_1)x + \varphi(L_2)x \tag{7.16}$$

and

$$\varphi(cL)x = \lim_{t\to 0} \frac{\pi(\exp t(cL))x - x}{t} = \lim_{ct\to 0} \frac{\pi(\exp ctL)x - x}{t/c} = c\varphi(L). \tag{7.17}$$

The mapping $(s, t) \mapsto \pi(\exp sL_1 \exp tL_2)x$ for $x \in H_\pi^\infty$ is smooth since it is composed of the smooth functions

$$(s, t) \mapsto (\exp sL_1, \exp tL_2), (g_1, g_2) \mapsto g_1g_2 \text{and} g \mapsto \pi(g)x.$$

Therefore the map

$$(s, t) \mapsto \langle \pi(\exp sL_1 \exp tL_2)x, y\rangle$$

is smooth for any $y \in H$. From calculus classes we know that $\frac{\partial^2}{\partial s \partial t} = \frac{\partial^2}{\partial t \partial s}$ in this case. This is proved using the mean-value theorem to find the limit as a diagonal limit. This tecnique can be applied here to get

$$
\begin{aligned}
\langle \varphi(L_1)\varphi(L_2)x, y\rangle &= \frac{\partial^2}{\partial s \partial t}\langle \pi(\exp sL_1 \exp tL_2)x, y\rangle \\
&= \lim_{t\to 0} t^{-2}\langle \pi(\exp tL_1)\pi(\exp tL_2)x - \pi(\exp tL_1)x - \pi(\exp tL_2)x + x, y\rangle
\end{aligned}
$$

The same calculations hold for $\varphi(L_2)\varphi(L_1)x$ and subtracting gives

$$
\begin{aligned}
&\varphi(L_1)\varphi(L_2)x - \varphi(L_2)\varphi(L_1)x \\
&\quad= \lim_{t\to 0}\frac{\pi(\exp tL_1)\pi(\exp tL_2)x - \pi(\exp tL_2)\pi(\exp tL_1)x}{t^2} \\
&\quad= \lim_{t\to 0}\pi(\exp tL_2 \exp tL_1)\frac{\pi(\exp(-tL_1)\exp(-tL_2)\exp(tL_1)\exp(tL_2))x - x}{t^2} \\
&\quad= \lim_{t\to 0}\pi(\exp tL_2 \exp tL_1)\frac{\pi(\exp(t^2[L_1, L_2] + O(t^3)))x - x}{t^2} \\
&\quad= \varphi([L_1, L_2]).
\end{aligned}
$$

The second last equality follows from [Hel78, Lemma 1.8(ii), p. 106]. □

The following is the same as Proposition 3.10 in [Kna86].

Proposition 7.23. Let (φ, H_π^∞) be the associated representation of $\mathfrak{g}$. Then each $\varphi(L)$ for $L \in \mathfrak{g}$ is skew-adjoint.

Proof. If x, y are in H_π^∞ then

$$
\begin{aligned}
\langle x, t^{-1}(\pi(\exp tL)y - y)\rangle &= \langle t^{-1}x, \pi(\exp tL)y\rangle - \langle t^{-1}x, y\rangle && (7.18)\\
&= \langle t^{-1}\pi(\exp(-tL))x, y\rangle - \langle t^{-1}x, y\rangle && (7.19)\\
&= -\langle (-t)^{-1}(\pi(\exp(-tL))x - x), y\rangle && (7.20)
\end{aligned}
$$

As $t \to 0$ it then follows that $\langle x, \varphi(L)y\rangle = -\langle \varphi(L)x, y\rangle$ which shows that $\varphi(L)$ is skew-adjoint. □

Since $\varphi(L)$ is skew-adjoint and densely defined by corollary 7.20 it is closable ([Ped89, 5.1.6]). I will now set out to show that the closure of $\varphi(L)$ is defined by the limit 7.8.

The results that follow are all inspired by [Seg51, p. 233ff]. The only difference is that I use left invariant vectorfields, instead of right invariant vectorfields.

Definition 7.24. For $f : G \to \mathbb{C}$ the adjoint of f is the function f^* defined by $f^*(g) = \overline{f(g^{-1})}\Delta(g)$.

Since Δ is smooth and $g \mapsto g^{-1}$ is smooth it is easily seen that if f is in $C_c^\infty(G)$ then so is f^*.

Lemma 7.25. For $f \in C_c^\infty(G)$ and $L \in \mathfrak{g}$ the adjoint of $\pi(L(f))$ is $\pi(R(f^*)) + \gamma\pi(f^*)$ where R corresponds to L.

Proof. Since $\pi(L(f))$, $\pi(R(f^*))$ and $\pi(f^*)$ are all continuous $(L(f), R(f^*)$ and f^* are smooth functions with compact support) I need only show that

$$\langle \pi(L(f))x, y\rangle = \langle x, \pi(R(f^*))y\rangle + \gamma\langle x, \pi(f^*)\rangle$$

for x, y in the Gårding space. The propositions 7.14 and 7.23 give

$$\langle \pi(L(f))x, y\rangle = \langle \varphi(L)\pi(f)x, y\rangle = \langle \pi(f)x, -\varphi(L)y\rangle \tag{7.21}$$

The following calculation shows that $\pi(f^*)$ is the adjoint of $\pi(f)$

$$\begin{aligned}
\langle \pi(f)x, y\rangle &= \int_G \langle f(a)\pi(a^{-1}x, y\rangle da && (7.22)\\
&= \int_G \langle x, \overline{f(a)}\pi(a)y\rangle da && (7.23)\\
&= \int_G \langle x, \overline{f(a^{-1})}\Delta(a)\pi(a^{-1})y\rangle da && (7.24)\\
&= \int_G \langle x, f^*(a)\pi(a^{-1})y\rangle da && (7.25)\\
&= \overline{\int_G \overline{\langle f^*(a)\pi(a^{-1})y, x\rangle} da} && (7.26)\\
&= \langle x, \pi(f^*)y\rangle. && (7.27)
\end{aligned}$$

The second equality is due to $\pi(g^{-1})^* = \pi(g)$ because π is unitary. Therefore

$$\langle \pi(f)x, -\varphi(L)y\rangle = \langle x, -\pi(f^*)\varphi(L)y\rangle = \langle x, \pi(R(f^*))y + \gamma\pi(f^*)y\rangle \tag{7.28}$$

by Proposition 7.14. □

Lemma 7.26. There exists a sequence $f_n \in C_c^\infty(G)$ of real non-negative functions and corresponding neighborhoods U_n of e such that $\lim \|f_n\|_1 = 1$, $\operatorname{supp} f_n \subseteq U_n$ and $\cap_n U_n = \{e\}$. For any $L \in \mathfrak{g}$ and corresponding right-invariant vectorfield R it holds that $\|L(f_n) + R(f_n)\|_1$ is bounded for $n \to \infty$.

Proof. I have already constructed a sequence that can be used, but I first need to show a few more properties. As in Remark 7.18 define

$$\phi_n(x) = \begin{cases} A^{-1}ne^{-(1-n^2x^2)^{-1}} & |x| < n^{-1} \\ 0 & |x| \geq n^{-1}. \end{cases}$$

Then for $\|x\| < 1/n$

$$\phi_n'(x) = \frac{-2n^2x}{(1-n^2x^2)^2}A^{-1}ne^{-(1-n^2x^2)^{-1}},$$

which is positive for $x \leq 0$ and negative for $x \geq 0$. Therefore

$$\begin{aligned}
\|\phi_n'\|_1 &= \int_0^{1/n} |\phi_n'(x)| dx \\
&= \int_{-1/n}^{0} -\phi_n'(x) dx + \int_{-1/n}^{1/n} \phi_n'(x) dx \\
&= -[\phi_n(x)]_0^{1/n} + [\phi_n(x)]_{-1/n}^{0} = 2(\phi_n(0) - \phi_n(1/n)) \\
&= 2\phi_n(0) = 2n/e = O(n).
\end{aligned}$$

If we let f_n be the sequence from (7.7) it then holds that

$$\left\| \frac{\partial f_n}{\partial x_i} \right\|_1 = O(n). \tag{7.29}$$

Let L_i be a basis for $\mathfrak{g}$ and let x_i be canonical coordinates in a neighbourhood N_e where $\exp : N_0 \to N_e$ is a diffeomorphism. According to (7.2) for f with support inside N_e it holds that

$$\int f(a) da = \int_{N_0} k(L) f(\exp L) dL.$$

where k is smooth and $k(0) \neq 0$.

There is an m such that for $n \geq m$ the mapping $a \mapsto \phi_n(x_i(a))$ will have support inside N_e for all i (since it can be regarded as a mapping from canonical coordinates which are in $\mathbb{R}$ to $\mathbb{R}$). For $n \geq m$ and $a \in N_e$ define

$$f_n(a) = k(0)^{-1} \prod_i \phi_n(x_i(a)).$$

Then $\lim_{n\to\infty} \|f_n\|_1 = 1$ and f_n is smooth with compact support.

By Proposition 7.9 we get that

$$\begin{aligned}
L_a(f) + R_a(f) &= \frac{d}{dt}\Big|_{t=0} f(a \exp(tL)) - \frac{d}{dt}\Big|_{t=0} f(\exp(tL)a) \\
&= \lim_{t\to 0} t^{-1}(f(a\exp(tL)) - f(a) - (f(\exp(tL)a) - f(a))) \\
&= \lim_{t\to 0} t^{-1}(f(a\exp(tL)) - f(\exp(tL)a))
\end{aligned}$$

Using the mean-value theorem (see [Pri84, Thm. 9.1]) we get that there is a $b \in \{a\exp(rL) | |r| < |s|\}$ such that

$$f(a\exp(tL)) = f(a) + \sum_i \frac{\partial f}{\partial x_i}\Big|_b (x_i(a\exp(sL)) - x_i(a))$$

Also there is a $c \in \{\exp(sL) | |s| < |t|\}$ such that

$$f(\exp(tL)a) = f(a) + \sum_i \frac{\partial f}{\partial x_i}\Big|_c (x_i(\exp(tL)a) - x_i(a))$$

When t tends to 0, b and c tend to a. Therefore we get

$$\begin{aligned} L_a(f) + R_a(f) &= \lim_{t\to 0} t^{-1}(f(a\exp(tL)) - f(\exp(tL)a)) \\ &= \sum_i \frac{\partial f}{\partial x_i}\Big|_a \lim_{t\to 0} t^{-1}(x_i(a\exp(tL)) - x_i(\exp(tL)a)). \end{aligned} \tag{7.30}$$

The function $g(a,t) = x_i(a\exp(tL)) - x_i(\exp(tL)a)$ is smooth. And

$$\begin{aligned} \frac{\partial g}{\partial t}\Big|_{t=0}(a) &= \lim_{t\to 0} t^{-1}(x_i(a\exp(tL)) - x_i(\exp(tL)a) - (x_i(a) - x_i(a))) \\ &= \lim_{t\to 0} t^{-1}(x_i(a\exp(tL)) - x_i(\exp(tL)a)) \end{aligned}$$

For a close enough to e write $a = \exp(\sum_i a_i L_i)$ and define $\|a\| = \max_i\{a_i\}$. Using the mean-value theorem on $\frac{\partial g}{\partial t}|_{t=0}$ we get that there exists an a' with $\|a'\| \leq \|a\|$ such that

$$\frac{\partial g}{\partial t}\Big|_{t=0}(a) = \sum_i a_i \frac{\partial}{\partial x_i}\Big|_{a'}\Big(\frac{\partial g}{\partial t}\Big|_{t=0}\Big)$$

since $\frac{\partial g}{\partial t}|_{t=0}(0) = 0$. The partial derivatives of $\frac{\partial g}{\partial t}|_{t=0}$ are bounded ($\{a' | \|a'\| \leq \|a\|$ is compact and the derivatives are continuous) so it $\frac{\partial g}{\partial t}|_{t=0} = O(\|a\|)$. From this it follows that

$$\lim_{t\to 0} t^{-1}(x_i(a\exp(tL)) - x_i(\exp(tL)a)) = \frac{\partial g}{\partial t}\Big|_{t=0}(a) = O(\|a\|).$$

Since f_n is zero unless $\|a\| < n^{-1}$ and $\|\frac{\partial f_n}{dx_i}\|_1 = O(n)$ by (7.29), the expression (7.30) gives

$$\|L(f_n) + R(f_n)\|_1 \leq \sum_i \|\frac{\partial f}{\partial x_i}\|_1 O(\|a\|) = \sum_i O(n)O(n^{-1}) = O(1).$$

So $L(f_n) + R(f_n)$ is bounded. □

Lemma 7.27. Let f_n be as in Lemma 7.26. For any $L \in \mathfrak{g}$ and the corresponding right invariant vectorfield R the strong convergence $\pi(L(f_n)) + \pi(R(f_n)) + \gamma\pi(f_n) \to 0$ holds.

Proof. Using $\phi(a) = \pi(a^{-1})x$ we get that

$$\pi(Lf_n) + \pi(Rf_n) + \gamma\pi(f_n) = \int_G (Lf_n + Rf_n + \gamma f_n)(a)\phi(a)x da.$$

$\phi : G \to H$ is a bounded continuous function so it is enough to show that

$$\lim_{n\to\infty} \int_G (Lf_n + Rf_n + \gamma f_n)(a) da \to 0 \tag{7.31}$$

because the preceeding lemma gives that $\|Lf_n + Rf_n + \gamma f_n\|_1$ is bounded.

$$\int_G L_a(f)da = \int_G \frac{d}{dt}|_{t=0} f(a\exp(tL))da = \lim_{t\to 0} t^{-1} \int_G f(a\exp(tL)) - f(a)da$$

Since the Haar-measure is right invariant the integral is 0.

$$\begin{aligned}\int_G Rf_n(a)da &= -\lim_{t\to 0} t^{-1}\int_G f_n(\exp(tL)a) - f_n(a)da \\ &= -\lim_{t\to 0} t^{-1}(\Delta(\exp tL)-1)\int_G f_n(a)da \\ &= -\gamma\int_G f_n(a)da\end{aligned}$$

From this follows the convergence (7.31). □

Lemma 7.28. $\varphi(L)^{**} \supseteq -\varphi(L)^*$ for all $L \in \mathfrak{g}$.

Proof. I want to show that for all $x \in \mathcal{D}(\varphi(L)^*)$ and $y \in \mathcal{D}(\varphi(L)^*)$ the following holds

$$(\varphi(L)^*x, y) = (x, \varphi(L)^*y)$$

This shows that $y \in \mathcal{D}(\varphi(L)^{**})$ and that $\varphi(L)^{**}y = \varphi(L)^*y$. It is seen by expanding $(\varphi(L)^*x, y) - (x, \varphi(L)^*y)$ in the following way

$$\begin{aligned}\langle\varphi(L)^*x, y) - \langle x, \varphi(L)^*y\rangle = \langle\varphi(L)^*x, y) &- \langle\varphi(L)^*x, \pi(g^*)y\rangle &(7.32)\\ &+ \langle\varphi(L)^*x, \pi(g^*)y\rangle - \langle x, \varphi(L)\pi(g^*)y\rangle &(7.33)\\ &+ \langle x, \varphi(L)\pi(g^*)y\rangle - \langle x, \pi(Lg^*)y\rangle &(7.34)\\ &+ \langle x, \pi(Lg^*)y\rangle + \langle x, \pi(Rg^* + \gamma g^*)y\rangle &(7.35)\\ &- \langle x, \pi(Rg^* + \gamma g^*)y\rangle + \langle\pi(Lg)x, y\rangle &(7.36)\\ &- \langle\pi(Lg)x, y\rangle + \langle\varphi(L)\pi(g)x, y\rangle &(7.37)\\ &- \langle\varphi(L)\pi(g)x, y\rangle - \langle\pi(g)x, \varphi(L)^*y\rangle &(7.38)\\ &+ \langle\pi(g)x, \varphi(L)^*y\rangle - \langle x, \varphi(L)^*y\rangle. &(7.39)\end{aligned}$$

The terms in (7.33) and (7.38) are zero since $\varphi(L)$ is skew-adjoint. Lemma 7.25 and Proposition 7.14 tell us that the terms (7.36), (7.34) and (7.37) are zero. Setting $g = f_n^*$ and letting $n \to \infty$ the terms (7.32), (7.35) and (7.39) tend to zero by Lemma 7.27. □

Since $\varphi(L)$ is a densely defined skew-adjoint operator it is closable and the following holds

Theorem 7.29. $\varphi(L)$ is essentially skew-adjoint for all $L \in \mathfrak{g}$.

Proof. Let $T = \varphi(L)$ on H_π^∞ then $T = -T^*$. T is densely defined so according to [Ped89, thm 5.1.5] T^* is a closed operator. So $\bar{T} \subseteq -T^*$. By the same theorem T^{**} is closed and equal to $\bar{T}$, since T^* is densely defined. Then $(\bar{T})^* \supseteq -T^{**} = -\bar{T}$. The preceeding lemma gives $\bar{T} = T^{**} \supseteq -T^*$, which readily gives $(\bar{T})^* \subseteq -T^{**} = -\bar{T}$. This shows that $\bar{T}$ is skew-adjoint and thus that $\varphi(L)$ is essentially skew-adjoint. □

Proposition 7.30. Defining the mapping π, from $\mathfrak{g}$ to the skew-adjoint operators on H, such that $\pi(L)$ is the closure of $\varphi(L)$, then for any $L \in \mathfrak{g}$, $\pi \circ \exp(L) = \exp \circ \pi(L)$.

Proof. Define the strongly continuous one-parameter unitary groups $V_t = \pi(\exp(tL))$ and $W_t = \exp(t\pi(L))$. Then for $y \in H_\pi^\infty$, $\lim_{t\to 0} t^{-1}(V_t y - y) = \varphi(L)y = \pi(L)y$ by the definition of $\varphi(L)$ given by (7.8). Also Theorem 6.7 with Remark 6.8 on page 54 gives

$$\pi(L)y = \lim_{t\to 0} \frac{W_t y - y}{t}$$

for $y \in H_\pi^\infty$. Stone's theorem (Theorem 6.9) tells us that there exists self-adjoint operators X, Y such that $V_t = \exp(itX)$ and $W_t = \exp(itY)$. The previous remarks ensure that H_π^∞ is included in $\mathcal{D}(X)$ and $\mathcal{D}(Y)$ and that they are equal on this set. Theorem 6.7 then gives

$$iXy = \lim_{t\to 0} \frac{V_t y - y}{t} = \lim_{t\to 0} \frac{\pi(\exp(tL))y - y}{t} = \pi(L)y$$

and similarly $iYy = \pi(L)y$ for $y \in H_\pi^\infty$. So $\pi(L) \subseteq iX$ and $\pi(L) \subseteq iY$. Since $\pi(L)$, iX and iY are skew-adjoint it follows that $-\pi(L) = \pi(L)^* \supseteq (iX)^* = -iX$ and $-\pi(L) \supseteq -iY$. Therefore $-iX = \pi(L) = -iY$ which then gives $V_t = W_t$. $\square$

The above tells us that $\pi(L)$ is given by the limit (7.8) used to define $\varphi(L)$ for all x for which the limit exists.

7.6 The Uncertainty Principle

In [Kra67] Kraus stated an uncertainty principle concerning operators which arise from Lie groups. Folland and Sitaram [FS97, Thm 2.4] later generalized this theorem to the following

Theorem 7.31. Let G be a connected Lie group with Lie algebra $\mathfrak{g}$, and let (H, π) be a unitary representation of G. Suppose that $L_1, L_2 \in \mathfrak{g}$ and that the linear span $\mathfrak{I}$ of L_1, L_2 and $[L_1, L_2]$ is an ideal in $\mathfrak{g}$. Then

$$\|Ax\|\|Bx\| \geq \frac{1}{2}|\langle Cx, x\rangle| \qquad \text{for all } x \in \mathcal{D}(A) \cap \mathcal{D}(B) \cap \mathcal{D}(C) \tag{7.40}$$

holds with $A = \pi(L_1)$,$B = \pi(L_2)$ and $C = \pi([L_1, L_2])$.

Folland and Sitaram discuss the necessity of $\mathfrak{I}$ to be and ideal in theorem 7.31. Here I will show that this is not necessary.

Lemma 7.32. With functions $f_n \in C_c^\infty(G)$ as in Lemma 7.26 it holds that $\pi(L)\pi(f_n)u \to \pi(L)u$ for $n \to \infty$ for all left-invariant vector fields $L \in \mathfrak{g}$ and all $u \in D(\pi(L))$.

Proof. Fix $L \in \mathfrak{g}$. Let $\gamma = \frac{d}{dt}\Delta(exp(tL))|_{t=0}$ where Δ is the modular function for the chosen left Haar-measure. Also define $f^*(g) = \overline{f(a^{-1})}\Delta(a)$ then $\pi(f)^* = \pi(f^*)$ and $f^{**} = f$. Let R be the right-invariant vector field corresponding to L. Note that for all $f \in C_c^\infty(G)$ it holds that $\pi(f)\pi(L) \subseteq -\pi(Lf^*)^*$, since $\langle \pi(f)\pi(L)u, v\rangle = \langle u, -\pi(Lf^*)v\rangle$ for all $u \in D(\pi(L))$ and $v \in H$. Hence

$$\pi(f)\pi(L) \subseteq -\pi(Rf) - \gamma\pi(f)$$

by Lemma 7.25. But then

$$\pi(L)\pi(f_n)u - \pi(f_n)\pi(L)u = \pi(Lf_n)u + \pi(Rf_n)u + \gamma\pi(f_n)u$$

for any $u \in D(\pi(L))$. By Lemma 7.27 the above tends to 0 and since $\pi(f_n)\pi(L)u \to \pi(L)u$ it follows that $\pi(L)\pi(f_n)u \to \pi(L)u$. $\square$

Given two left-invariant vectorfields L_1 and L_2, then for any $u \in D(\pi(L_1)) \cap D(\pi(L_2)) \cap D(\pi([L_1, L_2]))$ it follows that $\pi(f_n)u \to u$ by Lemma 7.17. By the lemma above it also holds that $\pi(L_1)\pi(f_n)u \to \pi(L_1)u$, $\pi(L_2)\pi(f_n)u \to \pi(L_2)u$ and $\pi([L_1, L_2])\pi(f_n)u \to \pi([L_1, L_2])u$. Since the Gårding vector $\pi(f_n)u$ is in $D([\pi(L_1), \pi(L_2)])$ by corollary 7.15 then Proposition 6.1 gives

Theorem 7.33. Let G be a Lie group with Lie algebra $\mathfrak{g}$, and let (H, π) be a unitary representation of G. Suppose that $L_1, L_2 \in \mathfrak{g}$. Then

$$\|Ax\|\|Bx\| \geq \frac{1}{2}|\langle Cx, x\rangle| \qquad \text{for all } x \in D(A) \cap D(B) \cap D(C)$$

holds with $A = \pi(L_1)$,$B = \pi(L_2)$ and $C = \pi([L_1, L_2])$.

Example 7.34 (The $(\mathbf{ax + b})$-group). Let G be the the following two-dimensional Lie subgroup of $GL(2, \mathbb{R})$:

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \middle| a > 0, y \in \mathbb{R} \right\}$$

Using (a, b) to denote such an element in G, it is seen that $(a, b)(c, d) = (ac, ad + b)$. A representation of this group on $L_2(\mathbb{R})$ is then defined by $\pi(a, b)f(x) = e^{ibe^x} f(x + \log a)$. To check that it is a representation we calculate

$$\begin{aligned}\pi(a, b)\pi(c, d)f(x) &= \pi(a, b)e^{ide^x} f(x + \log c) = e^{ibe^x} e^{ide^{x+\log a}} f(x + \log a + \log c)\\ &= e^{i(ad+b)e^x} f(x + \log ac) = pi((a, b)(c, d))f(x).\end{aligned}$$

By a change of variable

$$\|\pi(a, b)f\|_2^2 = \int_{\mathbb{R}} |f(x + \log a)|^2 dx = \|f\|_2^2,$$

which shows that the representation is unitary.

The Lie algebra of G is spanned by

$$L_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad L_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

with $[L_1, L_2] = L_2$. By direct calculation of the exponential map for matrices it follows that

$$\exp(tL_1) = \begin{pmatrix} e^t & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \exp(tL_2) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

Assume $f \in \mathcal{D}(\pi(L_1)) \cap \mathcal{D}(\pi(L_2))$ then

$$\pi(L_1)f(x) = \lim_{t\to 0} \frac{\pi(exp(tL_1))f(x) - f(x)}{t} = \lim_{t\to 0} \frac{f(x+t) - f(x)}{t} = f'(x),$$

and

$$\pi(L_2)f(x) = \lim_{t\to 0} \frac{\pi(exp(tL_2))f(x) - f(x)}{t} = \lim_{t\to 0} \frac{e^{ite^x} f(x) - f(x)}{t} = f(x)(e^{ite^x})' = ie^x f(x).$$

This shows that f has to be differentiable with $f'(x)$ and $e^x f(x)$ both in $L_2(\mathbb{R})$. The theorem then gives the following inequality

$$\|f'\|_2 \|e^x f\|_2 \geq |\langle e^x f, f\rangle|.$$

for functions f with $f'(x)$ and $e^x f(x)$ both in $L_2(\mathbb{R})$.

Example 7.35 (Heisenberg's uncertainty principle). Let G be the Heisenberg-Weyl group of upper triangular matrices

$$G = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \,\middle|\, (x, y, z) \in \mathbb{R}^3 \right\}.$$

Multiplication of two element can be written as $(x_2, y_2, z_2)(x_1, y_1, z_1) = (x_2 + x_1, y_2 + y_1, z_2 + z_1 + y_1 x_2)$. The following defines a representation of G on $L_2(\mathbb{R})$: $\pi(x, y, z)f(s) = e^i(z + sy)f(s + x)$. This is verified by

$$\begin{aligned} \pi(x_2, y_2, z_2)\pi(x_1, y_1, z_1)f(s) &= \pi(x_2, y_2, z_2)e^{i(z_1 + ty_1)} f(t + x_1) \\ &= e^{i(z_2 + ty_2)} e^{i(z_1 + (t + x_1)y_1)} f(t + x_2 + x_1) \\ &= e^{i(z_2 + z_1 + y_1 x_2 + t(y_2 + y_1))} f(t + x_2 + x_1) \\ &= \pi((x_2, y_2, z_2)(x_1, y_1, z_1))f(t) \end{aligned}$$

The representation is unitary.

The following vectors span $\mathfrak{g}$ (which is no longer necessary)

$$L_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, L_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad L_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

By direct calculation we get

$$\exp(tL_1) = \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \exp(tL_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \exp(tL_3) = \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

By using the same method as before it follows that

$$\pi(L_1)f(s) = f'(s), \pi(L_2)f(s) = isf(s) \quad \text{and} \quad \pi(L_3)f(s) = if(s).$$

Here it is important to note that $\mathcal{D}(\pi(L_3)) = L_2(\mathbb{R})$, so for all $f \in \mathcal{D}(\pi(L_1)) \cap \mathcal{D}(\pi(L_2))$ the uncertainty principle holds. Thus we have just obtained Heisenberg's inequality

$$\|f'\|_2 \|xf\|_2 \geq \|f\|_2^2,$$

for functions f with $f'(x)$ and $xf(x)$ both in $L_2(\mathbb{R})$. The domain for which the uncertainty principle holds is bigger than that of Example 6.2.

Part IV

Applications of Uncertainty Principles

Chapter 8

Signal Analysis

In this chapter I will explore some strengths and weaknesses of uncertainty principles applied to signal analysis. Uncertainty principles normally tell us what does not hold or cannot be done, but there are in fact constructive applications of uncertainty principles.

8.1 Uncertainty Principles

The classical Heisenberg uncertainty principle tells us that if we observe a signal only for a finite period of time, we will loose information about the frequencies the signal consists of.

Let f represent the (real valued) amplitude of a signal with $\|f\|_2 = 1$. Then $\hat{f}(-y) = \overline{\hat{f}(y)}$ so $|\hat{f}|$ is an even function. If the signal is localised in frequency around y_0, then it is also localised around $-y_0$ (see figure 5.1 on page 46). So $|\hat{f}|$ has most of its "mass" concentrated around y_0 and $-y_0$. If y_0 is large, then the variance is large. In this case Heisenberg's uncertainty principle (Theorem 3.2) does not tell us much about a lower bound on the variance of f.

I will now try to apply Theorem 5.2 to the same problem. Assume we have the same distance $b = y_0 + \delta$ between all points of measure. Then we can adjust δ to "miss" the peaks in $\pm y_0$ slightly. If the peaks are very sharp, then the sum in (5.2) will be small and thus the variance of f will be big. The last application shows that $|\hat{f}|$ cannot be sharply concentrated in two points. If $|\hat{f}|$ is concentrated in countable many points, we can adjust δ so we miss all these points slightly. Then the sum (5.2) is small, which shows that the variance of f has a lower bound. Thus if f is concentrated around one point $\hat{f}$ cannot be concentrated around countably many "evenly" distributed points. This is more information than Heisenberg's uncertainty principle could give.

There are other ways to handle this problem. Since $|\hat{f}|$ is even we can restrict ourselves to the variance defined by

$$V_+(|\hat{f}|) = \inf_{b>0} \int (y-b)^2 |\hat{f}(y)|^2 dy.$$

There have been proved uncertainty principles that give a lower bound on the product of $V_+(|\hat{f}|)$ and the normal variance of f. Other principles can be proved and have been referred to on p. 216 in [FS97].

There is also a problem with applying the principle in Theorem 4.17. We are able to produce bandlimited signals of a given length, which seems to be a contradiction. But we must remember that the Fourier transform is a mathematical tool, and that nature is not governed by it.

8.2 Recovering Missing Segments of a Bandlimited Signal

This is taken from section 4 in [DS89]. It shows a way to constructively use uncertainty principles. The uncertainty principle in use is found in section 4.4 of this thesis.

Let $s(t) \in L_2$ be a signal containing only frequencies in a measurable set W (this is called a bandlimited signal). Assume that the reciever is unable to observe the signal on a measurable set T. Also assume that the recieved signal $r(t)$ is contaminated with noise $n(t) \in L_2$ when $t \notin T$. When $t \in T$ nothing is observed so $n(t) = 0$ on T. The received signal is

$$r(t) = \begin{cases} s(t) + n(t), & t \notin T \\ 0 & t \in T. \end{cases} \tag{8.1}$$

Theorem 8.1. If W and T are measurable sets with $|T||W| < 1$ then there exists an operator Q and a constant $C \leq (1 + \sqrt{|T||W|})^{-1}$ such that

$$\|s - Qr\|_2 \leq C\|n\|_2. \tag{8.2}$$

Proof. Let Q be the following bounded operator $Q = (I - P_T P_W)^{-1}$. This operator exists since $\|P_T P_W\| \leq \sqrt{|W||T|} < 1$ and also $Q = \sum_{j=0}^{\infty} (P_T P_W)^j$ by [Ped89, Lemma 4.1.7]. (8.1) is equivalent to

$$r = (I - P_T)s + n \tag{8.3}$$

where P_T is the time-limiting operator defined in section 4.4. Since s is bandlimited to W it holds that $P_T P_W s = P_T s$ and using (8.3) gives

$$\begin{aligned} s - Qr &= s - Q(I - P_T)s - Qn \\ &= s - (I - P_T P_W)^{-1}(I - P_T P_W)s - Qn \\ &= -Qn. \end{aligned}$$

The norm of Q can be evaluated as follows

$$\|Q\| \leq \sum_{j=0}^{\infty} \|P_T P_W\|^j = (1 - \|P_T P_W\|)^{-1}. \tag{8.4}$$

The lemmas from section 4.4 give $\|P_T P_W\| \leq \sqrt{|W||T|}$ and so

$$\|s - Qr\|_2 = \|Qn\|_2 \leq \|Q\|\|n\|_2 \leq (1 - \sqrt{|W||T|})^{-1}\|n\|_2,$$

which ends the proof. $\square$

This shows that s can be reconstructed from r to a certain degree. In general we cannot get rid of the noise, but assume that $n = 0$ everywhere, then (8.2) shows that we can fully reconstruct s from r if $|W||T| < 1$.

Chapter 9

Quantum Mechanics

The first time I heard about uncertainty principles was in a course on quantum mechanics. This chapter will introduce the postulates of quantum mechanics and show how Heisenberg's uncertainty principle applies. I will also give an example which shows me must take care when applying uncertainty principles.

9.1 The Postulates of Quantum Mechanics in One Dimension

The following information is extracted from [Sch81] and I will not prove anything here. It is important to notice, that the postulates are the foundation of quantum mechanics, and cannot be proved but only verified experimentally.

Consider at particle with restricted motion along a line.

Postulate 1. There is a function $\psi(x,t)$ of position x and time t such that the probability of the particle to be in the interval I at time t is given by

$$\int_I |\psi(x,t)|^2 dx.$$

Since the particle must be somewhere on the line

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

So now we see that the Hilbert space $H = L_2(\mathbb{R})$ is of great importance.

Theorem 9.1. If $\psi(x,t)$ is continuous with respect to x and for some t

$$\int_{\mathbb{R}} |x||\psi(x,t)|^2 dx < \infty$$

then the mathematical expectation of position is given by

$$\bar{x} = \int_{\mathbb{R}} x|\psi(x,t)|^2 dx.$$

This can also be written $\bar{x} = \langle X\psi, \psi\rangle$ where X is the operator $X(\psi(x)) = x\psi(x)$. The inner product exists for all $\psi \in C_c^\infty$ so $\mathcal{D}(X)$ is dense.

Postulate 2. The probability that the momentum p of the particle is contained in the interval I is

$$\frac{1}{h}\int_I |\hat{\psi}(p/h,t)|^2 dp,$$

where h is Plancks constant and $\hat{\psi}$ is the Fourier transform with respect to x.

Note that $\|\hat{\psi}\|_2 = \|\psi\|_2 = 1$ so $\hat{\psi}$ is also a probability distribution.

Theorem 9.2. If $\hat{\psi}(p/h,t)$ is continuous with respect to p and for some t

$$\int_{\mathbb{R}} |p||\hat{\psi}(p/h,t)|^2 dp < \infty$$

then the mathematical expectation of momentum is given by

$$\bar{p} = \frac{1}{h}\int_{\mathbb{R}} p|\hat{\psi}(p/h,t)|^2 dp.$$

Let us rewrite $\bar{p}$

$$\begin{aligned}
\bar{p} &= \frac{1}{h}\int_{\mathbb{R}} p|\hat{\psi}(p/h,t)|^2 dp \\
&= h\int_R p|\hat{\psi}(p,t)|^2 dp \\
&= -ih\int_R \widehat{\frac{\partial\psi}{\partial x}}\hat{\bar{\psi}} dp \\
&= -ih\int_R \frac{\partial\psi}{\partial x}\bar{\psi} dp.
\end{aligned}$$

The third equality is accoring to the definition of a distributional derivative, and the last is Parsevals identity (Theorem 1.3).

We now see that the momentum corresponds to the operator $P = -ih\frac{\partial}{\partial x}$ and that $\bar{p} = \langle P\psi, \psi\rangle$ if the right hand side exists. The inner product exists for $\psi \in C_c^\infty$ so $D(P)$ is dense.

The operators for position and momentum both satisfy

Postulate 3. To every observable a there corresponds a self-adjoint operator A with dense domain such that

$$\bar{a} = \langle A\psi, \psi\rangle.$$

That A is self-adjoint ensures that $\langle A\psi, \psi\rangle$ is a real number for all $\psi \in \mathcal{D}(A)$.

9.2 The Heisenberg Uncertainty Principle in One Dimension

A good measure for how much an observable a (equivalent to an operator A) deviates from its mean value is the variance

$$\sigma^2(a) = \int_{\mathbb{R}} (x - \bar{a})^2 |\psi(x)|^2 dx$$

if it exists. This is the mean value of $(x - \bar{a})^2$ and for the operator A we find

$$\sigma^2(A) = \overline{(A - \bar{a})^2} = \langle (A^2 + \bar{a}^2 - 2\bar{a}A)\psi, \psi \rangle = \langle A^2\psi, \psi \rangle - \bar{a}^2.$$

Note that since $\bar{a}$ is real $A - \bar{a}$ is self-adjoint so

$$\sigma^2(A) = \|A - \bar{a}\|_2^2.$$

We have already seen the variance before in section 6.2, where it was denoted $\sigma_A(\psi)$.

From Proposition 6.3 we therefore get that

$$\sigma(A)\sigma(B) \geq \frac{1}{2} |\langle [A, B]\psi, \psi \rangle| \tag{9.1}$$

if $\psi \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$.

9.2.1 Position and Momentum

For the position and momentum operators we then get

$$[X, P]\psi = x\Big(- ih\frac{\partial \psi}{\partial x}\Big) + ih\frac{\partial}{\partial x}(x\psi) = ih\psi$$

so since $\|\psi\|_2 = 1$ we get

$$\sigma(X)\sigma(P) \geq \frac{1}{2} |\langle ih\psi, \psi \rangle| = \frac{h}{2}$$

This is the original form of Heisenberg's uncertainty principle. It is derived from the postulates of quantum mechanics and it is remarkable that the right hand side does not depend on ψ. Assume that we want to measure the position and momentum of several particles in the same state ψ. Then multiplication of the variances of these measurements will be greater than or equal to a given constant. It is not a statement about the accuracy of our instruments for measuring, but a statement about the nature of matter. With perfect intruments (if they do not interact with the particle for example) we would still have this very small uncertainty in our measurements. According to results in section 7.6 this is true for all $\psi \in \mathcal{D}(X) \cap \mathcal{D}(P)$.

The uncertainty principle (9.1) applies to any simultaneous measurement of an observable. The lower bound is only greater than zero if we want to measure observables whose corresponding operators do not commute. Notice that it only holds in general for $\psi \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$.

9.3 The Uncertainty Principle for Angular Momentum and Angle

This section is a summary of an article by Kraus [Kra65]. It shows there is a problem with the domains in the uncertainty principle for measurement of angular momentum and angle. This tells us that it is important to take care, when applying uncertainty principles for operators. A different principle is derived, but here another problem, which I will not explore, arises.

Let $L_z f(\theta) = -if'(\theta)$ be the angular momentum operator defined on differentiable functions in $L_2([-\pi,\pi])$ with $f(-\pi) = f(\pi)$ (for symmetry). Let $\varphi f(\theta) = \theta f(\theta)$ be operator corresponding to measurement of angle. As in example 6.5 these operators do not satisfy the uncertainty principle

$$\sigma(L_z)\sigma(\varphi) \geq \frac{1}{2},$$

for all $f \in \mathcal{D}(L_z) \cap \mathcal{D}(\varphi)$ (but is does hold on $\mathcal{D}([L_z,\varphi])$).

Instead we can make the following estimation

$$(i\langle Au, Bu\rangle - i\langle Bu, Au\rangle)^2 = (2\,\mathrm{Im}\langle Au, Bu\rangle)^2 \leq 4\|Au\|^2\|Bu\|^2,$$

which holds for all $u \in \mathcal{D}(A) \cap \mathcal{D}(B)$. Defining

$$\Phi_{A,B}(u,v) = i\langle Au, Bv\rangle - i\langle Bu, Av\rangle$$

then we get the following new uncertainty principle

$$\sigma(A)\sigma(B) \geq \frac{1}{2}|\Phi_{A,B}|.$$

For angular momentum and angle $\Phi_{L_z,\varphi}$ is

$$\begin{aligned}
\Phi_{L_z,\varphi}(f,g) &= i\int_{-\pi}^{\pi} -if'(x)x\overline{g(x)} - xf(x)\overline{-ig'(x)}dx \\
&= \int_{-\pi}^{\pi} f'(x)x\overline{g(x)} + xf(x)\overline{g'(x)}dx \\
&= [xf(x)\overline{g(x)}]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} -f(x)(x\overline{g'(x)} + \overline{g(x)}) + xf(x)\overline{g'(x)}dx \\
&= 2\pi f(\pi)\overline{g(\pi)} - \langle f, g\rangle.
\end{aligned}$$

We then get the uncertainty relation (with f a normed function)

$$\sigma(L_z)\sigma(\varphi) \geq \frac{1}{2}|1 - 2\pi|f(\pi)|^2|. \tag{9.2}$$

For $f(\pi) = 0$ this reduces to the normal uncertainty principle. But Kraus [Kra65, p. 376] argues that (9.2) has no physical interpretation.

List of Symbols

$\mathbb{N}$	Natural numbers $1, 2, 3, \ldots$
$\mathbb{Z}$	Integers $0, \pm 1, \pm 2, \pm 3, \ldots$
$\mathbb{R}$	Real numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{C}$	Complex numbers
$\mathcal{S}$	Schwartz functions (rapidly decreasing functions)
C_c^∞	Infinitely differentiable functions with compact support
C^∞	Infinitely differentiable functions
C_c	Functions with compact support
∂f	Partial derivative (distributional or normal) of f
f'	Derivative (distributional or normal) of f
$\operatorname{Im} z$	Imaginary part of $z \in \mathbb{C}$
$\operatorname{Re} z$	Real part of $z \in \mathbb{C}$
$\langle x, y \rangle$	Inner product of x and y
$[a, b]$	Closed interval in $\mathbb{R}$
$]a, b[$	Open interval in $\mathbb{R}$
$\|z\|$	Length of z in $\mathbb{C}$
$\|x\|$	Length of x in $\mathbb{R}^n$
$\|A\|$	Lebesgue measure of the set A
A°	Inner points of the set A
$\bar{A}$	Closure of the set A
∂A	Closure of the set A
$\in$	Member of
$\notin$	Not member of
$\subseteq$	Subset of
A^*	Adjoint of an operator A
$f(x) = O(x)$	$f(x)/x$ bounded for some limit
$\bar{z}$	Complex conjugate of $z \in \mathbb{C}$
$\hat{f}$	Fourier transform of f
$\mathcal{F}(f)$	Fourier transform of f
$\Psi(f)$	Fourier transform of f
L_p	Equivalence class of p-integrable functions
$\\|f\\|_p$	L_p-norm of f

List of Symbols

$f * g$	Convolution of functions f and g
$\mathcal{D}(A)$	Domain of an operator A
μ	Measure
lim	Limit
log	Natural logarithm
exp	Exponential function
e	Exponential function
$\mathrm{supp}(f)$	The support of f
1_I	Function equal to 1 on I and 0 elsewhere
∇f	n-tuple of partial derivatives of f
r_a	Right translation by a
l_a	Left translation by a
sgn	The modulus $z/\lvert z\rvert$ of $z \in \mathbb{C}$
End	The set of endomorphisms
Aut	The set of automorphisms
Ad	Adjoint representation
ad	The differential of Ad

Bibliography

[Art64] Emil Artin. *The gamma function*. Translated by Michael Butler. Athena Series: Selected Topics in Mathematics. Holt, Rinehart and Winston, New York, 1964.

[BG95] Carlos A. Berenstein and Roger Gay. *Complex analysis and special topics in harmonic analysis*. Springer-Verlag, New York, 1995.

[CP84] Michael G. Cowling and John F. Price. Bandwidth versus time concentration: the Heisenberg-Pauli-Weyl inequality. *SIAM J. Math. Anal.*, 15(1):151–165, 1984.

[CR97] Jens G. Christensen and Dan Rasmussen. Bachelor projekt: Invariant integration. Københavns Universitet, 1997.
A copy can be found at `http://www.math.ku.dk/~vepjan/bach.dvi`.

[DM72] H. Dym and H. P. McKean. *Fourier series and integrals*. Academic Press, New York, 1972.

[DS89] David L. Donoho and Philip B. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, 49(3):906–931, 1989.

[FS97] Gerald B. Folland and Alladi Sitaram. The uncertainty principle: a mathematical survey. *J. Fourier Anal. Appl.*, 3(3):207–238, 1997.

[Hal50] Paul R. Halmos. *Measure Theory*. D. Van Nostrand Company, Inc., New York, N. Y., 1950.

[Hel78] Sigurdur Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 80 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978.

[Hel84] Sigurdur Helgason. *Groups and geometric analysis*, volume 113 of *Pure and Applied Mathematics*. Academic Press Inc., Orlando, FL, 1984.

[Hir57] I. I. Hirschman, Jr. A note on entropy. *Amer. J. Math.*, 79:152–156, 1957.

[HR63] Edwin Hewitt and Kenneth A. Ross. *Abstract harmonic analysis. Vol. I: Structure of topological groups. Integration theory, group representations*. Die Grundlehren der mathematischen Wissenschaften, Bd. 115. Academic Press Inc., Publishers, New York, 1963.

Bibliography

[Kaw72] Tatsuo Kawata. *Fourier analysis in probability theory.* Academic Press, New York, 1972.

[Kna86] Anthony W. Knapp. *Representation theory of semisimple groups*, volume 36 of *Princeton Mathematical Series.* Princeton University Press, Princeton, NJ, 1986.

[Kra65] K. Kraus. Remark on the uncertainty between angle and angular momentum. *Z. Physik*, 188:374–377, 1965.

[Kra67] K. Kraus. A further remark on uncertainty relations. *Z. Physik*, 201:134–141, 1967.

[Ped89] Gert K. Pedersen. *Analysis now*, volume 118 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1989.

[Pri84] G. Baley Price. *Multivariable analysis.* Springer-Verlag, New York, 1984.

[PS72] G. Pólya and G. Szegő. *Problems and theorems in analysis. Vol. I: Series, integral calculus, theory of functions.* Springer-Verlag, New York, 1972.

[Rud87] Walter Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, 1987.

[Rud91] Walter Rudin. *Functional analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, 1991.

[RV97] Margit Rösler and Michael Voit. An uncertainty principle for ultraspherical expansions. *J. Math. Anal. Appl.*, 209(2):624–634, 1997.

[Sch81] Martin Schechter. *Operator methods in quantum mechanics.* North-Holland Publishing Co., New York, 1981.

[Seg51] I. E. Segal. A class of operator algebras which are determined by groups. *Duke Math. J.*, 18:221–265, 1951.

[Sel02] Kathi K. Selig. Uncertainty principles revisited. *Electron. Trans. Numer. Anal.*, 14:164–176 (electronic), 2002.

[ST83] Ian Stewart and David Tall. *Complex analysis.* Cambridge University Press, Cambridge, 1983.

[Str89] Robert S. Strichartz. Uncertainty principles in harmonic analysis. *J. Funct. Anal.*, 84(1):97–114, 1989.

[Tit37] Edward C. Titchmarsh. *Introduction to the theory of Fourier Integrals.* Oxford University Press, Oxford, 1937.

[War71] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups.* Scott, Foresman and Co., Glenview, Ill.-London, 1971.