BIRKHOFF NORMAL FORMS FOR THE GROSS-PITAEVSKII EQUATION

BIRKHOFF NORMAL FORMS WITH APPLICATION TO THE GROSS-PITAEVSKII EQUATION

By

Zhenbin Yan, Ph.D

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY | McMaster University
(2007 Mathematics) | Hamilton, Ontario

TITLE: Birkhoff Normal Form with Application to Gross Pitaevskii Equation

AUTHOR: Zhenbin Yan, B.Sc (Wuhan University), M.S. (Wuhan University, P.R.China),

SUPERVISOR: Dr. Walter Craig

NUMBER OF PAGES: ix, 190

TO MY PARENTS

Abstract

This thesis investigates a 1-dimensional Gross-Pitaevskii (GP) equation from the viewpoint of a system of Hamiltonian partial differential equations (PDEs). A theorem on Birkhoff normal forms is a particularly important goal of this study. The resulting system is a perturbed system of a completely resonant system, which we analyze, using several forms of perturbation theory.

In chapter two, we study estimates on integrals of products of four Hermite functions, which represent coefficients of mode coupling, and play an important role in the proof of the Birkhoff normal form theorem. This is a basic problem, which has a close relationship with a problem of Besicovitch, namely the behavior of the L^p norms of L^2 -normalized Hermite functions.

In chapter three we carefully reconsider the linear Schrödinger equation with a harmonic potential, and we introduce a family of Hilbert spaces for studying the GP equation, which generalize the traditional energy spaces in which one works. One unexpected fact is that these function spaces have a close relationship with the former works for the tempered distributions, in particular the N-representation theory due to B. Simon, and V. Bargmann's theory, which uncovers relationship between the tempered distributions and his function spaces through the so-called Segal-Bargmann

transformation. In addition, our function spaces have a nice relationship with the Sobolev spaces. In this chapter, a few other questions regarding these function spaces are discussed.

In chapter four the proof of the Birkhoff normal form theorem on spaces we have introduced are provided. The analysis is divided into two cases according to the regularity of the related function space. After proving the Birkhoff normal form theorem, we made an analysis of the impact of the perturbation on the main part of the GP system, which we remark is completely resonant.

Acknowledgements

I express my sincere gratitude and appreciation to my supervisor Dr. Walter Craig, for his inspiring guidance, encouragement and wise recommendations in the preparation of this thesis.

I would like to thank my committee members Dr. Alama and Dr. Pelinovsky, for their insightful comments and criticism of this thesis.

My special thanks go to Dr. Slim Ibrahim and Feng Su for their helpful discussion and kind help throughout this thesis. I would like also to thank all my colleagues and friends, in particular Yi Zhang, who have been standing by me during my Ph.D. studies at McMaster University.

Last but not least, I would like to thank and my families, for their unconditional love and support. Without their support, I couldn't have done it.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

Bose-Einstein condensation is an exotic quantum phenomenon which is now a subject of intense theoretical and experimental study. A Bose-Einstein condensate (BEC) is a state of matter formed by a system of bosons confined in an external potential and cooled to temperatures very near to absolute zero. This state of matter was first predicted as a consequence of quantum mechanics by Albert Einstein, building upon the work of Satyendra Nath Bose in 1925, hence the name. Seventy years later, the first such condensate was produced by Eric Cornell and Carl Wieman in 1995 at the University of Colorado at Boulder NIST-JILA lab. For this work, Cornell, Wieman and Wolfgang Ketterle at MIT were awarded the 2001 Nobel Prize in Physics in Stockholm, Sweden.

Before this famous experimental realization of BEC, a remarkable series of investigations were conducted, formulated in terms of the Gross-Pitaevskii (GP) equation [Gro][Pi], which turns out to provide a good description on the behavior of BEC's. This equation can be written as

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\Delta_x\psi + V(x)\psi + \lambda\left|\psi\right|^2\psi, \tag{1.1}$$

where m is the mass of the bosons, $V(x)$ is the external potential and λ is a coefficient representating the inter-particle interactions. The sign of the coefficient λ differs for different chemical elements. For example, it is negative for ^{7}Li atoms ([B-S-T-H][B-S-H]) as well as for ^{85}Rb, and positive for ^{87}Rb, ^{23}Na and ^{1}H. We would like to mention that the case of harmonic potential, that is, $V(x) = \frac{m}{2}\omega^2x^2$ $(\omega > 0)$, is one of the most important cases, as it models the magnetic field used to confine the particles.

From the point of view in mathematics, equation (1.1) is a nonlinear Schrödinger equation with a potential function $V(x)$. In the simplest cases, the potential function can be a constant function. J. Ginibre and G. Velo have considered this kind of nonlinear Schrödinger equation for a larger class of nonlinearities [G-V], and the local existence and uniqueness of solutions of the initial value problem in the Sobolev space $H^1(\mathbb{R}^n)$ were obtained in their work. In some cases, they proved the existence of the global solutions. The method can be summarized as follows: first study the corresponding linear PDE and the property of the corresponding Schrödinger kernel; then write the PDE as an equivalent integral equation in a suitable Banach space; finally use the property of the Schrödinger kernel and fixed point theorem to get the solution.

In 1979, D. Fujiwara proved in [F1][F2] that for potentials of quadratic growth, the Schrödinger kernel has the form

$$k(t,x,y) = (\frac{-i}{2\pi t})^{n/2} a(t,x,y) e^{iS(t,x,y)}$$

for a short of time t, where $S(t,x,y)$ is smooth and $a(t,x,y)$ is a bounded continuous function of t, x and y. This result was a kind of generalization of the Mehlor's formula [F-H], which provides the exact Schrödinger kernel for Harmonic potential $V(x) = \frac{m}{2}\omega^2 x^2$. On the basis of the work by D. Fujiwara, Yong-Geun Oh [O] made a further study of the following Cauchy problem in 1989

$$i\partial_t \psi = -\frac{1}{2}\Delta_x \psi + V(x)\psi - |\psi|^{p-1}\psi, \quad 1 \le p < 1 + \frac{4}{n}, \tag{1.2}$$

where the potential V is bounded below and satisfies the condition that $|D^\alpha V|$ is bounded for all $|\alpha| \ge 2$. In the function space of $D(\sqrt{\frac{1}{2}\Delta + V})$, the domain of definition of the square root of the Schrödinger kernel $\frac{1}{2}\Delta + V$, he proved the long time existence of the flow in that space.

In the recent years, more work have been done, focusing on the case of harmonic potentials. In 2002, R. Carles [C1] studied a nonlinear Schrödinger equation in the following form

$$\begin{cases} i\hbar\partial_t u^\hbar + \frac{\hbar^2}{2}\Delta_x u^\hbar = \frac{\omega^2}{2}x^2 u^\hbar + \lambda \left|u^\hbar\right|^{2\sigma} u^\hbar, \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^n, \\ u^\hbar|_{t=0} = u_0^\hbar, \end{cases} \tag{1.3}$$

where $\hbar > 0$, $\lambda \in \mathbb{R}$ and $\omega, \sigma > 0$. With the local existence of the equation already known, the author made use of the conservation laws of the above PDE to study the global well-posedness in the space $\Sigma := \{u \in L^2(\mathbb{R}^n) | \ xu, \ \nabla u \in L^2(\mathbb{R}^n)\}$. The author pointed out that there are following cases that a datum in the space Σ can admit global flows:

- If $\lambda \geq 0$, then the solution is defined globally in time;

- If $\lambda < 0$ and $\sigma < 2/n$, then the solution is also defined globally in time;

- If $\lambda < 0$ and $\sigma \geq 2/n$, then the solution is defined globally in time when the initial data $u_0^\hbar$ is sufficiently small.

He also found that finite blow up solution do indeed exist. The conserved energy can be written as

$$E^\hbar = \frac{1}{2}\left\|\hbar\nabla_x u^\hbar(t)\right\|_{L^2}^2 + \frac{\omega^2}{2}\left\|xu^\hbar(t)\right\|_{L^2}^2 + \frac{\lambda}{\sigma+1}\left\|u^\hbar(t)\right\|_{L^{2\sigma+2}}^{2\sigma+2}.$$

The author proved that if $\lambda < 0$, $\sigma \geq 2/n$ and $E^\hbar \leq \frac{\omega^2}{2}\left\|xu^\hbar(t)\right\|_{L^2}^2$, then the solution collapse at a time $t_*^\hbar \leq \pi/2\omega$. In that paper, he also provided upper bound and lower bound estimate of the breaking time.

In addition to the cubic nonlinearity for the Schrödinger equation with harmonic potential, some other nonlinearities may also be a good choice to model Bose-Einstein condensation. In [K-N-S-Q], the author proposes a quintic nonlinearity in space dimension one; and in [Z], the author suggests more generally the study of

$$\begin{cases} i\partial_t u + \frac{1}{2}\Delta_x u = \frac{\omega^2}{2}x^2 u + \lambda\left|u\right|^{4/n}u, \quad (t,x)\in\mathbb{R}\times\mathbb{R}^n, \\ u|_{t=0} = u_0. \end{cases} \tag{1.4}$$

Then in the year 2002, R. Carles also studied the partial differential equation (1.4) in [C2]. He found that there is a coordinate change which transforms the above equation into the following form

$$\begin{cases} i\partial_t v + \frac{1}{2}\Delta_x v = \lambda\left|v\right|^{4/n}v, \quad (t,x)\in\mathbb{R}\times\mathbb{R}^n, \\ v|_{t=0} = v_0, \end{cases} \tag{1.5}$$

while the latter equation had been extensively studied. A quick and complete reference can be found from the website maintained by Jim Colliander, Mark Keel, Gigliola Staffilani, Hideo Takaoka, and Terry Tao [Webpage1]. Here we just provide some basic information for this equation (1.5).

The local well-posedness for this equation (1.5) has been provided by J. Ginibre and G. Velo's work in [G-V]. There are predictions of the blow up time and rate by C. Sulum. In 1982, M. I. Weinstein proved that if the initial data comes from the space $\Sigma := \{u \in L^2(\mathbb{R}^n) |\ xu,\ \nabla u \in L^2(\mathbb{R}^n)\}$, then one can have global solution in the defocusing case ($\lambda > 0$); in the focusing case ($\lambda < 0$), if the initial data is small enough, then one can also get global solution; otherwise, the wave v may collapse in finite time. In more detail, qualifying this criterion in the focusing case is called as critical mass, which equals to the L^2 norm of the unique radial solution [Kw] of

$$\begin{cases} -\frac{1}{2}\Delta Q + Q = -\lambda\, |Q|^{4/n}\, Q, \quad \text{in } \mathbb{R}^n, \\ \qquad\qquad Q > 0, \quad \text{in } \mathbb{R}^n. \end{cases}$$

In particular, he pointed out that there is an initial data u_0 with this critical mass such that it leads to a blow up solution at finite time. In other words, this critical L^2 norm is sharp. In 1993, F. Merle [Me1][Me2] proved that up to invariants of (1.5), the blowing up solutions enlightened by Weinstein are the only ones. More refined descriptions for the blow-up solutions were obtained in recent years by F. Merle and P. Raphael [M-R1].

Since R. Carles have found the coordinate transformations between the equation (1.4) and (1.5) (in the space $C(I, \Sigma)$), many results obtained for the equation (1.5) can be transposed into its corresponding version for the equation (1.4) (see [C2]). Of these results, the existence of the critical mass in fact had been proved by Zhang Jian in [Z].

In this thesis, we turn our attention into the initial value problem of the equation (1.1) with the harmonic potential in one space dimension case. That is

$$\begin{cases} i\psi_t = \frac{1}{2}\psi_{xx} - \frac{x^2}{2}\psi - g|\psi|^2\psi & x \in \mathbb{R}^1 \\ \psi(x,0) = \psi_0(x) & \psi \text{ complex valued,} \end{cases} \tag{1.6}$$

where g is a constant. According to the materials introduced above, it is already well understood on the question of the global well-posedness: any initial data in the space $\Sigma := \{u \in L^2(\mathbb{R}^n) | \; xu, \; \nabla u \in L^2(\mathbb{R}^n)\}$ will result in a global flow in time in the same space. But beyond that, we are still interested in finding more information to describe the behaviors of the solutions on time. So we take the viewpoint from the Hamiltonian system to study this equation in this thesis.

1.2 Hamiltonian systems

Many important physic models can be regarded as Hamiltonian systems, ones often as nearly integrable. In the case of finite dimensional systems of ODE, A. N. Kolmogorov, V. I. Arnold and J. Moser [Ko][A][Mo] introduced a theory, which roughly speaking, states that for sufficiently small perturbations of an integrable system, almost all invariant tori are preserved. In other words, there exist abundant quasi-periodic motions for an integrable system under sufficient small perturbations. This theory now is known as KAM theory. Of course, there are also many initial data leading to a motion not quasi-periodic, but they will at least admit the Nekhoroshev stability. In his paper [N], Nekhoroshev showed that under a small perturbation of an integrable Hamiltonian system, the action variables of an arbitrary orbit change exponentially slowly.

When a Hamiltonian system corresponds to a PDE, things become much more complicated. During the last fifteen years the perturbation theory of Hamiltonian partial differential equations has been extensively studied, and the subject is still under development. Typical problems (but not restricted to them) in this field are as the following:

- Can we get long time stability of solutions of small perturbations of linear or integrable Hamiltonian PDEs? This is the PDE setting of the Nekhoroshev problem.

- Can we find invariant tori for a given Hamiltonian PDEs? This is the PDE setting of the KAM theory problem.

- As a flow in a given phase space, how fast its norm can grow as the time tends to infinity or a finite blow-up time?

One general philosophy is commonly used in the research work within this field: by carefully choosing symplectic transformations, one changes a Hamiltonian system into a form with a well understood part, usually a linear or integrable part, under a sufficient small perturbation. Birkhoff normal form theory in the different PDE settings [Bam2][Gre][B-G] can usually take an important role to fulfill this philosophy.

Here we would like to focus our attention on the study of nonlinear Schrödinger equations within the Hamiltonian PDE theory structure. In 1996, S. Kuksin and J. Pöschel [K-P] studied the following equation

$$iu_t = u_{xx} - mu - f(|u|^2)u \tag{1.7}$$

on the finite x-interval $[0, \pi]$ with Dirichlet boundary conditions. The parameter m is real and f is real analytic in some neighbourhood of the origin in $\mathbb{C}$. When $f(|u|^2) = |u|^2$, it is just the cubic nonlinear Schrödinger equation, which was already known to be integrable. By writing the solution as a sum of the L^2 normalized eigenfunction corresponding to the operator $-\Delta+m$, this equation takes the forms of an infinite dimensional Hamiltonian system. The authors first proved the existence of a symplectic transformation, which can turn the Hamiltonian function into its Birkhoff normal form up to order four. In this way, the original equation became an integrable Hamiltonian system with a perturbation of order at least six near the origin point. Then an infinite dimensional KAM-theorem was applied, with which the authors showed the existence of an invariant Cantor manifolds of quasi-periodic oscillations.

The same equation was studied also by D. Bambusi in 1999 [Bam1] towards a Nekhoroshev type result. He proved that if a solution initiates near a finite dimensional torus, or in other words, the initial energy concentrates essentially in some finitely many eigenmodes, then it will remain in a small neighbourhood of that torus for at least an exponentially long time. The first step of his proof consists in putting the nonlinear Schrödinger equation into a Birkhoff normal form up to an exponentially small remainder, in which the truncation up to the forth order provides an integrable system. Then the author, inspired by a variant of the technique of approximation by periodic orbits introduced by Lochak [L], constructed another normal form close to a fixed finite dimensional torus and showed the long time stability of the solution close to it. Later, J. Pöschel simplified D. Bambusi's proof and obtained a slight refinement of the theorem in [Po]. In doing this, one key step is to find a symplectic transformation which transforms the original Hamiltonian into an integrable one, plus a perturbation which is small.

In those works mentioned above, Birkhoff normal forms have played important roles. It is a natural choice to use this method to study the nonlinear harmonic oscillator (1.6). In doing so, we need answer the following questions: What is the function space that we take for the domain and range of the transformation? Does the necessary symplectic transformation really exist? What can we deduce out for the dynamics of the original system? We would like to point out that although our equation looks similar to the equation (1.7), the two are really very different.

To prove a Birkhoff normal form theory in infinite dimensional case, the main difficulty consists in studying nonresonance property that allows to remove from the nonlinearity all the relevant non-normalized monomials. For the equation (1.7), most monomials can be removed and the system will be transformed into an integrable one with a small perturbation. But for equation (1.6), the eigenvalues of the Schrödinger operator $-\Delta + x^2$ are in the form of $(2k+1)/2$ with $k \in \mathbb{Z}_+$. In this case, many nontrivial linear combination of the eigenvalues give zero implying the existence of many resonances. So there is no way to remove those corresponding resonant monomials by the procedure of canonical transformations. In fact our equation is called a completely resonant PDE, whose perturbation is much harder to understand than the nearly integrable cases.

There are also other technical difficulties quite different from those encountered in work on the equation (1.7). One comes from the nature of the eigenfunctions. For a problem which is essentially the Laplace operator with Dirichlet or periodic boundary condition, its eigenfunctions are very simple–trigonometric functions. In our case, the eigenfunctions of the Schrödinger operator $-\Delta + x^2$ are Hermite functions, which are more complicated. It makes the estimates required for our Birkhoff normal form theorem more complicated to deal with. The second difficulty comes from the choice of the phase spaces. Traditionally the study of the Gross Pitaevskii equation is based on the space $\Sigma := \{u \in L^2(\mathbb{R}^n) |\ xu,\ \nabla u \in L^2(\mathbb{R}^n)\}$, sometimes called as the virial space. Our choice should be consistent with this traditional space. And after it is chosen, we need to answer the question as to what these function spaces really are.

This thesis mainly focuses on the Birkhoff normal form theorem for the equation (1.1), and the above problems are addressed. In the chapter two, we study the estimate on the integrals of the product of four Hermite functions. It turns out to have relationship with a problem of Besicovitch: what is the behavior of the L^p norm of the Hermite functions? In particular, when all four Hermite functions are equal, the 4th root of the integral is just the L^4 norm of that Hermite function. This case was once studied by G. Freud and G. Németh [F-N]. We generalize that result into the case of integrals of the square of the products of two Hermite functions. For other cases, we also provide the estimate needed to prove the Birkhoff normal form theorem.

In the chapter three we carefully reconsider the linear Schrödinger equation with harmonic potential, and provide our choice for the phase spaces. One unexpected thing is that we realize that those function spaces have a very close relationship with the theory of tempered distributions, especially the N-representation theory due to B. Simon [Si] and V. Bargmann's function spaces together with the corresponding Segal-Bargmann transformation [Bar1][Bar2]. It is true that our function spaces can be a very good complement for the N-representation theory, and also a natural way to provide descriptions for the rapid decreasing functions and tempered distributions.

Besides that, our function spaces have very close relationship with the Sobolev spaces. It is proved that when the regularity index number is a nonnegative integer, the function space is just the intersection of the Sobolev space with its image under the Fourier transformation; when that number is a negative integer, then it is just the "summation" of those two Banach spaces. This result also seems to be true when the regularity index is any other real number. To this conjecture, we haven't yet found a proof. But we find the following fact that partly supports the conjecture: for any regularity index big enough our function space forms a subalgebra of the Sobolev spaces, which itself is a Banach algebra with respect to pointwise multiplication. In the rest of that chapter, a few other problems of the function spaces are also discussed.

In the chapter four the proof for our Birkhoff normal form theorem are provided. It is divided into two cases according to the regularity index of the related function space. In particular when the index is large, we need a more detailed understanding of the integrals of the product of four Hermite functions other than those provided in the chapter two. Those new estimates are provided in a separated subsection. After providing the Birkhoff normal form theorem, we give an analysis on the impact of the perturbation on the main part system, which is completely resonant. In general, this is a very difficult problem, and here we provide a few results that we have obtained.

Chapter 2

Hermite Functions

2.1 Introduction

This chapter focuses on the properties of the Hermite functions. At first, we review facts about the eigenfunctions of the operator $-\Delta+x^2$, which are Hermite functions. Then any reasonable function in our study can be represented as the summation of its projections to every eigenfunction space. Below we use $h_j(x)$ ($j = 0, 1, 2, \cdots$) to denote the j-th Hermite function with unit L^2 norm. For our Hamiltonian PDE

$$\begin{cases} i\psi_t = \frac{1}{2}\psi_{xx} - \frac{x^2}{2}\psi - g|\psi|^2\psi & x \in \mathbb{R}^1 \\ \psi(x,0) = \psi_0(x) & \psi \text{ complex valued,} \end{cases} \tag{2.1}$$

if we write $\psi(t,x) = \sum_{j\geq 0} q_j(t)h_j(x)$, then the Hamiltonian function of the system will be (for more details, see chapter four)

$$H = \sum_{j\geq 0}\omega_j |q_j|^2 + \frac{g}{2}\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_k}\overline{q_l}q_m q_n.$$

A basic problem is how the Hamiltonian function depends on the coordinates q_k. In particular, we need understand the behavior of the coefficients C_{klmn}, which are integrals of products of four Hermite functions.

We would like to remark that the above problem is related to the following two questions. One was posed by Besicovitch: what is the behavior of the L^p norm of the Hermite functions? In particular, we note that $C_{kkkk} = \|h_k(x)\|_{L^4}^4$. In 1948, Ida W. Busbridge [Bu] obtained a formula that can express the integral of the product of Hermite polynomials with weighted function $\exp\{-x^2/a\}$ $(a > 0)$ into the summation of a sequence. Unfortunately, those terms in the sequence are not in same signatures, which implies that the formula can't provide good answer to the Besicovitch's question, or to our question. In 1973, G. Freud and G. Németh [F-N] made a very exact estimate on the terms C_{kkkk}. And by the year 1984, the Besicovitch's question have been completely solved (see lemma 1 in [Ma]).

The other one is related to multilinear eigenfunction estimates. In paper [D-S], J.-M. Delort and J. Szeftel obtained an estimate for the integral of the product of the eigenfunctions of the Laplace-Beltrami operator on Zoll manifold, and they used that result to study *long-time existence for small initial data to nonlinear Klein-Gordon equations on tori and spheres.*

Returning to our question, we found a very exact estimate for the term C_{mmnn}, which generalizes the result by G. Freud and G. Németh. In particular, we get $C_{mmnn} \approx \frac{2}{\sqrt{m}}E(s)$, where $s = \sqrt{\frac{n}{m}} < 1$ and $E(s) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1-s^2\sin^2\theta}}$ is a version of an elliptic integral. To keep in the same style of the result by G. Freud and G. Németh, we confirm that $C_{mmnn} \lesssim \langle m\rangle^{-\frac{1}{2}} \ln^{1/2}\langle m\rangle \ln^{1/2}\langle n\rangle$. For general terms C_{klmn}, a simple application of the Hölder's inequality provide such an estimate: for nonnegative integers $k \geq l \geq m \geq n$, it is true that

$$|C_{klmn}| \lesssim \langle k \rangle^{-\frac{1}{4}} \langle l \rangle^{-\frac{1}{4}} (\ln \langle k \rangle \ln \langle l \rangle \ln \langle m \rangle \ln \langle n \rangle)^{\frac{1}{4}}. \tag{2.2}$$

We would like to point out that this estimate is good when the numbers k, l, m and n almost equal to each other, but it will be too big in some other situations. For example, if one of the numbers is much larger than all the others, a much better estimate is also provided in this thesis (see chapter four).

For the purpose of making this chapter self contained, many basic properties of the Hermite functions are also reviewed. Most of them are needed for the estimate of the coefficients C_{klmn}.

2.2 The quantum harmonic oscillator

In this section we derive the eigenvalues and eigenfunctions of the quantum harmonic oscillator. We consider the differential operator $A = \frac{-\Delta+|x|^2}{2}$ acting on the complex-valued function space $L^2(\mathbb{R}^n)$, which is equipped with inner product given by $(\varphi, \psi)_{L^2} = \int_{\mathbb{R}^n} \varphi(x)\overline{\psi(x)}dx$. By the spectral theory[T], A is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$ and it has compact resolvent, therefore $L^2(\mathbb{R}^n)$ has an orthogonal basis made up of eigenfunctions of A. Below we will describe that in the 1-dimensional case, these eigenfunctions are Hermite functions.

Let us introduce the two well known creation and annihilation operators

$$a = \frac{1}{i}(\frac{d}{dx} + x),\ a^* = \frac{1}{i}(\frac{d}{dx} - x),$$

with which we can write

$$A = (aa^* - I)/2 = (a^*a + I)/2,$$

such that

$$[A, a] = -a, [A, a^*] = a^*.$$

Suppose that φ_j is an eigenfunction of the operator A, i.e. $A\varphi_j = \lambda_j\varphi_j$ where $\varphi_j \in D(A) = \{u \in L^2(\mathbb{R}) | -\frac{d^2}{dx^2}u + x^2u \in L^2(\mathbb{R})\}$ and $\lambda_j \in \mathbb{R}$. Then $a\varphi_j$ is also an eigenfunction of the operator A. Since A is an strict elliptic operator, by the classical regularity theory of the second order elliptic operators, we can deduce that any eigenfunction of operator A must be smooth. Furthermore we can easily find that

$$\varphi_j \in D(\sqrt{A}) = \{u \in L^2(\mathbb{R}) | -\frac{d}{dx}u \in L^2(\mathbb{R}) \text{ and } xu \in L^2(\mathbb{R})\}).$$

By the commutator calculation above,

$$-a\varphi_j = [A, a]\varphi_j = A(a\varphi_j) - a(A\varphi_j) = A(a\varphi_j) - \lambda_j a\varphi_j.$$

Thus

$$A(a\varphi_j) = (\lambda_j - 1)a\varphi_j, \tag{2.3}$$

which implies that $a\varphi_j \in D(A)$ is also an eigenfunction of operator A for eigenvalue $\lambda_j - 1$.

Similarly, we also have

$$a^*\varphi_j \in D(A)$$

and

$$A(a^*\varphi_j) = (\lambda_j + 1)a^*\varphi_j. \tag{2.4}$$

Let us define the eigenfunctions space of the operator A as $Eigen(\lambda, A) =$

$\{u \in D(A) \mid Au = \lambda u\}$. Equations (2.3) and (2.4) imply that

$$\begin{cases} a : Eigen(\lambda, A) & \longrightarrow \quad Eigen(\lambda - 1, A) \\ a^* : Eigen(\lambda, A) & \longrightarrow \quad Eigen(\lambda + 1, A). \end{cases} \tag{2.5}$$

Since A is self-adjoint and $(Au, u) \geq \frac{1}{2}\|u\|_{L^2}^2$ for any $u \in D(A)$, it follows that any eigenvalue λ of operator A must satisfy $\lambda \geq \frac{1}{2}$.

Noticing

$$aa^* - I = a^*a + I = 2A,$$

we can conclude that a and a^* in (2.5) are both isomorphisms for those eigenvalues $\lambda \geq \frac{3}{2}$. On the other hand, a must annihilate $Eigen(\lambda_0, A)$ when λ_0 is the smallest element eigenvalue of the operator A. If $\varphi_0 \in Eigen(\lambda_0, A)$, then

$$a\varphi_0 = \frac{1}{i}(\frac{d}{dx}\varphi_0 + x\varphi_0) = 0.$$

By solving this ODE, we find

$$\varphi_0(x) = ce^{-\frac{x^2}{2}},$$

meaning that $\lambda_0 = \frac{1}{2}$ and $Eigen(\frac{1}{2}, A) = span\{e^{-\frac{x^2}{2}}\}$, which is a 1-dimensional space.

For any other eigenvalue λ, we apply the operator $a : Eigen(\lambda, A) \longrightarrow Eigen(\lambda - 1, A)$ repeatedly. After finitely many steps, it must end up with the mapping $a : Eigen(\frac{1}{2}, A) = span\{\varphi_0(x)\} \longrightarrow \{0\}$. And all the mappings in this process except the last one are in fact isomorphisms. Therefore, we can conclude that

$$\begin{aligned} Spec(A) &= \{\text{eigenvalues of } A\} = \{\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \cdots\} \\ &= \{k + \tfrac{1}{2}, k = 0, 1, 2, \cdots\} \end{aligned}$$

and

$$Eigen((k + \frac{1}{2}), A) = span\{\varphi(x) = (\frac{\partial}{\partial x} - x)^{k-1}e^{-\frac{x^2}{2}}\}.$$

Those eigenfunctions can be written in such forms

$$(\frac{\partial}{\partial x} - x)^k e^{-\frac{x^2}{2}} = (\frac{\partial}{\partial x} - x)^{k-1}((\frac{\partial}{\partial x}(e^{-x^2}))e^{\frac{x^2}{2}}) \tag{2.6}$$

$$= (\frac{\partial}{\partial x})^k (e^{-x^2} e^{\frac{x^2}{2}}) \tag{2.7}$$

$$= (-1)^k H_k(x) e^{-\frac{x^2}{2}}, \tag{2.8}$$

where $H_k(x)$ $(k = 0, 1, 2, \cdots)$ are Hermite polynomials, given by Rodrigues' formula[M-O-S]

$$H_k(x) = (-1)^k((\frac{\partial}{\partial x})^k(e^{-x^2}))e^{x^2}. \tag{2.9}$$

The first 6 Hermite polynomials are

$$H_0(x) = 1, \tag{2.10}$$

$$H_1(x) = 2x, \tag{2.11}$$

$$H_2(x) = 4x^2 - 2, \tag{2.12}$$

$$H_3(x) = 8x^3 - 12x, \tag{2.13}$$

$$H_4(x) = 16x^4 - 48x^2 + 12, \tag{2.14}$$

$$H_5(x) = 32x^5 - 160x^3 + 120x. \tag{2.15}$$

2.3 Properties of Hermite polynomials and functions

In this section, we will review some basic properties of the Hermite polynomials (functions), most of which come from the theory of special functions [M-O-S][A-S]. The materials below are organized in an order with the intention of being a self contained account and easy accessible. All the properties are provided with a short proof.

Proposition 2.3.1. *Hermite Polynomials are mutually orthogonal, under the inner product with respect to the weight functions* e^{-x^2}. *In particular, we have*

$$\int_{-\infty}^{+\infty} H_k(x)H_m(x)e^{-x^2}dx = \delta_{km}2^m m!\sqrt{\pi}. \tag{2.16}$$

Proof. Without loss of generality, we assume that $k \geqslant m$,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} H_k(x)H_m(x)e^{-x^2}dx \\
&= \int_{-\infty}^{+\infty} (-1)^k e^{x^2}((\frac{d}{dx})^k(e^{-x^2}))H_m(x)e^{-x^2}dx \\
&= \int_{-\infty}^{+\infty} e^{-x^2}((\frac{d}{dx})^k(H_m(x)))dx.
\end{aligned} \tag{2.17}$$

Since $H_m(x)$ is a polynomial with order m. so we have two possible cases.

If $k > m$, we must have $(\frac{d}{dx})^k H_m(x) = 0$, which implies $\int_{-\infty}^{+\infty} H_k(x)H_m(x)e^{-x^2}dx =$ 0.

If $k = m$, by using the fact that the leading term of $H_m(x)$ is $2^m x^m$, we have

$$\begin{aligned}
&\int_{-\infty}^{+\infty} H_k(x)H_m(x)e^{-x^2}dx \\
&= \int_{-\infty}^{+\infty} e^{-x^2}(\frac{d}{dx})^m(H_m(x))dx \\
&= 2^m m!\sqrt{\pi}. \qquad (2.18)
\end{aligned}$$

Hence (2.16) holds. □

Let us denote $h_k(x) = H_k(x)e^{-\frac{x^2}{2}}/(k!2^k\sqrt{\pi})^{1/2}$ as the normalized Hermite functions, which are also the eigenfunctions of the operator $A = (-\Delta + x^2)/2$. According to the result in spectral analysis, all these functions $\{h_0(x), h_1(x), \cdots, h_k(x), \cdots\}$ form an orthonormal basis of the function space $L^2(\mathbb{R}^1) = \{u : \mathbb{R} \to \mathbb{C} \mid \int_{-\infty}^{+\infty} |u(x)|^2 dx < +\infty\}$. This result can be easily extended to higher dimensions. In general, for the complex function space $L^2(\mathbb{R}^n) = \{u : \mathbb{R}^n \to \mathbb{C} \mid \int_{\mathbb{R}^n} |u(x)|^2 dx < +\infty\}$, there is an orthonormal basis as $\{h_{k_1}(x_1) \otimes h_{k_2}(x_2) \otimes \cdots \otimes h_{k_n}(x_n),\ k_j = 0, 1, 2, \cdots\}$.

Proposition 2.3.2. *The Hermite functions satisfy the ordinary differential equation:*

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0. \qquad (2.19)$$

Proof. Since $H_n(z)e^{-\frac{x^2}{2}}$ is an eigenfunction of the operator $A = \frac{-\Delta+x^2}{2}$ with eigenvalue $n + \frac{1}{2}$, we can just plug in to get the property above,

$$\begin{aligned}
(n+\frac{1}{2})H_n(x)e^{-\frac{x^2}{2}} &= \frac{-\Delta + x^2}{2}(H_n(x)e^{-\frac{x^2}{2}}) && (2.20)\\
&= (n+\frac{1}{2})H_n(x)e^{-\frac{x^2}{2}} - (H_n(x)e^{-\frac{x^2}{2}})'' + x^2(H_n(x)e^{-\frac{x^2}{2}}) && (2.21)\\
&= (n+\frac{1}{2})H_n(x)e^{-\frac{x^2}{2}} - (H_n'' - 2xH_n'(x) - H_n(x) && (2.22)\\
&+ x^2H_n(x))e^{-\frac{x^2}{2}} + x^2H_n(x)e^{-\frac{x^2}{2}}. && (2.23)
\end{aligned}$$

So we have

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0.$$

Using the ODEs above, we can get the explicit formula of Hermite functions. □

Proposition 2.3.3. *An explicit representation of the Hermite functions is that*

$$H_n(x) = \sum_{k=0}^{[\frac{n}{2}]} (-1)^k \frac{n!}{k!(n-2k)!}(2x)^{n-2k}, \tag{2.24}$$

where as usual $[\frac{n}{2}]$ *denote the biggest integer not greater than* $\frac{n}{2}$.

Proof. It is easy to see that the Hermite Polynomial $H_n(x)$ is an odd function when n is odd and an even function when n is even. Thus we can write $H_n(x)$ as

$$H_n(x) = \begin{cases} a_n x^n + a_{n-2}x^{n-2} + \cdots + a_2 x^2 + a_0 & n \text{ even}, \\ a_n x^n + a_{n-2}x^{n-2} + \cdots + a_3 x^3 + a_1 x & n \text{ odd}. \end{cases} \tag{2.25}$$

Through observing (2.9), we can find that the highest order term in $H_n(x)$ arises when every derivative falls on the factor e^{-x^2}. Thus we know $a_n x^n = (-2x)^n(-1)^n = (2x)^n$. By proposition (2.3.2), $H_n(x)$ satisfies the ordinary differential equation (2.19), in the case of n even, that is

$$\begin{aligned} 0 &= H_n''(x) - 2xH_n'(x) + 2nH_n(x) \quad (2.26) \\ &= n(n-1)a_n x^{n-2} + (n-2)(n-3)a_{n-2}x^{n-4} + \cdots + 2\cdot 1 a_2 \\ &+ 2na_n x^n + 2na_{n-2} + \cdots + 2na_0 \\ &- 2na_n x^n - 2(n-2)a_{n-2}x^{n-2} - \cdots - 2\cdot 2a_2 x^2. \quad (2.27) \end{aligned}$$

Comparing the coefficients of the polynomials on both sides, we get the combinational expressions

$$a_{n-2} = -\frac{n(n-1)}{2 \cdot 2 \cdot 1} a_n, \tag{2.28}$$

$$a_{n-4} = -\frac{(n-2)(n-3)}{2 \cdot 2 \cdot 2} a_{n-2}, \tag{2.29}$$

$$\cdots \tag{2.30}$$

$$a_0 = -\frac{2 \cdot 1}{2 \cdot 2 \cdot (\frac{n}{2})} a_2. \tag{2.31}$$

So we have

$$a_{n-2k} = (-1)^k \frac{n(n-1)\cdots(n-2k+1)}{4^k k!} a_n \tag{2.32}$$

$$= (-1)^k \frac{n!}{k!(n-2k)!} 2^{n-2k}, \qquad 0 \leqslant k \leqslant \frac{n}{2}. \tag{2.33}$$

The last equality above implies that we have proved (2.24) when n is even. In the case of n odd, we can repeat the process above again to arrive at (2.24). □

With the help of the explicit formula of Hermite functions, we can prove certain relationships between Hermite polynomials and their derivatives.

Proposition 2.3.4.

$$H'_{n+1}(x) = 2(n+1)H_n(x), \qquad (n = 0, 1, 2, \cdots). \tag{2.34}$$

Proof. This follows from the explicit formula for the Hermite polynomials.

$$\begin{aligned} H'_{n+1}(x) &= \tfrac{d}{dx}\Big(\sum_{k=0}^{[(n+1)/2]} \tfrac{(n+1)!}{k!(n+1-2k)!}(-1)^k (2x)^{n+1-2k} \Big) \\ &= \sum_{k=0}^{[n/2]} \tfrac{2(n+1)\, n!}{k!(n-2k)!}(-1)^k (2x)^{n-2k} \\ &= 2(n+1)\ H_n(x). \end{aligned}$$

□

We also have another proposition in a similar form, which is often very useful.

Proposition 2.3.5.

$$\frac{d}{dx}(H_n(x)e^{-x^2}) = (-1)\,H_{n+1}(x)e^{-x^2} \qquad n = 0, 1, 2, \cdots.$$

Proof. The result follows from the Rodrigues" formula,

$$\begin{aligned} \tfrac{d}{dx}(H_n(x)e^{-x^2}) &= (-1)^n(\tfrac{d^n}{dx^n}e^{-x^2})' \\ &= (-1)^{n+1}(\tfrac{d}{dx})^{n+1}(H_n(x)e^{-x^2}) \\ &= (-1)\,H_{n+1}(x)e^{-x^2}. \end{aligned}$$

□

The next two propositions are devoted to recurrence relations satisfied by of Hermite polynomials and the generating functions of these polynomials.

Proposition 2.3.6.

$$H_{n+1}(x) = 2x\,H_n(x) - 2n\,H_{n-1}(x). \tag{2.35}$$

Proof. A direct computation from proposition 2.3.4 gives

$$\begin{aligned} 0 &= H''_{n+1}(x) - 2x\,H'_{n+1}(x) + 2(n+1)H_{n+1}(x) \\ &= 2(n+1)2n\,H_{n-1}(x) - 2x2(n+1)H_n(x) + 2(n+1)H_{n+1}(x). \end{aligned}$$

So we get the desired result. □

Proposition 2.3.7. *We have*

$$\exp(-t^2 + 2tx) = \sum_{k=0}^{+\infty} H_k(x)\frac{t^k}{k!}. \tag{2.36}$$

Proof.

$$\begin{aligned}\exp(-t^2+2tx) &= \exp(x^2)\sum_{k=0}^{+\infty}\left(\tfrac{d}{dx}\right)^k \exp(-(t-x)^2)|_{t=0}\tfrac{t^k}{k!}\\ &= \exp(x^2)\sum_{k=0}^{+\infty}\left(\tfrac{d}{dx}\right)^k \exp(-x^2)(-1)^k\tfrac{t^k}{k!}\\ &= \sum_{k=0}^{+\infty} H_k(x)\tfrac{t^k}{k!}.\end{aligned}$$

□

Corollary 2.3.8. *In the particular case that t in proposition 2.3.7 equals ± 1, $\pm i$ respectively, we get these formulas*

$$\frac{1}{e}\sinh 2x = \sum_{k=0}^{+\infty}\frac{1}{(2k+1)!}H_{2k+1}(x), \tag{2.37}$$

$$\frac{1}{e}\cosh 2x = \sum_{k=0}^{+\infty}\frac{1}{2k!}H_{2k}(x), \tag{2.38}$$

$$e\sin 2x = \sum_{k=0}^{+\infty}(-1)^k\frac{1}{(2k+1)!}H_{2k+1}(x), \tag{2.39}$$

$$e\cos 2x = \sum_{k=0}^{+\infty}(-1)^k\frac{1}{2k!}H_{2k}(x). \tag{2.40}$$

The proposition 2.3.6 can be read as $H_1(x)H_n(x) = H_{n+1}(x) + 2nH_{n-1}(x)$. In general, any product of two Hermite polynomials can be represented as a summation of finitely many other Hermite polynomials. The details are given in the next proposition and its proof is very interesting to us.

Proposition 2.3.9. *We have*

$$H_m(x)H_n(x) = \sum_{k=0}^{\min(m\ n)} 2^k k!\binom{m}{k}\binom{n}{k}H_{m+n-2k}(x). \tag{2.41}$$

Proof. Let us introduce the integrals $I_{mnl} = \int_{-\infty}^{+\infty} H_m(x)\,H_n(x)\,H_l(x)e^{-x^2}dx$, which enable us to write the product as

$$H_m(x)H_n(x) = \sum_{l=0}^{+\infty} \frac{\int_{-\infty}^{+\infty} H_m(x)\, H_n(x)\, H_l(x) e^{-x^2} dx}{\int_{-\infty}^{+\infty} H_l(x)\, H_l(x) e^{-x^2} dx} H_l(x)$$
$$= \sum_{l=0}^{+\infty} \frac{I_{mnl}}{2^l l! \sqrt{\pi}} H_l(x).$$

So we only need to calculate I_{mnl}. In fact, all these integrals can be evaluated explicitly, as in the proposition below. □

Proposition 2.3.10. *$I_{mnl} \neq 0$ can only occur in case that $m+n+l$ is even, when we will have (let $s = (m+n+l)/2$)*

$$I_{mnl} = \begin{cases} \frac{2^s \sqrt{\pi} m! n! l!}{(s-m)!(s-n)!(s-l)!} & \text{if } s \geq \max(m,n,l), \\ 0 & \text{otherwise.} \end{cases} \tag{2.42}$$

Before evaluating the integral I_{mnl}, let us see how it can help us in writing the product as the summation of Hermite polynomials. According to proposition 2.3.10, I_{mnl} is equal to zero unless $l \equiv m+n (\bmod 2)$ and $|m-n| \leq l \leq m+n$. Let $k = \frac{m+n-l}{2}$ be an integer, then this is equivalent to say that $I_{mnl} \neq 0$ only occurs at $0 \leq k \leq \min(m,n)$. So further computations give

$$H_m(x)H_n(x) = \sum_{l=0}^{+\infty} \frac{I_{mn(m+n-2k)}}{2^{m+n-2k}(m+n-2k)!\sqrt{\pi}} H_{m+n-2k}(x)$$
$$= \sum_{l=0}^{+\infty} \frac{2^{m+n-k}(m+n-2k)! m! n! \sqrt{\pi}}{2^{m+n-2k}(m+n-2k)! k! (m-k)! (n-k)! \sqrt{\pi}} H_{m+n-2k}(x)$$
$$= \sum_{k=0}^{\min(m,n)} 2^k k! \binom{m}{k} \binom{n}{k} H_{m+n-2k}(x).$$

The only thing left is to complete the proof for the proposition 2.3.10.

Proof. Without loss of generality, assume that l is the smallest integer of the set

$\{m, n, l\}$. Then

$$
\begin{aligned}
I_{mnl} &= \int_{-\infty}^{+\infty} H_m(x)\, H_n(x)\, H_l(x) e^{-x^2} dx \\
&= \int_{-\infty}^{+\infty} H_m(x)\, H_n(x) \frac{d}{dx}(-H_{l-1}(x) e^{-x^2}) dx \\
&= \int_{-\infty}^{+\infty} (mH_{m-1}(x)\, H_n(x) + nH_m(x)\, H_{n-1}(x))\, H_{l-1}(x) e^{-x^2} dx \\
&= 2mI_{(m-1)n(l-1)} + 2nI_{m(n-1)(l-1)} \\
&= 2^2 m(m-1) I_{(m-2)n(l-2)} + 2 \cdot 2^2 mn I_{(m-1)(n-1)(l-2)} + 2^2 n(n-1) I_{m(n-2)(l-2)} \\
&= \cdots \\
&= \sum_{j_1+j_2=l}^{+\infty} 2^l \binom{l}{j_1} \binom{m}{m-j_1} \binom{n}{n-j_2} I_{(m-j_1)(n-j_2)0}.
\end{aligned}
$$

According to the orthogonality of Hermite functions, we know that $I_{(m-j_1)(n-j_2)0} \neq 0$ can only happen when $m - j_1 = n - j_2$ and $j_1 + j_2 = l$, that is, when. $j_1 = s - n$ and $j_2 = s - m$. Therefore we have the result that

$$
\begin{aligned}
I_{mnl} &= \frac{2^l l!}{(s-n)!(s-m)!} \frac{m!}{(s-l)!} \frac{n!}{(s-l)!} 2^{s-l} (s-l)! \sqrt{\pi} \\
&= \frac{2^s \sqrt{\pi} m! n! l!}{(s-m)!(s-n)!(s-l)!}.
\end{aligned}
$$

□

At the end of this section we shall introduce some other propositions concerning Hermite functions, which may not appear in the rest of this thesis but are so important that we feel it very necessary to mention. These include the relationship between Hermite functions and eigenfunctions of Fourier operators, the role of Hermite functions in probability theory related to the normal distribution, and the integral interpretation of Hermite functions.

Proposition 2.3.11. *[Webpage2] Let $\mathcal{F}$ be the Fourier transformation defined on $L^2(\mathbb{R})$, which is the unique continuous extension of the operator*

$$\mathcal{F}f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ix\cdot\xi} f(x)\, dx, \qquad f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R}). \tag{2.43}$$

Then the Hermite functions $\{h_k(x) \,:\, k = 0, 1, 2, \cdots\}$ are eigenfunctions of the Fourier transformation, and they satisfy

$$Eigen\{(-i)^k, \mathcal{F}\} = span\{\, h_{4m+k}(x),\, m = 0, 1, 2, \cdots, k = 0, 1, 2, 3.\}. \tag{2.44}$$

Proof. At first it is easy to check the case of $h_0(x) = \frac{\exp\{-x^2/2\}}{\sqrt[4]{\pi}}$:

$$\begin{aligned}
\mathcal{F}h_0(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ix\,\xi} \frac{\exp\{-x^2/2\}}{\sqrt[4]{\pi}} dx \\
&= \frac{1}{\sqrt{2\pi}\sqrt[4]{\pi}} \int_{-\infty}^{+\infty} \exp\{-x + i\xi)^2/2\} dx e^{-\frac{\xi^2}{2}} \\
&= \frac{1}{\sqrt[4]{\pi}} e^{-\frac{\xi^2}{2}}.
\end{aligned}$$

So $h_0(x)$ is an eigenfunction of Fourier transformation with eigenvalue 1.

Noticing that $h_k(x) = \frac{1}{\sqrt{2^k k!}}(-\frac{d}{dx} + x)^k\, h_0(x)$, we can further calculate

$$\begin{aligned}
\mathcal{F}h_k(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ix\cdot\xi} \frac{1}{\sqrt{2^k k!}} (-\frac{d}{dx} + x)^k\, h_0(x)\, dx \\
&= (2\pi)^{-\frac{1}{2}} \frac{1}{\sqrt{2^k k!}} (-i)^k\, (-\frac{d}{d\xi} + \xi)^k \int_{-\infty}^{+\infty} e^{-ix\cdot\xi} h_0(x) dx \\
&= (-i)^k \frac{1}{\sqrt{2^k k!}} (-\frac{d}{d\xi} + \xi)^k\, h_0(\xi) \\
&= (-i)^k h_k(x).
\end{aligned}$$

Since the Fourier transformation is a unitary operator on function space $L^2(\mathbb{R})$ by Parceval's identity, and $\{h_k(x),\, k = 0, 1, 2, 3, \cdots\}$ is an orthonormal basis of the same space, the operator $\mathcal{F}$ can be represented as an infinite diagonal unitary matrix

$$\mathcal{F} = (a_{jk}) = (\delta_{jk}(-i)^k). \qquad j, k = 0, 1, 2, \cdots. \tag{2.45}$$

Therefore

$$Eigen\{(-i)^k, \mathcal{F}\} = span\{\, h_{4m+k}(x),\; m = 0, 1, 2, \cdots,\; k = 0, 1, 2, 3.\}.$$

□

Remark 2.3.12. *In higher dimensional cases, similar results are also true. The eigenvalues of the Fourier transformation are still* $\{\pm 1, \pm i\}$, *and the only difference is that the eigenfunctions are Hermite functions in high dimensional cases, that is,* $\{h_{k_1}(x_1) \otimes h_{k_2}(x_2) \otimes \cdots \otimes h_{k_n}(x_n), \quad k_j = 0, 1, 2, \cdots\}$, *which form orthonormal bases in high dimensional cases.*

Next we will see Hermite polynomials also play an important role in the probability theory.

Proposition 2.3.13. *If we have a normal distribution* $p(x) = e^{-(x-\mu)^2}/\sqrt{\pi}$, *then the expectation of the Hermite polynomials are*

$$E(H_n(x)) = 2^n \mu^n. \tag{2.46}$$

Proof. It follows from proposition 2.3.7.

$$\begin{aligned}
E(H_n(x)) &= \int_{-\infty}^{+\infty} H_n(x) \frac{1}{\sqrt{\pi}} e^{-(x-\mu)^2} dx \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} \left(\sum_{k=0}^{+\infty} H_k(x) \frac{\mu^k}{k!} \right) H_n(x) dx \\
&= \sum_{k=0}^{+\infty} \left(\frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} \delta_{kn} 2^n n! \frac{\mu^k}{k!} \right) \\
&= 2^n \mu^n.
\end{aligned}$$

□

Finally let us state that Hermite polynomials have their form of integral interpretation.

Proposition 2.3.14.

$$H_n(x) = 2^n \int_{-\infty}^{+\infty} (x + it)^n \frac{e^{-t^2}}{\sqrt{\pi}} dt. \tag{2.47}$$

Proof. A direct computation with the help of explicit formulas of Hermite polynomials will work.

$$\begin{aligned}
&2^n \int_{-\infty}^{+\infty} (x + it)^n \frac{e^{-t^2}}{\sqrt{\pi}} dt \\
&= 2^n \int_{-\infty}^{+\infty} \sum_{k=0, k\, even}^{+\infty} \left(\binom{n}{k} x^{n-k} i^k t^k \right) \frac{e^{-t^2}}{\sqrt{\pi}} dt \\
&= 2^n \int_{-\infty}^{+\infty} \sum_{k=0}^{[\frac{n}{2}]} \left(\frac{n!}{(2k)!(n-2k)!} x^{n-2k} (-1)^k t^{2k} \right) \frac{e^{-t^2}}{\sqrt{\pi}} dt \\
&= \sum_{k=0}^{[\frac{n}{2}]} \left(\frac{2^n n!}{(2k)!(n-2k)!} x^{n-2k} (-1)^k \frac{\Gamma(k+1/2)}{\Gamma(1/2)} \right) \\
&= \sum_{k=0}^{[\frac{n}{2}]} \left(\frac{n!}{k!(n-2k)!} (2x)^{n-2k} (-1)^k \right) \\
&= H_n(x).
\end{aligned}$$

□

2.4 Some estimates of integrals of products of Hermite functions

Let us introduce the integral $C_{klmn} = \int_{-\infty}^{+\infty} h_k(x)h_l(x)h_m(x)h_n(x)\,dx$, which can be viewed as a symmetric four-tensor. It plays an important role in this thesis to obtain a good understanding of this four tensor.

In this section, we will pay attention to estimates of the entries C_{klmn}. A very special case, $k = l = m = n$, has been studied by G.Freud and G.Németh[F-N]. Their method depends heavily on the knowledge of special functions such as the Gamma and hypergeometric functions. A quick review will be provided in this section. Since their method can only work for estimate of C_{kkkk}, we will also provide our new method for estimate on C_{kkll}, which will recover the result by G.Freud and G.Németh if $k = l$. One advantage of our method is that it only depends on the properties listed in the last section, and doesn't require the knowledge of other special functions. Now let us begin with the result of G.Freud and G.Németh.

Theorem 2.4.1. *[F-N] We have*

$$C_{nnnn} = 2^{-1/2}\pi^{-3/2}\int_0^1 t^n(1-t)^{-1/2}\,{}_2F_1\left(\frac{1}{2},\frac{1}{2};1;t\right)dt, \tag{2.48}$$

where

$${}_2F_1\left(\frac{1}{2},\frac{1}{2};1;t\right) = \sum_{j=0}^{+\infty}\left(\frac{\Gamma(j+1/2)}{\Gamma(1/2)j!}\right)^2 t^j. \tag{2.49}$$

Corollary 2.4.2. *The sequence* $\{C_{nnnn}\}$ *is totally monotone and*

$$C_{nnnn} = 2^{-1/2}\pi^{-2}n^{-1/2}\log n + O(n^{-1/2}), \qquad \text{as } n \to +\infty. \tag{2.50}$$

Remark 2.4.3. *A sequence $c_0, c_1, c_2, \cdots$ of real numbers is called totally monotone if $\Delta^m c_n \geq 0$, $(m, n = 0, 1, 2, \cdots)$, where*

$$\Delta^m c_n = c_n - \binom{m}{1} c_{n+1} + \binom{m}{2} c_{n+1} - \cdots + (-1)^m \binom{m}{m} c_{n+m}.$$

Hausdorff showed that for every totally monotone sequence $c_0, c_1, c_2, \cdots$ there exists (essentially uniquely) a monotone nondecreasing real function $\phi(u)$, $0 \leq u \leq 1$, such that

$$c_n = \int_0^1 u^n d\phi(u), \quad n = 0, 1, 2, \cdots.$$

Conversely, if $\phi(u)$ is a monotone nondecreasing bounded real function on the interval $0 \leq u \leq 1$, then $\Delta^m c_n = \int_0^1 (1-u)^m u^n d\phi(u) \geq 0$, $m, n = 0, 1, 2, \cdots$, so that $c_0, c_1, c_2, \cdots$ is totally monotone.

We are most interested in the claim in the second part of the corollary above. A sketch of the proof towards this result is provided. The authors first use Mehler's generating series [E] to get

$$\begin{aligned}\sum_{m,n=0}^{+\infty} C_{mmnn} u^m v^n &= \frac{1}{\sqrt{2\pi}} \left((1-u)(1-v)(1-uv)\right)^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{2\pi}} \sum_{\lambda,\mu,\nu=0}^{+\infty} \frac{\Gamma(\lambda+1/2)\Gamma(\mu+1/2)\Gamma(\nu+1/2)}{\Gamma(1/2)\lambda!\Gamma(1/2)\mu!\Gamma(1/2)\nu!} (uv)^\lambda u^\mu v^\nu.\end{aligned}$$

Although general terms C_{mmnn} can't be represented in a clean formula by comparing the coefficients on both sides, it can be done for C_{nnnn}

$$C_{nnnn} = \frac{1}{\sqrt{2\pi}} \sum_{j=0}^{+\infty} \left(\frac{\Gamma(j+1/2)}{\Gamma(1/2)j!}\right)^2 \frac{\Gamma(n-j+1/2)}{\Gamma(1/2)(n-j)!}. \tag{2.51}$$

By making use of hypergeometric functions, the equation (2.51) can be written as

$$\begin{aligned}C_{nnnn} &= 2^{-1/2}\pi^{-3/2} \frac{(-1)^n}{n!} \int_0^1 t^n (1-t)^{-\frac{1}{2}} \left(\left(\frac{d}{dt}\right)^n t^{-\frac{1}{2}} (1-t)^n\right) dt \\ &= 2^{-1/2}\pi^{-3/2} \int_0^1 t^n (1-t)^{-\frac{1}{2}} {}_2F_1\left(\frac{1}{2}, \frac{1}{2}; 1; t\right) dt,\end{aligned}$$

which is just the result in theorem 2.4.1. Using Sterling's asymptotic series, we obtain

$$ {}_2F_1\left(\frac{1}{2},\frac{1}{2};1:t\right)=\frac{1}{\pi}\log\frac{1}{1-t}+O(1), \tag{2.52} $$

which holds uniformly in $t\in[0,1)$. A further computation concludes that

$$ \begin{aligned} C_{nnnn} &= 2^{-1/2}\pi^{-5/2}\int_0^1 t^n(1-t)^{-\frac{1}{2}}\log\frac{1}{1-t}dt+O(1)\int_0^1 t^n(1-t)^{-\frac{1}{2}}dt \\ &= 2^{-1/2}\pi^{-2}n^{-1/2}\log n+O(n^{-1/2}). \end{aligned} $$

In the process above, it is a very important step to establish the formula (2.51). Below we will provide our methods for estimate of C_{kkll}, and a similar formula will be obtained. First of all, let us review the fact that C_{mn00} can be computed exactly [E-M-O-T].

Proposition 2.4.4.

$$ \int_{-\infty}^{+\infty} H_m(x)H_n(x)e^{-2x^2}dx \tag{2.53} $$

$$ =\begin{cases} 0 & if\, m\not\equiv n(\text{mod}\,2), \\ (-1)^{\frac{m-n}{2}}2^{\frac{m+n-1}{2}}\Gamma(\frac{m+n+1}{2}) & if\, m\equiv n(\text{mod}\,2). \end{cases} \tag{2.54} $$

Proof. The first case is trivial. Let us now work in the second case $m\equiv n(\text{mod}\,2)$. By proposition 2.3.7 in the last section, we can introduce two extra parameter s and t to get

$$ \begin{aligned} \exp(-t^2+2tx) &= \sum_{m=0}^{+\infty} H_m(x)\frac{t^m}{m!}, \\ \exp(-s^2+2sx) &= \sum_{n=0}^{+\infty} H_n(x)\frac{s^n}{n!}. \end{aligned} $$

Their product is

$$ \exp(-(t^2+s^2)+2(t+s)\,x)=\sum_{m,n=0}^{+\infty} H_m(x)H_n(x)\frac{t^m s^n}{m!n!}, $$

and in particular,

$$\int_{-\infty}^{+\infty} \exp(-(t^2+s^2)+2(t+s)\,x)\exp(-2x^2)dx$$

$$= \sum_{m,n=0}^{+\infty} \left(\int_{-\infty}^{+\infty} H_m(x)H_n(x)\,e^{-2x^2}dx\right)\frac{t^m s^n}{m!n!}. \tag{2.55}$$

We can also write this integral in another formula by using the same proposition

$$\int_{-\infty}^{+\infty} \exp(-(t^2+s^2)+2(t+s)\,x)\exp(-2x^2)dx$$

$$= \int_{-\infty}^{+\infty} \left(\sum_{k=0}^{+\infty} H_k(\sqrt{2}x)\frac{(t+s)^k}{2^{k/2}k!}\right)\exp(-2x^2)dx\exp\left(-(t-s)^2/2\right).$$

Noticing that functions $H_k(\sqrt{2}x)$, $k=0,1,2,\cdots$, are orthogonal to each other with respect to weight function e^{-2x^2}, we can proceed as

$$\int_{-\infty}^{+\infty} \exp(-(t^2+s^2)+2(t+s)\,x)\exp(-2x^2)dx \tag{2.56}$$

$$= \exp\left(-(t-s)^2/2\right)\int_{-\infty}^{+\infty} \exp(-2x^2)dx \tag{2.57}$$

$$= \sqrt{\frac{\pi}{2}}\sum_{j=0}^{+\infty}\frac{(-1)^j}{j!2^j}(t-s)^{2j}. \tag{2.58}$$

By comparing equation (2.55) and (2.58), we deduce that

$$\begin{aligned}\int_{-\infty}^{+\infty} H_m(x)H_n(x)e^{-2x^2}dx &= \sqrt{\frac{\pi}{2}}m!n!C(t^m s^n;\exp\left(-(t-s)^2/2\right)\\ &= \sqrt{\frac{\pi}{2}}(m+n)!\frac{(-1)^{\frac{m-n}{2}}}{(\frac{m+n}{2})!2^{\frac{m+n}{2}}}\\ &= (-1)^{\frac{m-n}{2}}2^{\frac{m+n-1}{2}}\Gamma(\frac{m+n+1}{2}).\end{aligned}$$

In the process above, we use $C(t^m s^n;\exp\left(-(t-s)^2/2\right)$ to denote the coefficient of term $t^m s^n$ in the Taylor series of the entire function $\exp\left(-(t-s)^2/2\right)$ at the point $(0,0)$. □

Now we are in a good position to compute out C_{mn00}, writing it as $C_{mn00} = \frac{\int_{-\infty}^{+\infty} H_m(x)H_n(x)e^{-2x^2}dx}{\sqrt{2^m m!}\sqrt{2^n n!}\pi}$, from which the corollary follows.

Corollary 2.4.5. *We have*

$$C_{mn00} = \begin{cases} 0 & if\, m \neq n (\text{mod}\, 2), \\ (-1)^{\frac{m-n}{2}} \frac{\Gamma(\frac{m+n+1}{2})}{\sqrt{2}\pi\sqrt{m!n!}} & if\, m \equiv n (\text{mod}\, 2). \end{cases}$$

After computing out what is C_{mn00}, let us define the coefficients C_{mn}^p as numbers defined by the relation $h_m(x)h_n(x) = \sum_{p=0}^{+\infty} C_{mn}^p (h_p(\sqrt{2}x)2^{1/4})$. Since $\{h_p(x),\ p = 0, 1, 2, \cdots\}$ is an orthonormal basis of the function space $L^2(\mathbb{R})$, a scaling gives us another orthonormal basis as $\{h_p(\sqrt{2}x)2^{1/4},\ p = 0, 1, 2, \cdots\}$. One advantage to introduce these coefficients is that we can interpret C_{klmn} easily, as $C_{klmn} = \sum_{p=0}^{+\infty} C_{kl}^p C_{mn}^p$. In the following theorem, we will see that for m, n given, C_{mn}^p is nonzero only for finitely many p, and we can further compute out C_{mn}^p exactly in some special cases. Here let us use the notation $C(x^p, p(x))$ that denotes the coefficient of term x^p in the polynomial $p(x)$.

Theorem 2.4.6. *For given m, n, $C_{mn}^p \neq 0$ only occurs at $p = m + n - 2r$, where r is an integer satisfying $0 \leq r \leq [\frac{m+n}{2}]$. And in that case,*

$$C_{mn}^p = \frac{(m+n-2r)!\Gamma(r+1/2)(-1)^r C(x^{2r}; (1+x)^m(1-x)^n)}{2^{1/4}2^{\frac{m+n}{2}-r}\pi^{3/4}\sqrt{m!n!(m+n-2r)!}}. \tag{2.59}$$

To prove this theorem, we need introduce another sequence of integrals $I_{mn;p} = \int_{-\infty}^{+\infty} H_m(x)H_n(x)H_p(\sqrt{2}x)e^{-2x^2}dx$. It is easy to find the relationship between C_{mn}^p and $I_{mn,p}$ as

$$\begin{aligned} C_{mn}^p &= \int_{-\infty}^{+\infty} h_m(x)h_n(x)h_p(\sqrt{2}x)2^{1/4}dx \\ &= \frac{I_{mn,p}2^{1/4}}{2^{\frac{m+n+p}{2}}\sqrt{m!n!p!}\pi^{3/4}}. \end{aligned} \tag{2.60}$$

We will see that the result of the theorem 2.4.6 is equivalent to the following lemma, which will be proved.

Lemma 2.4.7. *We have*

$$I_{mn;p} = \begin{cases} 0 \\ \quad if\ p > m+n\ or\ p \neq m+n (\bmod 2), \\ \\ 2^{\frac{m+n-1}{2}} \Gamma(\frac{m+n-p+1}{2}) p! (-1)^{\frac{m-n-p}{2}} C(x^p, (1+x)^m (1-x)^n) \\ \quad if\ p \leq m+n\ and\ p = m+n (\bmod 2). \end{cases}$$

Proof. It is trivial to find $I_{mn;p} = 0$ when $p \neq m+n (\bmod 2)$. If $p > m+n$, then as a polynomial of order less than p, $H_m(x)H_n(x)$ can be written as the linear combination of the scaled Hermite polynomials $H_l(\sqrt{2}x)$, which are orthogonal to $H_p(\sqrt{2}x)$ in the meaning of integration with weighted function e^{-2x^2}. Thus we must have $I_{mn;p} = 0$ for those $p > m+n$. The most interesting case is the next one, the computation of $I_{mn;p}$ when $p = m+n (\bmod 2)$.

$$\begin{aligned} I_{mn,p} &= \int_{-\infty}^{+\infty} H_m(x)H_n(x) d(-H_{p-1}(\sqrt{2}x)e^{-2x^2}) \frac{1}{\sqrt{2}} \\ &= \sqrt{2}(mI_{(m-1)n;(p-1)} + nI_{m(n-1);(p-1)}) \\ &= \left(\sqrt{2}\right)^2 \left(m(m-1)I_{(m-2)n;(p-2)} + 2mnI_{(m-1)(n-1);(p-2)} + n(n-1)I_{m(n-2),(p-2)}\right) \\ &= \cdots \\ &= \left(\sqrt{2}\right)^p \sum_{\substack{d_1+d_2=p \\ 0\leq d_1\leq m,\ 0\leq d_2\leq n}} \frac{p!}{d_1!d_2!} \frac{m!}{(m-d_1)!} \frac{n!}{(n-d_2)!} I_{(m-d_1)(n-d_2);0}. \end{aligned}$$

By proposition 2.4.4, $I_{(m-d_1)(n-d_2);0}$ can be computed out exactly, so

$$I_{mn;p} = \sum_{\substack{d_1+d_2=p \\ 0\le d_1\le m,\ 0\le d_2\le n}} 2^{p/2} p!\Gamma(\frac{m+n-p+1}{2}) 2^{\frac{m+n-p-1}{2}} (-1)^{\frac{(m-d_1)-(n-d_2)}{2}}$$

$$\frac{m!}{d_1!(m-d_1)!}\frac{n!}{d_2!(n-d_2)!}$$

$$= 2^{\frac{m+n-1}{2}} p!\Gamma(\frac{m+n-p+1}{2})(-1)^{\frac{m-n-p}{2}}$$

$$\sum_{\substack{d_1+d_2=p \\ 0\le d_1\le m,\ 0\le d_2\le n}} \left((-1)^{d_2}\frac{m!}{d_1!(m-d_1)!}\frac{n!}{d_2!(n-d_2)!}\right)$$

$$= 2^{\frac{m+n-1}{2}} p!\Gamma(\frac{m+n-p+1}{2})(-1)^{\frac{m-n-p}{2}} C(x^p; (1+x)^m(1-x)^n).$$

□

It is easy to verify theorem 2.4.6 after we have proved lemma 2.4.7. We just need to use the relationship between C^p_{mn} and $I_{mn;p}$ (see 2.60), together with the following two basic propositions for polynomials,

$$C(x^p; (1+x)^m(1-x)^n) = (-1)^n C(x^{m+n-p}; (1+x)^m(1-x)^n), \tag{2.61}$$

$$C(x^p; (1+x)^m(1-x)^n) = (-1)^p C(x^p; (1-x)^m(1+x)^n). \tag{2.62}$$

In general, $C(x^p; (1-x)^m(1+x)^n)$ is hard to figure out. But the situation is very different for such a special case $m = n$,

$$C^{2m-2r}_{mm} = \frac{(2m-2r)! \cdot \Gamma(r+1/2) C(x^{2r};\ (1-x^2)^m)(-1)^r}{2^{1/4}\pi^{3/4} 2^{m-r} m! \sqrt{(2m-2r)!}}$$

$$= 2^{-\frac{1}{4}}\pi^{-\frac{1}{4}}\sqrt{\frac{(2m-2r)!}{2^{2m-2r}(m-r)!(m-r)!}\frac{(2r)!}{2^{2r}r!r!}} \tag{2.63}$$

$$= (2\pi)^{-\frac{1}{4}}\frac{\Gamma(r+1/2)}{\Gamma(1/2)r!}\left(\frac{\Gamma(m-r+1/2)}{\Gamma(1/2)(m-r)!}\right)^{\frac{1}{2}}. \tag{2.64}$$

Theorem (2.4.6) gives the following.

Corollary 2.4.8. *For Given m, $C^p_{mm} \neq 0$ only occurs at $p = 2r$, where r is an integer satisfying $0 \leq r \leq m$. And in that case, we have*

$$C^{2r}_{mm} = (2\pi)^{-\frac{1}{4}} \frac{\Gamma(m-r+1/2)}{\Gamma(1/2)(m-r)!} \left(\frac{\Gamma(r+1/2)}{\Gamma(1/2)r!}\right)^{\frac{1}{2}}. \tag{2.65}$$

According to theorem 2.4.6, every nontrivial entries C_{klmn} $(k+l = m+n (\bmod 2))$ in the four tensor can be represented as finite summations as follows

$$\begin{aligned} C_{klmn} &= \sum_{p=m+n(\bmod 2)}^{\min(m+n,k+l)} C^p_{kl} C^p_{mn} \\ &= (-1)^{\frac{k+l+m+n}{2}} 2^{-1/2} \pi^{-3/2} (m!n!k!l!)^{-1/2} \sum_{p=m+n(\bmod 2)}^{\min(m+n,k+l)} \frac{p!}{2^p} (-1)^p \\ &\Gamma\left(\frac{m+n-p+1}{2}\right) \Gamma\left(\frac{k+l-p+1}{2}\right) \\ &C(x^{m+n-p}; (1+x)^m (1-x)^n) C(x^{k+l-p}; (1+x)^k (1-x)^l). \end{aligned} \tag{2.66}$$

In particular, we have

$$C_{mmnn} = (2\pi)^{-\frac{1}{2}} \sum_{r=0}^{\min(m,n)} \frac{\Gamma(r+1/2)}{\Gamma(1/2)r!} \frac{\Gamma(m-r+1/2)}{\Gamma(1/2)(m-r)!} \frac{\Gamma(n-r+1/2)}{\Gamma(1/2)(n-r)!}. \tag{2.67}$$

When $m = n$, the equation above is just the formula (2.51), which appears in the paper [F-N].

Although it turns out that C_{klmn} in the formula (2.66) is still hard to be analyzed in general settings, we can use it to get good estimates on C_{mmnn}. Before we state our results for this, let us introduce two useful bits of notation. We denote $\langle k \rangle$ by $\sqrt{k^2+1}$ and use the symbol $a \approx b$ to describe the relationship of a and b as there exist two universal constants $c_1, c_2 > 0$ such that $c_1 b \leq a \leq c_2 b$. Similarly, we use $a \lesssim b$ ($a \gtrsim b$) with the meaning that there exist constants $c > 0$ such that $a \leq cb$ ($a \geq cb$). Now we can have such an estimate on C_{mmnn}.

Theorem 2.4.9. *For nonnegative integers $m \geq n$, we have*

$$0 < C_{mmnn} \lesssim \langle m \rangle^{-\frac{1}{2}} \ln^{1/2} \langle m \rangle \ln^{1/2} \langle n \rangle . \tag{2.68}$$

If m or n is zero, then the part $\ln^{1/2} \langle m \rangle \ln^{1/2} \langle n \rangle$ can be completely removed.

Proof. Let us start with formula (2.67). By Sterling's formula $n! = n^n e^{-n} \sqrt{2\pi n} e^{\theta_n} \approx n^n e^{-n} \sqrt{2\pi n}$, $\frac{1}{12n+1} < \theta_n < \frac{1}{12n}$(see [A-S]), we can rewrite equation (2.65) in the form of

$$C_{mm}^{2r} \approx \begin{cases} \frac{1}{\sqrt{m} 2^{1/4} \pi^{3/4}} & if\ r = 0, \\ \frac{1}{(2m)^{1/4} \sqrt{\pi}} & if\ r = m, \\ \frac{1}{r^{1/4}(m-r)^{1/2} 2^{1/4} \pi} & if\ 1 \leq r \leq m-1. \end{cases} \tag{2.69}$$

Formula (2.67) can therefore be written as

$$\begin{aligned} C_{mmnn} &= \sum_{r=0}^{n} C_{mm}^{2r} C_{nn}^{2r} \\ &\approx 2^{-1/2} \pi^{-2} \left(\frac{\sqrt{\pi}}{\sqrt{m}\sqrt{n}} + \frac{\sqrt{\pi}}{\sqrt{n}\sqrt{m-n}} + \sum_{r=1}^{n-1} \frac{1}{\sqrt{r}\sqrt{m-r}\sqrt{n-r}} \right). \end{aligned} \tag{2.70}$$

First let us see the case $m = n$. We have

$$C_{mmmm} = 2^{-1/2} \pi^{-2} \left(\frac{\sqrt{\pi}}{m} + \frac{\pi}{\sqrt{m}} + \sum_{r=1}^{m-1} \frac{1}{\sqrt{r}(m-r)} \right).$$

Noticing that the function $f(r) := \frac{1}{\sqrt{r}(m-r)}$ is decreasing on $(0, \frac{m}{3}]$ and then increasing on $[\frac{m}{3}, m)$, we can get

$$\begin{aligned} \sum_{r=1}^{n-1} \frac{1}{\sqrt{r}(m-r)} &\leq \int_0^{m-1} \frac{1}{\sqrt{r}(m-r)} dr + \left. \frac{1}{\sqrt{r}(m-r)} \right|_{r=m-1} - \left. \frac{1}{\sqrt{r}(m-r)} \right|_{r=\frac{m}{3}} \\ &\leq \frac{1}{\sqrt{m}} (\ln m + \ln 4) + \frac{1}{\sqrt{m-1}} \end{aligned}$$

and

$$\sum_{r=1}^{n-1}\frac{1}{\sqrt{r}(m-r)} \geq \int_1^{m-1}\frac{1}{\sqrt{r}(m-r)}dr + \frac{1}{\sqrt{r}(m-r)}\bigg|_{r=\frac{m}{3}}$$
$$\geq \frac{1}{\sqrt{m}}(\ln m + \ln\frac{9}{4}) + \frac{3\sqrt{3}}{2}\frac{1}{m\sqrt{m}}.$$

Thus we conclude that

$$C_{mmmm} \approx \frac{1}{\sqrt{m}}\ln m \approx \langle m\rangle^{-\frac{1}{2}}\ln\langle m\rangle\,. \tag{2.71}$$

In doing so, we have already recovered the result in [F-N] without use of properties of special functions. We proceed for the case $m > n$.

When $n = 0$, according to corollary 2.4.5, we have

$$C_{mm00} = \frac{\Gamma(m+1/2)}{\sqrt{2}\pi m!} \approx \frac{1}{\sqrt{m}}\frac{1}{\sqrt{2\pi}}. \tag{2.72}$$

When $n = 1$, we have

$$\begin{aligned} C_{mm11} &= C^0_{mm}C^0_{11} + C^2_{mm}C^2_{11} \\ &\approx 2^{-\frac{1}{2}}\pi^{-\frac{3}{2}}\frac{2}{\sqrt{m}}. \end{aligned} \tag{2.73}$$

Finally let us check on the case of $m > n \geq 2$,

$$\begin{aligned} \sum_{r=1}^{n-1}\frac{1}{\sqrt{r}\sqrt{m-r}\sqrt{n-r}} &\leq \left(\sum_{r=1}^{n-1}\frac{1}{r(m-r)}\right)^{1/2}\left(\sum_{r=1}^{n-1}\frac{1}{n-r}\right)^{1/2} \\ &= \frac{1}{\sqrt{m}}\left((1+\frac{1}{2}+\cdots+\frac{1}{n-1}) + (\frac{1}{m-n+1}+\cdots+\frac{1}{m-1})\right)^{1/2} \\ &\quad (1+\frac{1}{2}+\cdots+\frac{1}{n-1})^{1/2} && (2.74) \\ &\lesssim \frac{1}{\sqrt{m}}\ln^{1/2}\left((m-1)\frac{n-1}{m-n+1}\right)\ln^{1/2}(n-1) \\ &\lesssim \langle m\rangle^{-\frac{1}{2}}\ln^{1/2}\langle m\rangle\ln^{1/2}\langle n\rangle\,. && (2.75) \end{aligned}$$

By using the last result in the equality (2.70), we can conclude the statement in the theorem. □

Remark 2.4.10. *This estimate turns out to be very accurate up to the order of* $\langle m \rangle$, *and the logarithmitic term cannot be removed in the general cases. To see this, we only need compute the term* C_{mmnn} *in such two special cases. One is* C_{mm00}, *which is equal to* $\frac{\Gamma(m+1/2)}{\sqrt{2}\pi m!} = \frac{1}{\sqrt{m}}\frac{1}{\sqrt{2\pi}} + o(\frac{1}{\sqrt{m}})$ *as* $m \to +\infty$. *The other one is* $C_{mm(m-1)(m-1)}$. *We have*

$$\begin{aligned} C_{mm(m-1)(m-1)} &\approx 2^{-\frac{1}{2}}\pi^{-2}\left(\frac{\sqrt{\pi}}{\sqrt{m}\sqrt{m-1}} + \frac{\sqrt{\pi}}{\sqrt{m-1}} + \sum_{r=1}^{m-2}\frac{1}{\sqrt{r}\sqrt{m-r}\sqrt{m-1-r}}\right) \\ &\approx 2^{-\frac{1}{2}}\pi^{-2}\left(\frac{\sqrt{\pi}}{m-1} + \frac{\pi}{\sqrt{m-1}} + \sum_{r=1}^{m-2}\frac{1}{\sqrt{r}(m-1-r)}\right) \\ &\approx \langle m \rangle^{-\frac{1}{2}} \ln^{1/2}\langle m \rangle \ln^{1/2}\langle m-1 \rangle . \end{aligned}$$

Remark 2.4.11. *It turns out that* C_{mmnn} *has another interesting estimate if we make use of the theory of elliptic integrals. Since the function* $f(r) := \frac{1}{\sqrt{r}\sqrt{m-r}\sqrt{n-r}}$ *is integrable on* $(0,n)$ *and its monotonicity periods are easy to analyze, we can find that* $C_{mmnn} \approx 2^{-\frac{1}{2}}\pi^{-2}\int_0^n \frac{1}{\sqrt{r}\sqrt{m-r}\sqrt{n-r}}dr$. *The integral part can be represented by elliptic integral (see [J]) as* $\int_0^n \frac{1}{\sqrt{r}\sqrt{m-r}\sqrt{n-r}}dr = \frac{2}{\sqrt{m}}\int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1-s^2\sin^2\theta}} = \frac{2}{\sqrt{m}}E(s)$, *where* $s = \sqrt{\frac{n}{m}} < 1$. *Thus we get a very exact estimate on* C_{mmnn}.

By using Hölder inequality, we can easily deduce from the result above an estimate for the general term C_{klmn}. Assume that $k \geq l \geq m \geq n$, then

$$\begin{aligned} |C_{klmn}| &\leq C_{kknn}^{1/2} C_{llmm}^{1/2} \\ &\lesssim \langle k \rangle^{-\frac{1}{4}} \langle l \rangle^{-\frac{1}{4}} \left(\ln\langle k \rangle \ln\langle l \rangle \ln\langle m \rangle \ln\langle n \rangle\right)^{\frac{1}{4}} . \end{aligned} \tag{2.76}$$

Here we list this result as a corollary of the theorem above.

Corollary 2.4.12. *For nonnegative integers $k \geq l \geq m \geq n$, we have*

$$|C_{klmn}| \lesssim \langle k \rangle^{-\frac{1}{4}} \langle l \rangle^{-\frac{1}{4}} \left(\ln \langle k \rangle \ln \langle l \rangle \ln \langle m \rangle \ln \langle n \rangle\right)^{\frac{1}{4}}. \tag{2.77}$$

If the smallest number n is equal to 0, then the part of $(\ln \langle k \rangle \ln \langle n \rangle)^{1/4}$ can be taken off.

We should mention that the estimate in (2.77) is not bad when these integers k, l, m, n are almost equal to each other or the integer pairs (k, n) and (l, m) are not far away. But it turns out that the entries C_{klmn} in the four tensor are very complicated, and show completely different property in some other occasions. In the next chapter, we will see that when one of the integers $\{k, l, m, n\}$ is much larger than all the others, C_{klmn} will show a fast decreasing property.

2.4.1 The estimate of C_{klmn} in a special case

In the second chapter, we have got an estimate on general coefficients C_{klmn}, which is sharp when $(k, l) = (m, n)$. But it turns out that the four tensor C_{klmn} shows very different behavior in other situations. In this section, a different method from that in the second chapter will be adopted, which is based on the generating functions for the Hermite polynomials. We will show that when one of the indices, say k, is much larger than all the others, the coefficient C_{klmn} will be rapidly decreasing.

Proposition 2.4.13. *We have*

$$C_{klmn} = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{2^p} \sqrt{k!l!m!n!}\, C\left(t^k s^l u^m v^n;\ f(t, s, u, v)\right) & \text{if } k + l + m + n \equiv 0 (\operatorname{mod} 2) \\ 0 & \text{otherwise,} \end{cases} \tag{2.78}$$

where $p = (k+l+m+n)/2$, $f(t,s,u,v)=\exp\left(\frac{(t+s+u+v)^2}{2} - (t^2+s^2+u^2+v^2)\right)$, *and* $C\left(t^k s^l u^m v^n;\ f(t,s,u,v)\right)$ *denotes the coefficient of term* $t^k s^l u^m v^n$ *in the Taylor expansion of the entire function* $f(t,s,u,v)$ *at point* $(0,0,0,0)$.

Proof. Only the first case is nontrivial. Let $H_k(x)H_l(x)H_m(x)H_n(x) = \sum_{j=0}^{\infty} B^j_{klmn}\, H_j(\sqrt{2}x)$. Then

$$C_{klmn} = \int_{-\infty}^{+\infty} B^0_{klmn}\, H_0(\sqrt{2}x)\, e^{-2x^2} dx / \left(2^{\frac{k+l+m+n}{2}} \sqrt{k!l!m!n!\pi}\right)$$
$$= B^0_{klmn} / (\sqrt{\pi 2} \cdot 2^{\frac{k+l+m+n}{2}} \sqrt{k!l!m!n!}).$$

By the property of the generating function for Hermite functions, we have

$$exp\{-t^2+2xt\} = \sum_{k=0}^{+\infty} \frac{H_k(x)}{k!} t^k$$
$$exp\{-s^2+2xs\} = \sum_{l=0}^{+\infty} \frac{H_l(x)}{l!} s^l$$
$$exp\{-u^2+2xu\} = \sum_{m=0}^{+\infty} \frac{H_m(x)}{m!} u^m$$
$$exp\{-v^2+2xv\} = \sum_{n=0}^{+\infty} \frac{H_n(x)}{n!} v^n.$$

Taking the product of these identities, we get

$$\exp\left(-(t^2+s^2+u^2+v^2)+2(t+s+u+v)x\right) \tag{2.79}$$
$$= \sum_{k,l,m,n=0}^{+\infty} \frac{H_k(x)H_l(x)H_m(x)H_n(x)}{k!l!m!n!} t^k s^l u^m v^n, \tag{2.80}$$

from which we find,

$$L.H.S. = \exp\left(-\frac{(t+s+u+v)^2}{2} + 2\frac{(t+s+u+v)}{\sqrt{2}}\sqrt{2}x\right)$$
$$\times \exp\left(\frac{(t+s+u+v)^2}{2} - (t^2+s^2+u^2+v^2)\right)$$
$$= \sum_{j=0}^{+\infty} \frac{H_j(\sqrt{2}x)}{j!}\left(\frac{t+s+u+v}{\sqrt{2}}\right)^j f(t,s,u,v)$$

and

$$R.H.S. = \sum_{k,l,m,n,j=0}^{+\infty} \frac{B^j_{klmn} H_j(\sqrt{2}x)}{k!l!m!n!} t^k s^l u^m v^n$$
$$= \sum_{j=0}^{+\infty} H_j(\sqrt{2}x)\left(\sum_{k,l,m,n=0}^{+\infty} \frac{B^j_{klmn}}{k!l!m!n!} t^k s^l u^m v^n\right).$$

Through comparing the coefficients, we get for each $j \in \mathbb{Z}_+$

$$\frac{1}{j!}\left(\frac{t+s+u+v}{\sqrt{2}}\right)^j f(t,s,u,v) = \sum_{k,l,m,n=0}^{+\infty} \frac{B^j_{klmn}}{k!l!m!n!} t^k s^l u^m v^n$$

In particular, when $j = 0$ we have

$$f(t,s,u,v) = \sum_{k,l,m,n,j=0}^{+\infty} \frac{B^0_{klmn} H_j(\sqrt{2}x)}{k!l!m!n!} t^k s^l u^m v^n.$$

Thus

$$C_{klmn} = \frac{1}{\sqrt{2\pi}}\sqrt{k!l!m!n!}\,2^{-(k+l+m+n)/2} C(t^k s^l u^m v^n;\ f(t,s,u,v)).$$

□

Corollary 2.4.14. *For general nontrivial term C_{klmn} ($k+l+m+n$ even), we have the estimate*

$$|C_{klmn}| \le \frac{1}{\sqrt{2\pi}}\left(\frac{(2p)!}{2^{2p}p!p!}\right)\sqrt{\frac{p!p!}{k!l!m!n!}}, \tag{2.81}$$

where $p=(k+l+m+n)/2$.

Proof. By the Taylor expansion series formula for exponential functions, we write

$$C(t^k s^l u^m v^n;\ f(t,s,u,v)) = C\left(t^k s^l u^m v^n;\ \frac{1}{p!}\left(\frac{(t+s+u+v)^2}{2} - (t^2+s^2+u^2+v^2)\right)^p\right).$$

We observe that the coefficient of term $t^k s^l u^m v^n$ in $\left(\frac{(t+s+u+v)^2}{2} - (t^2+s^2+u^2+v^2)\right)^p$ can always be controlled by that in $\left(\frac{(t+s+u+v)^2}{2}\right)^p$, therefore

$$\begin{aligned} |C(t^k s^l u^m v^n;\ f(t,s,u,v))| &\le C\left(t^k s^l u^m v^n;\ \left(\frac{(t+s+u+v)^2}{2}\right)^p\right) \\ &= \frac{1}{p!}\frac{1}{2^p}\frac{(k+l+m+n)!}{k!l!m!n!}. \end{aligned}$$

According to the result of the last proposition, we get

$$\begin{aligned} |C_{klmn}| &\le \frac{1}{\sqrt{2\pi}}\frac{1}{2^{2p}p!}\frac{(2p)!}{\sqrt{k!l!m!n!}} \\ &= \frac{1}{\sqrt{2\pi}}\frac{(2p)!}{2^{2p}p!p!}\sqrt{\frac{p!p!}{k!l!m!n!}}. \end{aligned}$$

□

Although the inequality in (2.81) only provides a poor estimate for term C_{klmn}, when k,l,m,n are almost equal to each other, it provides a much sharper estimate than that in last section when one of k,l,m,n is much larger than all of the others.

Corollary 2.4.15. *There is a positive integer N and a real number $a>1$ such that if $k > N(l+m+n)$ then*

$$|C_{klmn}| \lesssim a^{-k}. \tag{2.82}$$

Proof. By using the Sterling's formula, the estimate (2.81) can be continued as

$$|C_{klmn}| \lesssim \frac{1}{2\pi} \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt[4]{klmn}} \sqrt{\frac{p^p p^p}{k^k l^l m^m n^n}}$$
$$\lesssim \sqrt{\frac{p^p p^p}{k^k l^l m^m n^n}}.$$

Let $I = \sqrt{p^{2p}/k^k l^l m^m n^n}$, then we have

$$\ln \frac{1}{I} = \frac{p}{2} \left(\frac{k}{p} \ln \frac{k}{p} + \frac{l}{p} \ln \frac{l}{p} + \frac{m}{p} \ln \frac{m}{p} + \frac{n}{p} \ln \frac{n}{p} \right),$$

which can be regarded as the product of the number $p/2$ and the sum of the values of the function $f(x) = x \ln x$ at the points of k/p, l/p, m/p and n/p. From the fact that $p = (k+l+m+n)/2$, we know $k/p + l/p + m/p + n/p = 2$, and when k is much bigger than l, m and n, then k/p is close to 2 and l/p, m/p and n/p will fall into the internal $[0, 1]$, on which the function $f(x)$ is nonpositive and convex. Thus we have

$$\frac{k}{p} \ln \frac{k}{p} + \frac{l}{p} \ln \frac{l}{p} + \frac{m}{p} \ln \frac{m}{p} + \frac{n}{p} \ln \frac{n}{p}$$
$$\geq \frac{k}{p} \ln \frac{k}{p} + \frac{l+m+n}{p} \ln \frac{l+m+n}{3p}$$
$$= \frac{k}{p} \ln \frac{k}{p} + \left(2 - \frac{k}{p}\right) \ln \left(\frac{2 - \frac{k}{p}}{3} \right).$$

This quantity tends to $2 \ln 2$ as k/p tends to 2. Thus there is a positive integer N sufficiently large and a positive number $\delta > 0$ such that if $k > N(l+m+n)$ then

$$\ln \frac{1}{I} \geq \frac{p}{2} \delta \geq \frac{k}{4} \delta > 0.$$

For example, we give a choice of the pair of (N, δ); we can choose $N = 7$ (which implies $k/p > 1.75$) and

$$\delta = 0.358 < 1.75 \ln 1.75 + 0.25 \ln (0.25/3).$$

Let $a = e^{\delta/4} > 1$, then the inequality (2.82) holds. $\square$

At last, before we end up this section, we provide a few values of nontrivial entries C_{klmn} $(k + l + m + n$ even)

$$\begin{array}{lll} C_{0000} = \frac{1}{\sqrt{2\pi}} & C_{0011} = \frac{1}{2\sqrt{2\pi}} & C_{0002} = -\frac{1}{4\sqrt{\pi}} \\ C_{0022} = \frac{3}{8\sqrt{2\pi}} & C_{0112} = \frac{1}{8\sqrt{\pi}} & C_{1111} = \frac{3}{4\sqrt{2\pi}} \\ C_{0222} = \frac{1}{32\sqrt{\pi}} & C_{1122} = \frac{7}{16\sqrt{2\pi}} & C_{2222} = \frac{41}{64\sqrt{2\pi}} \end{array} \tag{2.83}$$

Chapter 3

Function Spaces

3.1 Introduction

Let us consider such a linear Schrödinger equation of the following form (Schrödinger equation with quadratic potential), also known as the problem of a quantum harmonic oscillator

$$\begin{cases} i\psi_t = \frac{1}{2}\psi_{xx} - \frac{x^2}{2}\psi & x \in \mathbb{R}^n \\ \psi(x,0) = \psi_0(x) & \psi \text{ complex valued.} \end{cases} \tag{3.1}$$

This equation can be solved completely in several ways, one, using the theoretic structure of (semi)groups [H-P]. By writing the equation in the form of $-i\psi_t = A\psi = \frac{(-\Delta+x^2)}{2}\psi$, we express its unique solution as $\psi(x,t) = e^{iAt}\psi_0(x)$, where the operator e^{iAt} is a strongly continuous group (C_0 group) on the function space $L^2(\mathbb{R}^n)$. In particular, for any time $t \in \mathbb{R}$ the operator e^{iAt} is a bounded linear mapping on $L^2(\mathbb{R}^n)$, and it satisfies

(i) when $t = 0$, the mapping is the identity operator on $\mathbb{R}^n$;

(ii) $\forall t, s \geq 0$, $e^{iA(t+s)} = e^{iAt} \circ e^{iAs}$;

(iii) $\forall \varphi \in L^2(\mathbb{R}^n) : ||e^{iAt}\varphi - \varphi||_{L^2(\mathbb{R}^n)} \to 0$, as $t \to 0$.

A basic property of this group is that all the operators are unitary.

Proposition 3.1.1. *The C_0 group is made up of unitary operators on $L^2(\mathbb{R}^n)$, which means for any time $t \in \mathbb{R}$ the operator e^{iAt} is an isomorphism on the function space $L^2(\mathbb{R}^n)$, and it preserves the inner product $(\phi, \varphi) = \int_{-\infty}^{+\infty} \phi(x)\overline{\varphi(x)}dx$ on $L^2(\mathbb{R}^n)$, that is, $\left(e^{iAt}\phi, e^{iAt}\varphi\right) = (\phi, \varphi)$.*

Meanwhile this operator e^{iAt} has an explicit form of its kernel, given by Mehlor's formula[F-H][C1].

Proposition 3.1.2. *Mehlor's formula is an explicit expression for the kernel of the operator e^{iAt} for $0 < t < \pi$:*

$$\psi(t, x) = \frac{1}{(-2i\pi \sin t)^{n/2}} \int_{\mathbb{R}^n} e^{\frac{-i}{\sin t}\left(\frac{x^2+y^2}{2} \cos t - x \cdot y\right)} \psi_0(y) dy, \tag{3.2}$$

where ψ_0 is any function in the space $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

From the proposition above. we can see that the operator in fact can be defined on the function space $L^1(\mathbb{R}^n)$. An easy application of the Mehlor's formula is that this operator shows a dispersive property for a period of time.

Proposition 3.1.3. *For $0 < t \leq \frac{\pi}{2}$, this C_0 group $\{e^{iAt}\}$ satisfies*

$$\left\|e^{iAt}\right\|_{L^1 \to L^\infty} \leq c\,|t|^{-n/2}. \tag{3.3}$$

As we have introduced before, Hermite functions are eigenfunctions of the operator $A = (-\Delta + x^2)/2$. By using this fact, we can write the solution of (3.1) in another form. For simplicity, we always work on the 1-dimensional case below without further specification. *But all the results in the following sections are in fact true for* n-dimensional cases.

Since $\{h_k(x), k = 0, 1, 2 \cdots\}$ is an orthonormal basis of the function space $L^2(\mathbb{R})$, we can uniquely write the initial data $\psi_0(x) \in L^2(\mathbb{R})$ as a series $\psi_0(x) = \sum_{k\geq 0}^{+\infty} q_k(0)h_k(x)$, where it converges with respect to the L^2-norm. Then we can write the solution as

$$\psi(t,x) = e^{iAt}\psi_0(x) = \sum_{k\geq 0}^{+\infty} q_k(0)e^{i\omega_k t}h_k(x), \tag{3.4}$$

where $\omega_k = k + 1/2, \; k = 0, 1, 2, \cdots$ is the k-th eigenvalues of operator A. In particular, we have the proposition.

Proposition 3.1.4. *For the C_0 group $\{e^{iAt}\}$ acting on $L^2(\mathbb{R})$, it has such periodic properties*

$$\exp(iA(t + 2\pi)) = -\exp(iAt), \tag{3.5}$$

$$\exp(iA(t + 4\pi)) = \exp(iAt). \tag{3.6}$$

A natural idea is that this result may be generalized for a wider family of functions than L^2. Formally we can have such a result: if the initial data

$$\psi_0(x) = \sum_{k\geq 0}^{+\infty} c_k h_k(x) \tag{3.7}$$

for sequences $\{c_k\}$ with some properties, then the solution $\sum_{k\geq 0}^{=\infty} c_k e^{i\omega_k t}h_k(x)$ should be the solution of the partial differential equation (3.1). In the following section, we will introduce a family of Hilbert spaces $X^s(\mathbb{R}^n)$ (a Hilbert scale) based on this representation of the initial data in decomposition with respect to Hermite functions.

It is an unexpected thing that this Hilbert scale turns out to have a close relationship with the theory of rapidly decreasing functions and tempered distributions. In fact, the representation (3.7) is true for every tempered distribution and the equality holds in the sense of Hermite expansion series for tempered distribution. Thus the PDE can have a uniquely global solution for any initial data which is a tempered distribution. When we turn to study the Hilbert scale itself, it is found that these Hilbert functions have many very good properties. For example, they can determine the topological structure of the Fréchet space $S(\mathbb{R}^n)$ and that of the tempered distribution $S'(\mathbb{R}^n)$; many familiar operators like the annihilation operator (lowering operator) and creation operator (raising operator) can be regarded as homomorphism on the Hilbert scale; in particular the Fourier transformation can be regarded as essentially a member in the unitary (semi)group of $e^{\imath At}$.

One particularly interesting thing is that the function spaces $X^s(\mathbb{R}^n)$ have close relationships with Sobolev spaces $H^s(\mathbb{R}^n)$. When $s \in \mathbb{Z}_+$, they are just the space $H^s(\mathbb{R}^n) \cap \mathcal{F}(H^s(\mathbb{R}^n))$; when $s \in \mathbb{Z}_-$, then they are the space $H^s(\mathbb{R}^n) + \mathcal{F}(H^s(\mathbb{R}^n))$. If s is not an integer, similar results seem to be correct too. In that case, we provide some partial results in the direction of trying to prove it. Some other properties are also considered, such as the relationship of the spaces $X^s(\mathbb{R}^n)$ and $L^p(\mathbb{R}^n)$ and the definition of the product of two functions in $X^s(\mathbb{R}^n)$. In the process of tracing back to the theory of rapid decreasing functions and tempered distribution, we realize that the function spaces $X^s(\mathbb{R}^n)$ are "essentially" the V. Bargmann's function spaces F_n^ρ, which were used to analyze the properties of the tempered distributions in [Bar1][Bar2].

The following sections are organized in this way. In section 3.2, we provide definitions of the function spaces $X^s(\mathbb{R}^n)$ and prove that they form a Hilbert scale, whose intersection is the rapid decreasing functions space $S(\mathbb{R}^n)$ and whose union is the tempered distribution space $S'(\mathbb{R}^n)$. In section 3.3, we study on the relationship of topological structure of the function space $X^s(\mathbb{R}^n)$ and those of the space $S(\mathbb{R}^n)$ and $S'(\mathbb{R}^n)$. Besides that we consider on many basic properties of the function spaces $X^s(\mathbb{R}^n)$ from the viewpoint of Hilbert scale, and many familiar operators are also considered when regarded as a homomorphism of the Hilbert scale. In section 3.4, we study on the relationship of the function spaces $X^s(\mathbb{R}^n)$ and the Sobolev spaces $H^s(\mathbb{R}^n)$. In the last section (section 3.5), we study on several problems: the first one is the relationship between the space $X^s(\mathbb{R}^n)$ and the space $L^p(\mathbb{R}^n)$; the second one is about the definition of the product of the two functions from the space $X^s(\mathbb{R}^n)$; the last one is about the proof that for any $s \in \mathbb{R}$ the function spaces $X^s(\mathbb{R}^n)$ are isomorphic to the V. Bargmann's function spaces F_n^ρ as two Hilbert spaces. A brief review on V. Bargmann's work [Bar1][Bar2] are provided in this section. And then the difference of our work from that of V. Bargmann, together with B. Simon's work [Si], are pointed out .

3.2 Definitions of function spaces

In the last section, some review on the properties of the operator $e^{\imath At}$ are provided. Now another observation on the operator $A = (-\Delta + x^2)/2$ is that it is an positive definite essentially self adjoint unbounded operator on the function space $L^2(\mathbb{R})$, and the inverse of its square root operator $B = A^{-1/2} = ((-\Delta + x^2)/2)^{-1/2}$ is a compact operator on the same space. Let us recall that $L^2(\mathbb{R})$ is a separable complex Hilbert space with inner product $(u,v) = \int_{-\infty}^{+\infty} u(x)\ \overline{v(x)}dx$, and it has an orthonormal basis $\{h_k(x),\ k = 0,1,2,\cdots\}$ Then the operator B can be represented in the form $Bu = \sum_{k=0}^{+\infty}(k+\frac{1}{2})^{-1/2}\ (u, h_k(x))h_k(x)$. Below let us denote the space $L^2(\mathbb{R})$ by $E^0(\mathbb{R})$, or just E^0 when the space is already taken as $\mathbb{R}$. Now we can define a scale of Hilbert space[Mi, p.93], [Bo], [E-K-M-Y, p.61], [B-H-K, p.143].

Definition 3.2.1. *Let* $E_{+\infty} = \bigcap_{n=0}^{+\infty} B^n E_0$ *and for* $u, v \in E_{+\infty}$ *and* $\forall\alpha \in \mathbb{R}$, *define* $(u,v)_\alpha = (B^{-\alpha}u, B^{-\alpha}v)$. *Denote by* E_α *as the function space which is the completion of the space* $E_{+\infty}$ *equipped with the* $\|\|_{E_\alpha} = (\,,)_\alpha^{1/2}$ *topology. The family of* $\alpha \to E_\alpha$ *is called the* ***Hilbert scale*** *defined by* B, *and* $E_{+\infty}$ *is called the* ***center***.

It follows that $E_{+\infty} = \bigcap_{\alpha\in\mathbb{R}} E_\alpha$, and if we suppose that $E_{+\infty}$ carries the weakest topology in which all inclusions $E_{+\infty} \to E_\alpha$ are continuous, then we can get that $\bigcup_{\alpha\in\mathbb{R}} E_\alpha = E_{-\infty}$ is its dual space. Besides that, since $\|u\|_{E_\alpha} = \left\|B^\beta u\right\|_{E_{\alpha+\beta}}$ for all $u \in E_{+\infty}$ and $\alpha, \beta \in \mathbb{R}$ it follows that the operator B^β extends to an isometry from E_α to $E_{\alpha+\beta}$, which we will also denote by B^β.

In the following parts, we will use another natural way to define our function space and then go back to check that our definition fits in the above theoretic structure.

As we have mentioned, the operator $A = (-\Delta + x^2)/2$ is a positive definite self-adjoint unbounded operator on the function space $L^2(\mathbb{R})$. In terms of the orthogonal basis $\{h_k(x),\ k = 0,1,2,\cdots\}$, the operator A can be interpreted as an infinite matrix $A = (a_{jk})$ where $a_{jk} = \delta_{jk}\ (k + 1/2)$ and $j, k = 0,1,2,\cdots$. Then we can define an linear operator A^s for any $s \geq 0$ through such an infinite matrix,

$A^s = (b_{jk})_{j,k=0}^{+\infty} = \left(a_{jk}^s\right)_{j,k=0}^{+\infty}$. In other words,

$$A^s(u) = A^s\left(\sum_{k=0}^{+\infty} q_k h_k(x)\right) = \sum_{k=0}^{+\infty} q_k (k+1/2)^s h_k(x), \tag{3.8}$$

where its definition domain can be chose as $D(A^s) = \{u \in L^2(\mathbb{R}) |\ u = \sum_{k=0}^{+\infty} q_k h_k(x)$ satisfying $\sum_{k=0}^{+\infty} |q_k|^2 (k+1/2)^{2s} < +\infty\ \}$. It is easy to verify such two facts for this operator: (i) its adjoint operator $(A^s)^*$ satisfies $D((A^s)^*) = D(A^s)$ and $(A^s)^* = A^s$; (ii) $(A^s(u), u) = \sum_{k=0}^{+\infty} |q_k|^2 (k+1/2)^s \geq 0$. That means the linear operator defined by (3.8) is a positive unbounded self-adjoint operator. Below we define our function space like this.

Definition 3.2.2. *For any $s \geq 0$, if any function $u \in L^2(\mathbb{R})$ is in the definition domain of the operator A^s, then we say that u is in the* ***function space*** $\mathbf{X}^{2s}(\mathbb{R})$. *And we endow this function space with such an inner product, $(u, v)_{X^{2s}} = (A^s u, A^s v)_{L^2(\mathbb{R})}$ for all $u, v \in X^{2s}(\mathbb{R})$.*

In general, if T is an unbounded self-adjoint operator on a Hilbert space H with its inner product $(\cdot,\cdot)$, then its definition domain $D(T)$ can also become a Hilbert space when we equip this linear subspace with the inner product $(u, v)_{D(T)} = (Tu, v) + (u, v)$. In the definition of our function space, there is no second term for the inner product. But it turns out that our defined space is still a Hilbert space. This is due to the fact that our operator A^s is in some sense "strict positive" and the second term (u, v) can thus be absorbed into the first term $(A^s u, A^s v)$.

Theorem 3.2.3. *For $s \geq 0$ the function space $X^{2s}(\mathbb{R})$ forms a Hilbert space under the inner product $(\cdot,\cdot)_{X^{2s}}$. And this space is isomorphic to the space $l_2^s(\mathbb{Z}_+; \mathbb{C}) := \{q = (q_0, q_1, q_2, \cdots) \in \mathbb{Z}_+^{\mathbb{C}}$ satisfying $\sum_{k=0}^{+\infty} |q_k|^2 \langle k \rangle^s < +\infty\}$ under the mapping T :*

$X^{2s}(\mathbb{R}) \longrightarrow l_2^s(\mathbb{Z}_+;\mathbb{C})$ *by* $T\left(\sum_{k=0}^{+\infty} q_k h_k(x)\right) = (q_0, q_1, q_2, \cdots)$.

Proof. By the property of the operator A^s, we can have that for any $u = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^{2s}(\mathbb{R})$, $(A^s u, A^s u) = \sum_{k=0}^{+\infty} (k+1/2)^{2s} \; |q_k|^2 \approx \sum_{k=0}^{+\infty} \langle k \rangle^{2s} \; |q_k|^2$, where the last term is the standard inner product on the space $l_2^s(\mathbb{Z}_+;\mathbb{C})$ (without further clarification we always writing $l_2^s(\mathbb{Z}_+)$ instead of the notation $l_2^s(\mathbb{Z}_+;\mathbb{C})$ in the following part of this thesis for simplicity). According to this, it is easy to verify that the mapping T is a well defined continuous linear operator and it is injective.

Conversely, since s is nonnegative, for any $q = (q_0, q_1, q_2, \cdots) \in l_2^s(\mathbb{Z}_+)$ we can verify that the corresponding function $u(x) = \sum_{k=0}^{+\infty} q_k h_k(x)$ are really in the function space $X^{2s}(\mathbb{R})$, which means the mapping T is also surjective. We can easily further verify that the inverse of mapping T is also continuous. Thus the function space $X^{2s}(\mathbb{R})$ is isomorphic to $l_2^s(\mathbb{Z}_+)$, which obviously is a Hilbert space. □

According to the theorem above, we know that every function in $X^{2s}(\mathbb{R})$ $(s \geq 0)$ can be represented as $\sum_{k=0}^{+\infty} q_k h_k(x)$, where $q = (q_0, q_1, q_2, \cdots)$ is the unique correspondence in the function space $l_2^s(\mathbb{Z}_+)$. So a natural question next is what kind of functions are really in our function space $X^{2s}(\mathbb{R})$ and whether they have close relationship with other function spaces which are already familiar to us. The answer turns out to be yes.

Theorem 3.2.4. *If $2s$ is an integer, say $2s = n \in \mathbb{Z}_+$, then our function space $X^n(\mathbb{R})$ in fact is*

$$X^n(\mathbb{R}) = \left\{ u \in L^2(\mathbb{R}) \mid x^\alpha \left(\frac{d}{dx}\right)^\beta u(x) \in L^2(\mathbb{R}) \text{ for all indices } 0 \leq \alpha + \beta \leq n \right\},$$

and there is an equivalent norm $\|u\|_n^2 = \sum_{0\le\alpha+\beta\le n} \left\|x^\alpha u^{(\beta)}\right\|_{L^2}^2 \approx \|u\|_{X^n}^2 = (A^s u, A^s u)$ *on this function space.*

Remark 3.2.5. *In particular, we have such two special cases: one is the case* $n = 0$, *we have* $X^0(\mathbb{R}) = L^2(\mathbb{R})$*; the other one is the case* $n = 1$*, then we have* $X^1(\mathbb{R}) = \{u \in L^2(\mathbb{R}) \mid \partial_x u \in L^2(\mathbb{R}), xu \in L^2(\mathbb{R})\} = \{u \in H^1(\mathbb{R}) \mid xu \in L^2(\mathbb{R})\}$*, which is the virial space in [M-R2].*

Proof. Let us define $V^n(\mathbb{R}) = \left\{u \in L^2(\mathbb{R}) \mid x^\alpha \partial_x^\beta u(x) \in L^2 \text{ for all indices } 0 \le \alpha + \beta \le n\right\}$ as a function space equipped with the inner product $(u,v)_{V^n} = \sum_{0\le\alpha+\beta\le n} \left(x^\alpha \partial_x^\beta u, x^\alpha \partial_x^\beta u\right)_{L^2}$. It is easy to check that this function space V^n with the given inner product is a Hilbert space. So the only thing left to show is that these two function spaces are in fact the same one. Let us prove this by the method of induction.

When $n = 0$, then obviously we have the operator $A^0 = Id$, thus function space $X^0(\mathbb{R}) = L^2(\mathbb{R}) = V^1(\mathbb{R})$.

In the case of $n = 1$, for those good functions u in the function space X^1, say $u \in S$, we have

$$\begin{aligned}(u,u)_{X^1} &= \left(\frac{(-\Delta + x^2)}{2} u, u\right) \\ &= \frac{1}{2}\left(\int_{-\infty}^{+\infty} \left|\frac{d}{dx} u\right|^2 dx + \int_{-\infty}^{+\infty} |xu|^2\, dx\right) \\ &= (u,u)_{V^1}.\end{aligned}$$

We claim that the function space S is both a dense subset of $X^1(\mathbb{R})$ and $V^1(\mathbb{R})$ with respect to their norms respectively. This is because:

(i) In the case of space $X^1(\mathbb{R})$, for any $u = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^1(\mathbb{R})$, we have

$u_N = \sum_{k=0}^{N} q_k h_k(x) \in S$, and $\|u - u_N\|_{X^1}^2 = \sum_{k=N+1}^{+\infty} |q_k|^2 \, (k + 1/2) \longrightarrow 0$ as $N \longrightarrow +\infty$;

(ii) In the other case, we can first find two C_0^∞ function (cut-off functions): one is $0 \leq \chi(x) \leq 1$ satisfying $\chi(x) = 0$ if $|x| \geq 2$ and$\chi(x) = 1$ if $|x| \leq 1$; the other one is $0 \leq J(x) \leq 1$ satisfying $J(x) = 0$ if $|x| \geq 1$ and $\int_{-\infty}^{+\infty} J(x)dx = 1$. Let $\chi_N(x) = \chi(x/N)$ and $J_{1/M}(x) = MJ(xM)$. Then for any $v \in V^1(\mathbb{R})$, we have a family of rapid decreasing functions (in fact C_0^∞ functions) $J_{1/M}(x) * (\chi_N(x)v)$. It is easy to verify that for any given $f, g \in V^1(\mathbb{R})$, we have $\|\chi_N(x)f - f\|_{V^1} \longrightarrow 0$ as $N \longrightarrow +\infty$ and $\left\|J_{1/M}(x) * g - g\right\|_{V^1} \longrightarrow 0$ as $M \longrightarrow +\infty$. Then we can choose a sequence of functions $J_{1/M(N)}(x) * (\chi_N(x)v) \longrightarrow v$ in the function space $V^1(\mathbb{R})$.

The claim implies that $X^1(\mathbb{R})$ and $V^1(\mathbb{R})$ are respectively the complement of the function space S under the norm $\|\cdot\|_{X^1}$ and $\|\cdot\|_{V^1}$, which are in fact equivalent to each other on S. By the uniqueness of the complement of a normed space, we can draw the conclusion that the function spaces $X^1(\mathbb{R})$ and $V^1(\mathbb{R})$ are in fact identical.

Suppose our conclusion is true for all the cases $0 \leq n \leq k$, let us prove it is also true for case $n = k + 1$. Again, it is still true that S is a dense subset of both spaces $X^{k+1}(\mathbb{R})$ and $V^{k+1}(\mathbb{R})$ under the norm $\|\cdot\|_{X^{k+1}}$ and $\|\cdot\|_{V^{k+1}}$ respectively. In fact we can use the exactly same method as in the case X^1 and V^1 to get the desired approximation sequence. We claim that $\|\cdot\|_{X^{k+1}}$and $\|\cdot\|_{V^{k+1}}$ are equivalent norms on S. By the uniqueness nature of the completion of a normed space, we can deduce that the two normed spaces, S with restricted norm from $\|\cdot\|_{X^{k+1}}$and $\|\cdot\|_{V^{k+1}}$, are in fact identical. Below let us prove our claim.

For $u \in S$, we have

$$
\begin{aligned}
2\|u\|_{X^{k+1}}^2 &= 2\left(A^{k+1}u, u\right) \\
&= \left(A^k(-\frac{d}{dx})\frac{d}{dx}u, u\right) + \left(A^k x \cdot xu, u\right) \\
&= \left(A^{k-1}(-\frac{d}{dx})A\frac{d}{dx}u, u\right) + \left(A^{k-1}x\frac{d}{dx}u, u\right) \\
&+ \left(A^{k-1}xAxu, u\right) + \left(A^{k-1}(-\frac{d}{dx})xu, u\right) \\
&= \left(A^{k-2}(-\frac{d}{dx})A^2\frac{d}{dx}u, u\right) + 2\left(A^{k-2}xA\frac{d}{dx}u, u\right) + \left(A^{k-2}(-\frac{d}{dx})\frac{d}{dx}u, u\right) \\
&+ \left(A^{k-2}xA^2xu, u\right) + 2\left(A^{k-2}(-\frac{d}{dx})Axu, u\right) + \left(A^{k-2}x \cdot xu, u\right) \\
&= \cdots \\
&= \sum_{d\ even}^{k}\left(\left((-\frac{d}{dx})A^{k-d}\frac{d}{dx}u, u\right)\binom{k}{d} + \left(xA^{k-d}xu, u\right)\binom{k}{d}\right) \\
&+ \sum_{d\ odd}^{k}\left(\left(xA^{k-d}\frac{d}{dx}u, u\right)\binom{k}{d} + \left((-\frac{d}{dx})A^{k-d}xu, u\right)\binom{k}{d}\right) \\
&\leq \sum_{d\ even}^{k}\binom{k}{d}\left(\left\|\frac{d}{dx}u\right\|_{X^{k-d}}^2 + \|xu\|_{X^{k-d}}^2\right) \\
&+ \sum_{d\ odd}^{k} 2\binom{k}{d}\left(A^{(k-d)/2}\frac{d}{dx}u, A^{(k-d)/2}\frac{d}{dx}u\right)^{1/2}\left(A^{(k-d)/2}xu, A^{(k-d)/2}xu\right)^{1/2} \\
&\leq C\sum_{d\ even}^{k}\binom{k}{d}\left(\left\|\frac{d}{dx}u\right\|_{V^{k-d}}^2 + \|xu\|_{V^{k-d}}^2\right) \\
&+ \sum_{d\ odd}^{k} 2\binom{k}{d}\left(\left(A^{(k-d)/2}\frac{d}{dx}u, A^{(k-d)/2}\frac{d}{dx}u\right) + \left(A^{(k-d)/2}xu, A^{(k-d)/2}xu\right)\right) \\
&\leq C\sum_{d=0}^{k}\binom{k}{d}\left(\left\|\frac{d}{dx}u\right\|_{V^{k-d}}^2 + \|xu\|_{V^{k-d}}^2\right) \\
&\leq C\|u\|_{V^{k+1}}^2 .
\end{aligned}
$$

In the process above, we have used the fact that $A(-\frac{d}{dx}) = (-\frac{d}{dx})A + x$ and $A \cdot x = x \cdot A + (-\frac{d}{dx})$.

Similarly, we have

$$
\begin{aligned}
\|u\|^2_{V^{k+1}} &\leq \left\|\frac{d}{dx}u\right\|^2_{V^k} + \|xu\|^2_{V^k} \\
&\lesssim \left(\left(A^k\frac{d}{dx}u, \frac{d}{dx}u\right) + \left(A^k xu, xu\right)\right) \\
&= \left(\left(-\frac{d}{dx}\right)A^k\frac{d}{dx}u, u\right) + \left(xA^k xu, u\right) \\
&= \sum_{d\ even}^{k}\left(\left(\frac{d}{dx}\left(-\frac{d}{dx}\right)A^{k-d}u, u\right) + \left(x^2A^{k-d}u, u\right)\right)\binom{k}{d} \\
&+ \sum_{d\ odd}^{k}\left(\left(\left(\frac{d}{dx}\right)xA^{k-d}u, u\right) + \left(x\left(-\frac{d}{dx}\right)A^{k-d}u, u\right)\right)\binom{k}{d} \\
&= \sum_{d\ even}^{k}2\left(\left(A^{k+1-d}u, u\right) + \left(x^2A^{k-d}u, u\right)\right)\binom{k}{d} + \sum_{d\ odd}^{k}\left(A^{k-d}u, u\right)\binom{k}{d} \\
&\leq 2\sum_{d=0}^{k}\left(A^{k+1}u, u\right)\binom{k}{d} \\
&\lesssim \|u\|^2_{X^{k+1}}.
\end{aligned}
$$

Thus we have finished the proof. $\square$

Let us consider the intersection of the function space $X^n(\mathbb{R})$, then the last theorem implies that

$$
\begin{aligned}
X^{+\infty}(\mathbb{R}) &= \bigcap_{n=1}^{+\infty}X^n(\mathbb{R}) \\
&= \{u \in L^2(\mathbb{R}) \mid x^\alpha\left(\frac{d}{dx}\right)^\beta u \in L^2(\mathbb{R}) \text{ for all indices } \alpha \text{ and } \beta\ \} \\
&= \{u \in H^{+\infty}(\mathbb{R}) \mid x^\alpha\left(\frac{d}{dx}\right)^\beta u \in L^2(\mathbb{R}) \text{ for all indices } \alpha \text{ and } \beta\ \} \\
&= \{u \in C^{+\infty}(\mathbb{R}) \mid x^\alpha\left(\frac{d}{dx}\right)^\beta u \in L^\infty(\mathbb{R}) \text{ for all indices } \alpha \text{ and } \beta\ \} \\
&= S(\mathbb{R}).
\end{aligned}
$$

It is clear now that our spaces have close relationships with the Schwartz function spaces. Furthermore, it is a natural question to ask what is the relationship of the topological structure of $X^n(\mathbb{R})$ and that of the function space $S(\mathbb{R})$.

As we known, the standard topology on $S(\mathbb{R})$ is given by a family of seminorms $\|u\|_{\alpha,\beta,\infty} = \sup\left|x^\alpha u^{(\beta)}(x)\right|$, which make the space $S(\mathbb{R})$ become a Fréchet space. Since the space $S(\mathbb{R})$ turns out to be intersections of all the Hilbert spaces $X^n(\mathbb{R})$, we can have another choice to provide it a topological structure. Let $\|\cdot\|_n$ be the restriction of $\|\cdot\|_{X^n}$ on the function space $S(\mathbb{R})$, then they form a sequence of norms on that space, which induces a topology on $S(\mathbb{R})$ like this: every open set is the union of the sets in the form of $\{f \in S(\mathbb{R}) \mid \|f\|_{n_1} < \delta_1, \cdots, \|f\|_{n_k} < \delta_k\}$.

In fact, this topology on $S(\mathbb{R})$ is the weakest topology satisfying that all the inclusion mappings $S(\mathbb{R}) \hookrightarrow X^n(\mathbb{R})$ are continuous. It turns out that these two topological structure are in fact the same one. Below let us prove it by starting with the definition on what are two equivalent sequences of seminorms on a topological space.

Definition 3.2.6. *We call two families of seminorms $\{\rho_\alpha\}_{\alpha\in A}$ and $\{d_\beta\}_{\beta\in B}$ on a vector space X equivalent if they generate the same natural topology.*

It is often useful to know such a proposition.

Proposition 3.2.7. *[9] Let $\{\rho_\alpha\}_{\alpha\in A}$ and $\{d_\beta\}_{\beta\in B}$ be two families of seminorms. The following statement are equivalent:*

(a) The families are equivalent families of seminorms.

(b) Each ρ_α is continuous in the d-natural topology and each d_β is continuous in the ρ-natural topology.

(c) For each $\alpha \in A$, there are $\beta_1, \cdots, \beta_n \in B$ and $C > 0$ so that for all $x \in X$

$$\rho_\alpha(x) \leq C\left(d_{\beta_1}(x) + \cdots + d_{\beta_n}(x)\right),$$

and for each $\beta \in B$, there are $\alpha_1, \cdots . \alpha_m \in A$ and $D > 0$ so that for all $x \in X$

$$d_\beta(x) \leq D\left(\rho_{\alpha_1}(x) + \cdots + \rho_{\alpha_m}(x)\right).$$

Obviously, it is an equivalent relation to say if two families of seminorms are equivalent. So if we want to show two families of seminorms are equivalent, we can choose another family of seminorms as a bridge, which is equivalent to both of the two target families. That is the way we proceed.

Lemma 3.2.8. *The families of seminorms $\{\|u\|_{\alpha,\beta,\infty}\}$ and $\{\|u\|_{\alpha,\beta,2}\}$ on $S(\mathbb{R})$ are equivalent.*

Proof. For any given function $f \in S(\mathbb{R})$, we have for any $x \in \mathbb{R}$

$$\begin{aligned}
\left|x^\alpha f^\beta(x)\right| &= \left|\int_{-\infty}^x \alpha x^{\alpha-1} f^{(\beta)}(x)dx + \int_{-\infty}^x x^\alpha f^{(\beta+1)}(x)dx\right| \\
&\leq \left|\alpha \int_{-\infty}^x \frac{(x^2+1)}{(x^2+1)} x^{\alpha-1} f^{(\beta)}(x)dx\right| + \left|\int_{-\infty}^x \frac{(x^2+1)}{(x^2+1)} x^\alpha f^{(\beta+1)}(x)dx\right| \\
&\leq \alpha\left(\left\|x^{\alpha+1} f^{(\beta)}(x)\right\|_{L^2} + \left\|x^{\alpha-1} f^{(\beta)}(x)\right\|_{L^2}\right) \int_{-\infty}^{+\infty} (x^2+1)^{-1}dx \\
&+ \left(\left\|x^\alpha f^{(\beta+1)}(x)\right\|_{L^2} + \left\|x^{\alpha+2} f^{(\beta)}(x)\right\|_{L^2}\right) \int_{-\infty}^{+\infty} (x^2+1)^{-1}dx \\
&\leq C\left(\|f\|_{\alpha+1,\beta,2} + \|f\|_{\alpha-1,\beta,2} + \|f\|_{\alpha,\beta+1,2} + \|f\|_{\alpha+2,\beta+1,2}\right).
\end{aligned}$$

That means $\|f\|_{\alpha,\beta,\infty} \leq C\left(\|f\|_{\alpha+1,\beta,2} + \|f\|_{\alpha-1,\beta,2} + \|f\|_{\alpha,\beta+1,2} + \|f\|_{\alpha+2,\beta+1,2}\right)$ is true for any index α and β.

And for the seminorm $\|u\|_{\alpha,\beta,2}$ we have

$$\left(\int_{-\infty}^{+\infty} \left|x^\alpha f^{(\beta)}(x)\right|^2 dx\right)^{1/2}$$
$$= \left(\int_{-\infty}^{+\infty} x^{2\alpha} f^{(\beta)}(x)\overline{f^{(\beta)}(x)}\frac{(x^2+1)}{(x^2+1)}dx\right)^{1/2}$$
$$\leq \left(\int_{-\infty}^{+\infty}(x^2+1)^{-1}dx\right)^{1/2}\left(\left\|x^{\alpha+1}f^{(\beta)}(x)\right\|_{L^\infty} + \left\|x^\alpha f^{(\beta)}(x)\right\|_{L^\infty}\right)$$
$$\leq C\left(\|f\|_{\alpha+1,\beta,\infty} + \|f\|_{\alpha,\beta,\infty}\right).$$

From these two inequalities, we can conclude that the two families of seminorms $\{\|u\|_{\alpha,\beta,\infty}\}$ and $\{\|u\|_{\alpha,\beta,2}\}$ will induce same topology on the function space $S(\mathbb{R})$. □

Note that the result in theorem 3.2.4 states that $\|u\|_n^2 \approx \sum_{0\leq\alpha+\beta\leq n} \|u\|_{\alpha,\beta,2}^2$, which implies that the family of norm $\{\|u\|_n\}$ induces the same topology on $S(\mathbb{R})$ as that induced by the norms $\{\|u\|_{\alpha,\beta,2}\}$. So we can state our results as follows.

Theorem 3.2.9. *We have*

$$X^{+\infty}(\mathbb{R}) = \bigcap_{n=0}^{+\infty} X^n(\mathbb{R}) = S(\mathbb{R}).$$

And if we endow the Schwartz function space with the weakest topology such that all the inclusion mappings $S(\mathbb{R}) \hookrightarrow X^n(\mathbb{R})$ are continuous, then this topology coincides with the standard topology which is induced by the family of seminorms $\{\|u\|_{\alpha,\beta,\infty}\}$.

Theorem (3.2.3) states that every function space $X^n(\mathbb{R})$ are "essentially" the Hilbert space $l_2^{n/2}(\mathbb{Z}_+)$. So the theorem above can have such a corollary.

Corollary 3.2.10. *Suppose that $\varphi \in L^2(\mathbb{R})$ and denote*

$$q_n = \int_{-\infty}^{+\infty}\varphi(x)h_n(x)dx \in \mathbb{C}.$$

Then $\varphi(x) \in S(\mathbb{R})$ *if and only if* $q = (q_0, q_1, \cdots) \in l_2^{+\infty}(\mathbb{Z}_+) = \bigcap_{n=0}^{+\infty} l_2^{n/2}(\mathbb{Z}_+)$. *And in that case* $\sum_{n=0}^{N} q_n h_n(x)$ *converges in the topology of* $S(\mathbb{R})$ *to* $\varphi(x)$ *as* $N \to +\infty$, *which can be denoted as* $\varphi(x) = \sum\limits_{n=0}^{N} q_n h_n(x)$.

This result in above corollary was once obtained by B. Simon in 1971 [Si] in a slightly different form. In his paper, another sequence of norms $\|\varphi\|_m = \left(|q_n|^2 (n+1)^m\right)^{1/2}$ were directly defined on the Schwartz functions space $S(\mathbb{R})$, and the author proved that every function $\varphi \in S(\mathbb{R})$ corresponds to a point in the sequence space satisfying $\sum\limits_{n=0}^{+\infty} |q_n|^2 (n+1)^m < +\infty$ is true for any $m \in \mathbb{Z}_+$. In his method all the viewpoints come from Schwartz function space itself, and the function space is not complete when it is equipped with only one norm $\|\cdot\|_m$.

However, our proof of the same result essentially comes from taking the Schwartz function space as the center of an Hilbert scale. In this way, it is easy to understand that the space $S(\mathbb{R})$ has many good properties such as separability, completeness, reflexivity etc. Since $S(\mathbb{R})$ is isomorphic to $l_2^{+\infty}(\mathbb{Z}_+)$, it is easy to conjecture that $S'(\mathbb{R}) \cong l_2^{-\infty}(\mathbb{Z}_+)$ is true. This space $S'(\mathbb{R})$, which is much larger than the space $L^2(\mathbb{R})$, provides us enough region to construct our Hilbert scale, especially the Hilbert space $X^s(\mathbb{R})$ for $s < 0$. In the following parts, we will first review Barry Simon's N-representation theory for S', then give out the definitions of our function spaces and finally prove that our function spaces form an Hilbert scale. Let us start with such a definition.

Definition 3.2.11. *A countable family of norms* $\|\cdot\|_k$ *is called directed if for any finite set* $k_1, \cdots, k_r$ *there is a* k *and a constant* C *so that* $\|f\|_{k_1} + \cdots + \|f\|_{k_r} \leq C \|f\|_k$.

Directed families are very useful because they provide a simple description of open sets and further of continuous functional.

Lemma 3.2.12. *A linear map $T : E \longrightarrow \mathbb{C}$ with E a countably normed space with a directed family of norms $\left\{\|\cdot\|_j\right\}$ is continuous if and only if $\exists C > 0, k$ such that*

$$|Tx| \leq C \|x\|_k .$$

It is very easy to find that our norms $\{\|\cdot\|_{X^k}\}$ with restriction on $S(\mathbb{R})$ are in fact a directed countable family of norms. So this lemma and the directed nature of $\|\cdot\|_{X^k}$ enable us to prove such a characterization of the space $S'(\mathbb{R})$. (Although very similar, it is in a little different form from the result in Barry Simon's work.)

Theorem 3.2.13. *Suppose f is a tempered distribution. Let $b_n = f(h_n(x))$, then there exists a real number s satisfying $\sum\limits_{n=0}^{+\infty} |b_n|^2 \langle n\rangle^{2s} < +\infty$, i.e. $b = (b_0, b_1, \cdots) \in l_2^s(\mathbb{Z}_+)$, and $f(\varphi) = \sum\limits_{n=0}^{+\infty} a_n b_n$ if a_n is the n-th Hermite coefficient of φ. Conversely, if $b = (b_0, b_1, \cdots) \in l_2^s(\mathbb{Z}_+)$ for some $s \in \mathbb{R}$, then the mapping $\varphi \longrightarrow \sum\limits_{n=0}^{+\infty} a_n b_n$ defines a tempered distribution.*

Proof. Since $f \in S'(\mathbb{R})$ and $\|\cdot\|_{X^m}$ is directed norms, by the lemma we get $|f(\varphi)| \leq C \|\varphi\|_{X^m}$ for some $m \in \mathbb{Z}_+$. Noticing $\|\varphi\|_{X^m} = \left(\sum\limits_{n=0}^{+\infty} |a_n|^2 (n+1/2)^m\right)^{1/2} \approx \|(a_0, a_1, \cdots)\|_{l_2^{m/2}(\mathbb{Z}_+)}$ and $S(\mathbb{R})$ is a dense subset of $X^m(\mathbb{R})$, we have a unique extension of f into a continuous linear functional on $X^m(\mathbb{R})$, or "essentially" on the space $l_2^{m/2}(\mathbb{Z}_+)$. As we know the dual space of $l_2^{m/2}(\mathbb{Z}_+)$ can be regarded as the Hilbert space $l_2^{-m/2}(\mathbb{Z}_+)$, so the distribution f has such a representation: let $b_n = f(h_n(x))$ then we have $\sum\limits_{n=0}^{+\infty} |b_n|^2 \langle n\rangle^{-m} < +\infty$, and $f(\varphi) = \sum\limits_{n=0}^{+\infty} a_n b_n$ where a_n is the n-th Hermite coefficient of $\varphi \in S(\mathbb{R})$.

Conversely, if $b = (b_0, b_1, \cdots)$ is an element in $l_2^s(\mathbb{Z}_+)$, there always exists an integer $m \geq 0$ such that $b \in l_2^{-m}(\mathbb{Z}_+)$. Then we only need make such a computation

$$\begin{aligned} \left|\sum_{n=0}^{+\infty} a_n b_n\right| &= \left|\sum_{n=0}^{+\infty} a_n \langle n\rangle^{m/2} b_n \langle n\rangle^{-m/2}\right| \\ &\leq \|b\|_{l_2^{-m}(\mathbb{Z}_+)} \|a\|_{l_2^m(\mathbb{Z}_+)} \\ &\lesssim \|\varphi\|_{X^m}, \end{aligned}$$

so that the mapping $\varphi \longrightarrow \sum_{n=0}^{+\infty} a_n b_n$ is a continuous linear functional on $S(\mathbb{R})$. □

This theorem can be viewed as in the dual form of the N-representation for the rapidly decreasing functions. It has such a corollary [Si].

Corollary 3.2.14. *$S(\mathbb{R})$ is dense in $S'(\mathbb{R})$ in the weak topology on S'. In particular, if $f \in S'(\mathbb{R})$ and $b_n = f(h_n(x))$ then $\sum_{n=0}^{N} b_n h_n(x) \longrightarrow f$ in the weak topology.*

Proof. The weak topology on $S'(\mathbb{R})$ is the weakest topology to make all the functionals $T_\varphi : f \in S'(\mathbb{R}) \longrightarrow f(\varphi)$ on $S(\mathbb{R})$ be continuous. So it is only left to check that we have $\sum_{n=0}^{N} a_n b_n \longrightarrow \sum_{n=0}^{+\infty} a_n b_n$, which is clearly true. □

Based on these two results about characterization of the tempered distribution, it is natural to come to the following definition [Si].

Definition 3.2.15. *For any function $f \in S'(\mathbb{R})$, we call $b_k = f(h_k(x))$ as the **k-th Hermite coefficient** of the tempered distribution f, and we write $f = \sum_{k=0}^{+\infty} b_k h_k(x)$ as its **Hermite expansion.***

What is more interesting to us is that we can generalize the definition for our function spaces into the cases with index $s < 0$.

Definition 3.2.16. *For any real number $s \in \mathbb{R}$, we call a function $f \in S'(\mathbb{R})$ is in the **function space** $\mathbf{X}^{2s}(\mathbb{R})$ if its Hermite coefficients $\{b_k = f(h_k(x))\}$ satisfies*

$$\sum_{k=0}^{+\infty} |b_k|^2 (k + 1/2)^{2s} < +\infty. \tag{3.9}$$

This definition is compatible with the definition once given for the cases $s \geq 0$, or in other words, $\{f \in S'(\mathbb{R}) \mid \sum_{k=0}^{+\infty} |b_k|^2 (k + 1/2)^{2s} < +\infty\} = \{f \in L^2(\mathbb{R}) | \sum_{k=0}^{+\infty} |b_k|^2 (k + 1/2)^{2s} < +\infty\}$. It is due to the fact: when $s \geq 0$, the inequality $\sum_{k=0}^{+\infty} |b_k|^2 (k + 1/2)^{2s} < +\infty$ implies that $b = (b_0, b_1, \cdots) \in l_2(\mathbb{Z}_+)$ and $\sum_{k=0}^{+\infty} b_k h_k(x)$ is an L^2-function.

Like the cases of $s \geq 0$, there is also a natural inner product on the function space $X^{2s}(\mathbb{R})$ for $s < 0$

$$(u, v)_{X^{2s}} = \sum_{k=0}^{+\infty} q_k \overline{p_k} (k + 1/2)^{2s},$$

where $\{q_k\}$and $\{p_k\}$ are Hermite coefficients of function u and v respectively. It is easy to get the following theorem.

Theorem 3.2.17. *For any $s \in \mathbb{R}$, the function space $X^{2s}(\mathbb{R})$ is a Hilbert space, which is isomorphic to the Hilbert space $l_2^s(\mathbb{Z}_+)$, i.e. $X^{2s}(\mathbb{R}) \cong l_2^s(\mathbb{Z}_+)$.*

To prove this theorem, we only need repeat the proof for the case $s \geq 0$. Again, it is the mapping $T : f(x) \longrightarrow (b_0, b_1, \cdots)$ that provides the isomorphism from $X^{2s}(\mathbb{R})$ to $l_2^s(\mathbb{Z}_+)$.

Up to now, a family of Hilbert spaces parameterized on real number s have been defined rigorously. We are in a good position to check if these Hilbert spaces are an Hilbert scale. The answer is yes.

Theorem 3.2.18. *The family of Hilbert spaces $X^{2s}(\mathbb{R})$ form an Hilbert scale, and they satisfy* $X^{+\infty}(\mathbb{R}) = \bigcap_{s\in\mathbb{R}} X^{2s}(\mathbb{R}) = S(\mathbb{R})$, *and* $X^{-\infty}(\mathbb{R}) = \bigcup_{s\in\mathbb{R}} X^{2s}(\mathbb{R}) = S'(\mathbb{R})$.

Proof. It is straightforward to prove this theorem, so below we just provide a sketch of the proof. According to the definition of the Hilbert scale, we need start with a compact operator. Let $B = A^{-1/2} = ((-\Delta + x^2)/2)^{-1/2}$, an operator acting on a separable complex Hilbert space $E_0 = L^2(\mathbb{R}) = X^0(\mathbb{R})$. This operator turns out to be compact, positive and self-adjoint, and can be represented as an infinite dimensional Matrix $diag\left((k+1/2)^{-1/2}; k = 0, 1, 2, \cdots\right)$. By the definition of Hilbert scale, we can have

(i) $E_{+\infty} = \bigcap_{n=0}^{+\infty} B^n E_0 = \bigcap_{n=0}^{+\infty} X^n(\mathbb{R})$; so this space (center of the Hilbert scale) is just $S(\mathbb{R})$.

(ii) The space E_α equals to the completion of the space $E_{+\infty} = S(\mathbb{R})$ equipped with the norm $\|u\|_\alpha = (B^{-\alpha}u, B^{-\alpha}u)_{L^2}^{1/2}$. From the theory of N-representation of rapid decreasing functions, we can see that $E_{+\infty} = S(\mathbb{R})$ is essentially the space $l_2^{+\infty}(\mathbb{Z}_+)$, and the norm $\|\cdot\|_\alpha$ on $E_{+\infty}$ is just the norm $\|\cdot\|_{X^\alpha}$of the space $X^\alpha(\mathbb{R})$ with restriction on $E_{+\infty} = S(\mathbb{R})$. Since $l_2^{+\infty}(\mathbb{Z}_+)$ is a dense subset of $l_2^{\alpha/2}(\mathbb{Z}_+)$, and further same is true for $S(\mathbb{R}) \subset X^\alpha(\mathbb{R})$, we must have the function space $X^\alpha(\mathbb{R})$ is equivalent to the space E_α in the sense of isomorphism.

Finally, $\bigcap_{s\in\mathbb{R}} X^{2s}(\mathbb{R}) = S(\mathbb{R})$ comes from the fact of $X^s(\mathbb{R}) \subset X^t(\mathbb{R})$ for any $s > t$; $\bigcup_{s\in\mathbb{R}} X^{2s}(\mathbb{R}) = S'(\mathbb{R})$ comes from the N-representation theorem for $S'(\mathbb{R})$. □

3.3 A Hilbert scale

In the last section, it took us a long way to bring out the definition of our function spaces. During the process, we have mentioned some properties of the function spaces $X^{2s}(\mathbb{R})$, like the characterization of the functions in $X^n(\mathbb{R})$ and the relationship of the topological structure between the space $S(\mathbb{R})$ and $X^n(\mathbb{R})$. In these section, we will provide a more systematic description for the properties of these function spaces from the view point of Hilbert scale. At first, we will give out two basic property coming from the property of a Hilbert scale.

Proposition 3.3.1. *For any $s \in \mathbb{R}$, the Hilbert space $X^{-2s}(\mathbb{R})$ is the dual space of the Hilbert space $X^{2s}(\mathbb{R})$.*

Since $X^{2s}(\mathbb{R})$ is isomorphic to the space $l_2^s(\mathbb{Z}_+)$, so we only need to prove that $l_2^{-s}(\mathbb{Z}_+) \cong l_2^s(\mathbb{Z}_+)$, which is obvious. The details are omitted here.

Proposition 3.3.2. *If $s < t$ in $\mathbb{R}$, then the imbedding $i : X^{2t}(\mathbb{R}) \hookrightarrow X^{2s}(\mathbb{R})$ is a compact mapping.*

To prove this proposition, we need a lemma such as the one below. It is a very basic fact in functional analysis and its proof can be found in any textbook.

Lemma 3.3.3. *Let X, Y be two Banach spaces and T is a continuous mapping from X to Y, i.e. $T_n \in \mathcal{L}(X,Y)$. If there is a sequence of continuous finite rank mapping $T_n \in \mathcal{L}(X,Y)$ such that $\|T_n - T\| \longrightarrow 0$ as $n \longrightarrow +\infty$, then the operator T is a compact operator.*

Now let us see how can we get to the proposition.

Proof. Since the space $X^{2s}(\mathbb{R})$ is isomorphic to $l_2^s(\mathbb{Z}_+)$, with a little abuse of notation, we only need to prove the imbedding $i : l_2^t(\mathbb{Z}_+) \hookrightarrow l_2^s(\mathbb{Z}_+)$ is compact. There is a sequence of mapping $i_N : l_2^t(\mathbb{Z}_+) \hookrightarrow l_2^s(\mathbb{Z}_+)$ defined by $i_N(q_0, q_1, q_2, \cdots) = (q_0, q_1, \cdots q_N, 0, 0, \cdots)$. This sequence of mappings are all continuous and they satisfy

$$\begin{aligned}
\lim_{N\to+\infty} \|\imath - i_N\| &= \lim_{N\to+\infty} \inf_{q\neq 0} \left(\frac{\sum_{k\geq N+1}^{+\infty} |q_k|^2 \langle k \rangle^{2s}}{\sum_{k=0}^{+\infty} |q_k|^2 \langle k \rangle^{2t}} \right)^{1/2} \\
&= \lim_{N\to+\infty} \inf_{q\neq 0} \left(\frac{\sum_{k\geq N+1}^{+\infty} |q_k|^2 \langle k \rangle^{2t} \langle k \rangle^{2s-2t}}{\sum_{k=0}^{+\infty} |q_k|^2 \langle k \rangle^{2t}} \right)^{1/2} \\
&\leq \lim_{N\to+\infty} N^{-(t-s)} \\
&= 0 \qquad (\text{since } t > s).
\end{aligned}$$

By the lemma above, we know that all the mappings $i : X^{2t}(\mathbb{R}) \hookrightarrow X^{2s}(\mathbb{R})$ are compact imbeddings. ◻

Another very important property of Hilbert scales is that there usually exist many good continuous linear mapping between the Hilbert spaces in those scales.

Definition 3.3.4. *Given two Hilbert scales $\{E_s\}$, $\{F_s\}$ and a linear map $L : E_{+\infty} \longrightarrow F_{-\infty}$, we denote by $\|L\|_{s_1,s_2} \leq \infty$ its norm as a map $E_{s_1} \longrightarrow F_{s_2}$. We say that the map L defines a* ***morphism of order*** d *of the scales $\{E_s\}$ and $\{F_s\}$ for $s \in [s_0, s_1]$, if $\|L\|_{s,s-d} < +\infty$ for each $s \in [s_0, s_1]$ with some fixed $-\infty \leq s_0 \leq s_1 \leq +\infty$.*

We should be careful on the case of $s_0 = -\infty$ or $s_1 = +\infty$, since $E_{\pm\infty}, F_{\pm\infty}$ are given no norms. So if $s_0 = -\infty$, we apply this definition for $s > s_0$ and similarly $s < +\infty$ if $s_1 = +\infty$. Sometimes this morphism can be invertible or even better, and here we provide below such two definitions.

Definition 3.3.5. *If in addition the inverse map L^{-1} exists and defines a morphism of order $-d$ of the scales $\{F_s\}$, $\{E_s\}$ for $s \in [s_0 + d, s_1 + d]$, we say that L defines an* ***isomorphism of order*** $\mathbf{d}$ *of the two scales. If $\{F_s\} = \{E_s\}$, then an isomorphism is called an* ***automorphism****.*

Definition 3.3.6. *If in addition we have $(Lu, Lu)_{F_{s-d}} = (u, u)_{E_s}$ for any $s \in [s_0, s_1]$ and $u \in E_s$, we say that L defines an* ***isometry of order*** $\mathbf{d}$ *for $s \in [s_0, s_1]$, and the operator L is* ***isometric****. If $\{E_s\}$ and $\{F_s\}$ are complex Hilbert scales, and $L(E_s) = F_{s-d}$, then this isometric operator is said to be* ***unitary****.*

Due to the structure of the Hilbert scale, for any morphism L, it can naturally induce a sequence of adjoint operators on their dual spaces. It turns out that these operators also form a morphism.

Definition 3.3.7. *If $L : E_s \longrightarrow F_{s-d}$ is a morphism of order d for $s \in [s_0, s_1]$, then the adjoint maps $L^* : (F_{s-d})^* = F_{-s+d} \longrightarrow (E_s)^* = E_{-s}$ form a morphism of the scales $\{F_s\}$ and $\{E_s\}$ of the same order d for $s \in [-s_1 + d, -s_0 + d]$. We call it the* ***adjoint morphism****.*

Definition 3.3.8. *A morphism L of a Hilbert scale $\{E_s\}$, complex or real, is called* ***symmetric (antisymmetric)*** *if $L = L^*$ (respectively $L = -L^*$) on the space $E_{+\infty}$. In particular, a linear operator $L : E_{s_0} \longrightarrow E_{s_0-d}$ is called* ***symmetric (anti-symmetric)*** *if $L = L^*$ (respectively $L = -L^*$) on the space $E_{+\infty}$. Furthermore, if L is a symmetric morphism of $\{X_s\}$ of order d for $s \in [s_0, d - s_0]$, then L^* is also a morphism of order d for $s \in [s_0, d - s_0]$ and $L = L^*$ as the scale's morphism. We call such a morphism* ***selfadjoint (Anti-selfadjoint)*** *morphisms are defined similarly.*

To check if a morphism on a complex Hilbert space is symmetric (antisymmetric), we usually do this way: let us introduce a sesquilinear form f on $E_s \times E_{-s} \longrightarrow \mathbb{C} : (u, v) \longmapsto \tau v(u)$, where τ is the natural mapping from $E_{-s} \longrightarrow (E_s)^*$ satisfying $\tau v(\cdot) = (\cdot, v)_{E_s}$; so then we only need check if this equality $f(Lu, v) = f(u, Lv)$ for any $u, v \in E_{+\infty}$.is satisfied.

In the following part, we will discuss on several morphisms on the scales $\{X^s(\mathbb{R})\}$, which is of particular interest from the viewpoint of analysis.

First of all, let us consider the operator $A^{\sigma/2}$. At the beginning of the last section, we have defined the operator $A^{\sigma/2} = \left(\frac{(-\Delta+x^2)}{2}\right)^{\sigma/2}$ as a linear mapping from $X^\sigma(\mathbb{R}) \longrightarrow X^0(\mathbb{R}) = L^2(\mathbb{R})$. Now with the help of the introduction of the Hilbert scale $X^s(\mathbb{R})$, we can generalizing its definition as in this formula: if $\varphi(x) = \sum_{k=0}^{+\infty} q_k h_k(x) \in S'(\mathbb{R})$ (i.e. a tempered distribution), we define

$$A^{\sigma/2}\varphi(x) := \sum_{k=0}^{+\infty} (k + 1/2)^{\sigma/2} q_k h_k(x) \in S'(\mathbb{R}). \tag{3.10}$$

Then we can have such a proposition.

Proposition 3.3.9. *For any $\sigma \in \mathbb{R}$, and $s \in \mathbb{R}$, $A^{\sigma/2} : S'(\mathbb{R}) \longrightarrow S'(\mathbb{R})$ defines an automorphism of order σ on the scales $\{X^s(\mathbb{R})\}$, which is isometric, selfadjoint and satisfying that each operator $A^{\sigma/2} : X^{s+\sigma} \longrightarrow X^s$ is a unitary mapping.*

Proof. It is a straightforward work to check that this operator is an automorphism on the Hilbert scale and each operator is a unitary mapping. To see it is also selfadjoint, let us introduce the sesquilinear mapping on the product space $X^s \times X^{-s}$ for any $s \in \mathbb{R}$:

$$f(u, v) = \sum_{k=0}^{+\infty} a_k \overline{b_k},$$

where $u = \sum_{k=0}^{+\infty} a_k h_k(x) \in X^s(\mathbb{R})$, and $v = \sum_{k=0}^{+\infty} b_k h_k(x) \in X^{-s}(\mathbb{R})$. So we only need $f(Au, v) = f(u, Av)$ is true for any $u \in X^s(\mathbb{R})$,and $v \in X^{-s+\sigma}(\mathbb{R})$, which can be easily verified. □

So until now, we have such three kinds of interpretations for our function space:

(i) $X^s(\mathbb{R}) = \{u \in S'(\mathbb{R}) \mid A^{s/2}u \in L^2(\mathbb{R})\}$;

(ii) $X^s(\mathbb{R}) = \{u \in S'(\mathbb{R}) \mid u = \sum_{k=0}^{+\infty} q_k h_k(x), q = (q_0, q_1, \cdots) \in l_2^{s/2}(\mathbb{Z}_+)\}$;

(iii) $X^s(\mathbb{R}) = \{f \in S'(\mathbb{R}) \mid q_k = f(h_k(x))$ such that $\sum_{k=0}^{+\infty} |q_k|^2 (k + 1/2)^s < +\infty\}$.

In the following let us discuss an interesting automorphism on our Hilbert scale—Fourier transformation. Let us recall that

$$\mathcal{F}(h_k)(x) = (-i)^k h_k(x), \tag{3.11}$$

which can induce such a proposition.

Proposition 3.3.10. *Fourier transformation defines an automorphism of order* 0 *on the Hilbert scale* $\{X^s(\mathbb{R})\}$, *which is surjective and isometric. In other words, each operator* $\mathcal{F} : X^s \longrightarrow X^s$ *is a unitary mapping.*

Proof. For any $\varphi = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^s(\mathbb{R})$, its Fourier transformation is

$\mathcal{F}\varphi = \sum_{k=0}^{+\infty} \left(q_k \left(-i\right)^k \right) h_k(x)$, which is in fact is also in the same space, since

$$\begin{aligned}
\|\mathcal{F}\varphi\|_{X^s} &= \left(\sum_{k=0}^{+\infty} \left| q_k \left(-i\right)^k \right|^2 (k+1/2)^s \right)^{1/2} \\
&= \left(\sum_{k=0}^{+\infty} |q_k|^2 (k+1/2)^s \right)^{1/2} \\
&= \|\varphi\|_{X^s} .
\end{aligned}$$

And if $\psi = \sum_{k=0}^{+\infty} p_k h_k(x)$ is any other element in $X^s(\mathbb{R})$, we have

$$\begin{aligned}
(\mathcal{F}\varphi, \mathcal{F}\psi)_{X^s(\mathbb{R})} &= \sum_{k=0}^{+\infty} q_k \left(-i\right)^{k2} \overline{p_k \left(-i\right)^{k2}} (k+1/2)^s \\
&= \sum_{k=0}^{+\infty} q_k \overline{p_k} (k+1/2)^s \\
&= (\varphi, \psi)_{X^s(\mathbb{R})} .
\end{aligned}$$

Together with the fact that the Fourier transformation is surjective on each space $X^s(\mathbb{R})$, this equality implies the operator is unitary. □

Remark 3.3.11. *In the classical theory of Fourier analysis, we know such two facts:*

(i) Fourier transformation is a unitary mapping on L^2 space;

(ii) Fourier transformation gives homeomorphisms on both Schwartz function space $S(\mathbb{R})$ (with standard topology) and tempered distribution space $S'(\mathbb{R})$ (with weak topology).

So all these facts can be viewed as special cases $s = 0, \pm\infty$ of the above proposition, with the attention that there is no longer inner products on $S(\mathbb{R})$ and $S'(\mathbb{R})$.

Let us recall that our function spaces essentially come from the analysis on the operator $A = (-\Delta + x^2)/2$. Between the operator A and $\mathcal{F}$, we have such an commutative relationship

$$\mathcal{F} \circ (-\Delta_x + x^2)/2 = (-\Delta_\xi + \xi^2)/2,$$

that is, $[\mathcal{F}, A] = 0$. In fact there is more fundamental relationship between them: the Fourier transformation can be represented by the operator A. Besides these a well known fact is that the operators $e^{\imath At}$ form a C_0 group of unitary mappings on L^2 space. All these facts can be uniformly treated together.

Proposition 3.3.12. *For any $t \in \mathbb{R}$, the operator $e^{\imath At}$ defines a surjective isometry on the Hilbert scale $\{X^s(\mathbb{R})\}$. If we fix the space $X^s(\mathbb{R})$ $s \in \mathbb{R}$, the operators $e^{\imath At}$ form a C_0 group of unitary mappings on $X^s(\mathbb{R})$ space.*

Proof. For any $\varphi = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^s(\mathbb{R})$, we have the formula

$$e^{\imath At}\varphi = \sum_{k=0}^{+\infty} q_k e^{\imath \omega_k t} h_k(x). \tag{3.12}$$

Then it is easy to verify that $e^{\imath At}$ is onto on each $X^s(\mathbb{R})$ and $\left(e^{\imath At}\varphi, e^{\imath At}\psi\right)_{X^s} = (\varphi, \psi)_{X^s}$ for any $\varphi, \psi \in X^s(\mathbb{R})$. These facts imply that $e^{\imath At}$ defines an surjective isometry on the Hilbert scale. To prove it in fact induces a C_0 group, the only nontrivial part is to verify that $e^{\imath At}\varphi \longrightarrow \varphi$ as $t \longrightarrow 0$.

For any given $\varphi \in X^s(\mathbb{R})$, we can choose N big enough such that $\left\| \sum_{k \geq N+1}^{+\infty} q_k h_k(x) \right\|_{X^s}^2 < \varepsilon/8$. Then we have

$$\begin{aligned} \left\| e^{\imath At}\varphi - \varphi \right\|_{X^s}^2 &\leq \sum_{k=0}^{N} |q_k|^2 \left| e^{\imath \omega_k t} - 1 \right|^2 + 4 \left\| \sum_{k \geq N+1}^{+\infty} q_k h_k(x) \right\|_{X^s}^2 \\ &< \varepsilon/2 + 4 \cdot \varepsilon/8 = \varepsilon \end{aligned}$$

as $t \to 0$. Thus we have finished our proof. □

Remark 3.3.13. *We can also apply the operator e^{iAt} for the cases $s = \pm\infty$. Although there is no inner product any more, thus not a chance to become a unitary mapping, the operator can still be a isomorphism in each case. Here we assumed that $S(\mathbb{R})$ is given with standard topology and tempered distribution space $S'(\mathbb{R})$ is given with weak topology. Some other special interest is paid to the choice of time t. It is already stated that $\{e^{iAt}\}$ are operators with period 4π satisfying $e^{iAt}|_{t=0} = Id$ and $e^{iAt}|_{t=2\pi} = -Id$. Now let $t = \frac{3}{2}\pi$, from the equality (3.12) we get*

$$\begin{aligned} e^{iAt}\varphi\big|_{t=\frac{3}{2}\pi} &= \sum_{k=0}^{+\infty} q_k \exp\{i(k+1/2)\cdot\frac{3}{2}\pi\}h_k(x) \\ &= \exp\{i\frac{3}{4}\pi\}\sum_{k=0}^{+\infty} q_k(-i)^k h_k(x) \\ &= e^{i\frac{3}{4}\pi}\mathcal{F}(\varphi). \end{aligned}$$

So we have the following corollary.

Corollary 3.3.14. *When $t = \frac{3}{2}\pi$, the operator e^{iAt} turns out to be essentially a Fourier transformation. In particular, for any $\varphi \in S'(\mathbb{R})$*

$$e^{iA\frac{3}{2}\pi}\varphi = e^{i\frac{3}{4}\pi}\mathcal{F}(\varphi). \tag{3.13}$$

Remark 3.3.15. *This result have a little different form in higher dimensional cases: for any $\varphi \in S'(\mathbb{R}^n)$*

$$e^{iA\frac{3}{2}\pi}\varphi = e^{i\frac{3}{4}n\pi}\mathcal{F}(\varphi). \tag{3.14}$$

As a tempered distribution, it admits multiplication with certain smooth functions, for example the function $g(x) = x$; it can also be taken derivatives for arbitrary times. These two kinds of operations can also be regarded as morphisms on the Hilbert scale $\{X^s(\mathbb{R})\}$. Let us denote by $T_x : S'(\mathbb{R}) \longrightarrow S'(\mathbb{R})$ and

$D: S'(\mathbb{R}) \longrightarrow S'(\mathbb{R})$ as the operator of multiplication with function $g(x) = x$ and of taking derivative for one time, then we have such results.

Proposition 3.3.16. *The operator T_x defines a selfadjoint morphism of order 1 on the Hilbert scale $\{X^s(\mathbb{R})\}$.*

Proof. For any $u = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^s(\mathbb{R})$, we have $T_x(u) = \sum_{m=0}^{+\infty} \langle T_x(u), h_m \rangle h_m(x) = \sum_{m=0}^{+\infty} \langle u, xh_m \rangle h_m(x)$. By the property of the Hermite functions, we have

$$xh_m(x) = \sum_{m=0}^{+\infty} \sqrt{\frac{m+1}{2}} h_{m+1}(x) + \sum_{m=1}^{+\infty} \sqrt{\frac{m}{2}} h_{m-1}(x). \tag{3.15}$$

Thus we get

$$T_x(u) = \sum_{m=0}^{+\infty} \left(\sqrt{\frac{m+1}{2}} q_{m+1} + \sqrt{\frac{m}{2}} q_{m-1} \right) h_m(x), \tag{3.16}$$

where q_{-1} is by defaut taken as 0. Then it turns out that $\|T_x(u)\|_{X^{s-1}} \leq \left\| \sum_{m=0}^{+\infty} \sqrt{\frac{m+1}{2}} q_{m+1} h_m(x) \right\|_{X^{s-1}} + \left\| \sum_{m=0}^{+\infty} \sqrt{\frac{m}{2}} q_{m-1} h_m(x) \right\|_{X^{s-1}} \lesssim \|u\|_{X^s}$ and for the sesquilinear mapping f and any $v = \sum_{k=0}^{+\infty} p_k h_k(x) \in X^{1-s}(\mathbb{R})$, we have $f(xu, v) = \sum_{m=0}^{+\infty} \left(\sqrt{\frac{m+1}{2}} q_{m+1} \overline{p_m} + \sqrt{\frac{m}{2}} q_{m-1} \overline{p_m} \right) = f(u, xv)$. □

Proposition 3.3.17. *The operator D defines an anti-selfadjoint morphism of order 1 on the Hilbert scale $\{X^s(\mathbb{R})\}$.*

The proof for this proposition is quite similar to the last one. So we only provided the essential equality needed as

$$Du = \sum_{m=0}^{+\infty} \left(-\sqrt{\frac{m+1}{2}} q_{m+1} + \sqrt{\frac{m}{2}} q_{m-1} \right) h_m(x). \tag{3.17}$$

Remark 3.3.18. *The compositions of operator T_x and D are also morphisms on the same Hilbert scale. For example, a polynomial with order n, which in fact the elements of the tempered distribution, now can be regarded as a morphism of order n; differential operators with polynomial coefficients are also morphisms with finite order. Besides these, the annihilation operator (lowering operator) $a = \frac{1}{i}(D + T_x)$ and creation operator (raising operator) $a^* = \frac{1}{i}(D - T_x)$ are both morphisms of order 1; they are conjugate to each other in the sense that.$f(au, v) = f(u, a^*v)$ for all $(u, v) \in X^s(\mathbb{R}) \times X^{1-s}(\mathbb{R})$.*

In a Hilbert scale, the space $E_{\pm\infty}$ play two special roles in this family of Hilbert spaces. One natural question is what are the relationships of these two spaces and the other Hilbert spaces such as topological structure, characterization of continuous linear mappings etc. In our case, the space $X^{\pm\infty}(\mathbb{R}) = S(\mathbb{R})$ or $S'(\mathbb{R})$ are important to the analysis and have been studied heavily. Now the introduction of the Hilbert scale provides us a new point of view to understand the spaces $S(\mathbb{R})$ and $S'(\mathbb{R})$. In the following part of this section, we will first use this new viewpoint to review Barry Simon's work on the N-representation theory with application in the analysis on the function spaces $S(\mathbb{R})$ and $S'(\mathbb{R})$. After that, we will study the relationship of the topological structures of these two spaces and the Hilbert function spaces $X^s(\mathbb{R})$.

In 1971, Barry Simon's once used the N-representation theory to analyze the structure of the function space $S(\mathbb{R})$ and $S'(\mathbb{R})$ (see [Si]). From the standpoint of Hilbert scale $\{X^s(\mathbb{R})\}$, we can easily rewrite most of his results as follows:

- Each Hilbert space $X^s(\mathbb{R})$ is separable, and $S(\mathbb{R}) = X^{+\infty}(\mathbb{R})$ is also separable under its standard topology.

• Similarly, since $\sum_{k=0}^{N} q_k h_k(x) \longrightarrow \sum_{k=0}^{+\infty} q_k h_k(x)$ in the weak topology $\sigma(S'(\mathbb{R}), S(\mathbb{R}))$, $S'(\mathbb{R}) = X^{-\infty}(\mathbb{R})$ with the weak topology is also separable.

• In particular $S(\mathbb{R}) = X^{+\infty}(\mathbb{R})$ is dense in $S'(\mathbb{R}) = X^{-\infty}(\mathbb{R})$ with weak topology.

• For any function $f \in S'(\mathbb{R})$. Then $\exists s \in \mathbb{R}, m \in \mathbb{Z}_+$ and a continuous bounded function $g \in X^m(\mathbb{R})$ such that $f = A^s g$. (In the next section we will discuss the relationship between the space $X^s(\mathbb{R})$ and Sobolev space, and see a more dedicated description on the function g)

• Under the standard topology, a subset $B \subset S(\mathbb{R})$ is bounded if and only if B is bounded in each Hilbert space $X^n(\mathbb{R}), n \in \mathbb{Z}_+$.

• Under the weak topology $\sigma(S'(\mathbb{R}), S(\mathbb{R}))$, a subset $C \subset S(\mathbb{R})$ is bounded if and only if $\exists m \in \mathbb{Z}_+$ such that C is a bounded set in Hilbert space $X^{-m}(\mathbb{R})$. (in fact, this is also true for Mackey topology or strong topology)

• Any closed and bounded subset of $S(\mathbb{R}) = X^{+\infty}(\mathbb{R})$ in the standard topology is a compact set.

Now let us further consider the choice of the topological structure on the spaces $X^{\pm\infty}(\mathbb{R})$. In the case of the space $S(\mathbb{R})$, we have proved that if it is endowed with the weakest topology such that $i_n : S(\mathbb{R}) \longrightarrow X^n(\mathbb{R})$ are all continuous then that topology is just the usual Fréchet topology on it. This is still true when we considering more imbeddings i_s.

Theorem 3.3.19. *If we endow the space $S(\mathbb{R}) = X^{+\infty}(\mathbb{R})$ with the weakest topology such that $i_s : S(\mathbb{R}) \longrightarrow X^s(\mathbb{R})$ are all continuous for $s \in \mathbb{R}$, then that topology is* *just the usual Fréchet topology on it. And it has such a property: for any topological space E, a mapping $f : E \longrightarrow S(\mathbb{R})$ is continuous if and only if for any $s \in \mathbb{R}$ the mapping $i_s \circ f : E \longrightarrow X^s(\mathbb{R})$ is continuous.*

Proof. For the first statement it is enough to prove that when we endow the space $S(\mathbb{R})$ with the weakest topology such that $i_n : S(\mathbb{R}) \longrightarrow X^s(\mathbb{R})$ are all continuous for $n \in \mathbb{Z}_+$, then it makes the mappings i_s continuous. This is easy. Since for any $s \in \mathbb{R}$, there is an $n \in \mathbb{Z}_+$ bigger than s, so the mapping i_s can be written as the composition of the mappings i_n and the compact imbedding form the space $X^n(\mathbb{R})$ to $X^s(\mathbb{R})$, thus it must be continuous.

For the second statement, if we have $f : E \longrightarrow S(\mathbb{R})$ is continuous, then of course $i_s \circ f : E \longrightarrow X^s(\mathbb{R})$ is also continuous. Reversely, with the given topology any open set U on $S(\mathbb{R})$ can be written in the form of the union of the sets in the form of $U_{s_1 s_2 \cdots s_m} = \{u \in S(\mathbb{R}) \mid \|u\|_{s_1} < \varepsilon_1, \|u\|_{s_2} < \varepsilon_2, \cdots, \|u\|_{s_m} < \varepsilon_m\}$, then if every $i_s \circ f : E \longrightarrow X^s(\mathbb{R})$ is continuous we have $f^{-1}(U)$ is the union of the sets in the form of $f^{-1}(U_{s_1 s_2 \cdots s_m})$. But we have $f^{-1}(U_{s_1 s_2 \cdots s_m}) = \bigcap_{j=1}^{m} \left(i_{s_j} \circ f\right)^{-1} (\{u \in S(\mathbb{R}) \mid \|u\|_{s_j} < \varepsilon_j\})$, which is always open, so the set $f^{-1}(U)$ must be open. That means the second statement is also true. □

When we consider the similar question for the case of the space $S'(\mathbb{R})$, the situation becomes a little more complicated. Unlike the case of compactly supported functions the inductive limit is not a possible choice. Although we have $X^{-\infty}(\mathbb{R}) = \bigcap_{s \in \mathbb{R}} X^s(\mathbb{R})$, the restriction of the topology of $X^t(\mathbb{R})$ to $X^s(\mathbb{R})$ (assuming $t < s$) is not the given topology on $X^s(\mathbb{R})$, thus the necessary setting for inductive limit topology is missing. But we still have many possible choices such as the weak topology $\sigma(S'(\mathbb{R}), S(\mathbb{R}))$, Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$, and strong topology $\beta(S'(\mathbb{R}), S(\mathbb{R}))$. Now let us choose the topology (denoted as $\widetilde{\tau}$) as the strongest locally convex topology on $S'(\mathbb{R})$ so that the injections $i_s : X^s(\mathbb{R}) \longrightarrow X^{-\infty}(\mathbb{R})$ are continuous, then it will finally turn out that this topology is just equal to $\tau(S'(\mathbb{R}), S(\mathbb{R}))$ and $\beta(S'(\mathbb{R}), S(\mathbb{R}))$.

Below let us start with the existence of such a topology. It is useful to provide such a characterization of the continuous linear mapping between two locally convex topological spaces.

Lemma 3.3.20. *Let E and F be locally convex spaces with families of semi-norms $\{\rho_\alpha\}_{\alpha\in A}$ and $\{d_\beta\}_{\beta\in B}$. Then a linear mapping map $T : E \longrightarrow F$, is continuous if and only if for any $\beta \in B$, there are $\alpha_1, \cdots . \alpha_n \in A$ and $C > 0$ with*

$$d_\beta(Tx) \leq C \left(\rho_{\alpha_1}(x) + \cdots + \rho_{\alpha_n}(x)\right). \tag{3.18}$$

This lemma is basic in the theory of locally convex spaces and can be found in many textbooks, so we don't provide its proof here. With the help of this lemma, we can guarantee the existence of the topology $\widetilde{\tau}$.

Proposition 3.3.21. *The topology $\widetilde{\tau}$ exists and it satisfies: for any locally convex space F, a linear mapping $g : S'(\mathbb{R}) \longrightarrow F$ is continuous if and only if for any $s \in \mathbb{R}$ the mapping $g|_{X^s} = g_s : X^s(\mathbb{R}) \longrightarrow F$ is continuous.*

Proof. According to the lemma, if a locally convex space $S'(\mathbb{R})$ with a family of seminorms $\{d_\beta\}_{\beta\in B}$ satisfy that all the inclusions $i_s : X^s(\mathbb{R}) \longrightarrow X^{-\infty}(\mathbb{R})$ are continuous, then for each $\beta \in B$, and $s \in \mathbb{R}$, there exists$C > 0$ with $d_\beta(i_s(x)) \leq C \|x\|_{X^s}$. Now let us consider the sets of all the possible seminorms $\Im = \{\text{seminorm } d \mid \forall s \in \mathbb{R}, \exists C_s \in \mathbb{R}$

so that $d_\beta(i_s(x)) \le C_s \|x\|_{X^s} \}$. This set is not empty and it can separate points in $S'(\mathbb{R})$, since for any function $u \in S(\mathbb{R})$, the seminorm $d_u(x) = |x(u)|$ is an element in $\Im$. Then the locally convex space with such a family of seminorms will induce the desired topology $\widetilde{\tau}$. It is strongest in the meaning: any other locally convex topology on $S'(\mathbb{R})$ to make all the inclusions $\imath_s : X^s(\mathbb{R}) \longrightarrow X^{-\infty}(\mathbb{R})$ continuous must be a weaker topology than it.

For the second statement if we have a locally convex space F and a continuous mapping $g : S'(\mathbb{R}) \longrightarrow F$, then $g|_{X^s} = g \circ i_s$ is of course continuous too. Reversely, if we have a locally convex space F a mapping $g : S'(\mathbb{R}) \longrightarrow F$ so that for any $s \in \mathbb{R}$ the mapping $g|_{X^s} = g_s : X^s(\mathbb{R}) \longrightarrow F$ is continuous, then we want to prove that for any open set U of F containing zero point, we can find an open set V in $\widetilde{\tau}$ satisfying $g(V) \subset U$. Without loss of generality let us assume U is a circled convex open set of F, then $g^{-1}(U) = \bigcup_{s\in\mathbb{R}} (g|_{X^s})^{-1}(U)$ is a circled convex set containing zero point in $S'(\mathbb{R})$ and for each $s \in \mathbb{R}$, $(g|_{X^s})^{-1}(U)$ is an open set in $X^s(\mathbb{R})$. It is easy to verify that $g^{-1}(U)$ is an absorbing subset of $S'(\mathbb{R})$ with the property that if $x \in g^{-1}(U)$ then $tx \in g^{-1}(U)$ for all $0 \le t \le 1$. Then the Minkowski functional $\widetilde{d}(x) = \inf\{\lambda \,|\, x \in \lambda g^{-1}(U)\}$ is a seminorm on $S'(\mathbb{R})$. We claim that it is in the set $\Im$.

Since for any real number s we have $g^{-1}(U) \cap X^s(\mathbb{R}) = (g|_{X^s})^{-1}(U)$, which is an open set in of $X^s(\mathbb{R})$ due to continuity of the mapping $g|_{X^s}$, we can always find a open set of $X^s(\mathbb{R})$ satisfying $\{x \,|\, \|x\|_s < \delta_s\} \subset g^{-1}(U) \cap X^s(\mathbb{R})$. Then we have $\widetilde{d}(i_s(x)) = \inf\{\lambda \,|\, x \in \lambda\,(g|_{X^s})^{-1}(U)\} \le \inf\{\lambda \,|\, x \in \lambda\{\|x\|_s < \delta_s\}\}$ is true for all $x \in X^s(\mathbb{R})$, which implies that $\widetilde{d}(i_s(x)) \le \delta_s^{-1} \|x\|_s$. So our claim has been proved. In this way, we can find an open set $\{x \in S'(\mathbb{R}) \,|\, \widetilde{d}(x) < 1\} \subset g^{-1}(U)$ which is mapped into the open set U of F. That means the mapping $g : S'(\mathbb{R}) \longrightarrow F$ is continuous. □

The remaining task is to identify this topology. From the proof above, it can be deduced that the weak topology $\sigma(S'(\mathbb{R}), S(\mathbb{R}))$ is weaker than the topology $\widetilde{\tau}$, since all the seminorms $d_u(x) = |x(u)|$ are in $\Im$ and these norms can define the topology $\sigma(S'(\mathbb{R}), S(\mathbb{R}))$. Another comparison of the topologies can be made between $(S'(\mathbb{R}), \widetilde{\tau})$ and the strong topology $\beta(S'(\mathbb{R}), S(\mathbb{R}))$. The latter one can be generated by the seminorms $\{\rho_A \,|\, A \subset S(\mathbb{R}) \text{ is bounded under standard topology}\}$ where $\rho_A(f) = \sup_{x\in A} |f(x)|$. Due to the N-representation theory of the function space $S(\mathbb{R})$, any bounded set on $S(\mathbb{R})$ must be bounded in each Hilbert space $X^n(\mathbb{R}), n \in \mathbb{Z}_+$. So we have for any $n \in \mathbb{Z}_+$

$$\rho_A(f) = \sup_{x\in A} |f(x)| \leq \|f\|_{X^{-n}} \sup_{x\in A} \|x\|_{X^n} \lesssim \|f\|_{X^{-n}},$$

and it can be further improved as for any $s \in \mathbb{R}$ $\rho_A(f) \lesssim \|f\|_{X^{-s}}$ by using the compact imbedding property of the Hilbert scale $\{X^s\}$. Recall the definition of the family of norms $\Im$, we know that all these ρ_A are also in $\Im$, which implies that the strong topology $\beta(S'(\mathbb{R}), S(\mathbb{R}))$ is also weaker than the topology $\widetilde{\tau}$ on $S'(\mathbb{R})$.

When we compare the topology $(S'(\mathbb{R}), \widetilde{\tau})$ and the Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$, let us consider the dual space of $(S'(\mathbb{R}), \widetilde{\tau})$. We find such a proposition.

Proposition 3.3.22. *The dual space of $(S'(\mathbb{R}), \widetilde{\tau})$ is just the function space $S(\mathbb{R})$.*

Proof. Let T is a continuous linear functional on $(S'(\mathbb{R}), \widetilde{\tau})$. By the proposition above, we have that the restrictions $T|_{X^s} : X^s(\mathbb{R}) \longrightarrow \mathbb{C}$ are all continuous, which means $T|_{X^s}$ is essentially an element in $X^s(\mathbb{R})$. So there exists $b_s = (b_{s0}, b_{s1}, \cdots) \in$

$l_2^{-s/2}(\mathbb{Z}_+)$ such that $T|_{X^s}(f) = \sum_{k=0}^{+\infty}(\overline{b_{sk}}\, q_k)$ where $f \in X^s(\mathbb{R})$ and q_k is the k-th Hermite coefficient of f. Notice this is true for every $s \in \mathbb{R}$, we can deduce that all these b_s are equal and they correspond to a unique element in $S(\mathbb{R})$, which implies that $T(f) = T_\varphi(f) = f(\varphi)$ for some $\varphi \in S(\mathbb{R})$. Reversely for any $\varphi \in S(\mathbb{R})$, it induces a linear mapping $T_\varphi(f) = f(\varphi)$ on $S'(\mathbb{R})$. For any $s \in \mathbb{R}$ we have $|T_\varphi(f)| \leq \|f\|_{X^s}$ $\|\varphi\|_{X^{-s}}$, so T_φ is continuous on the topological space $(S'(\mathbb{R}), \widetilde{\tau})$. $\square$

Recall that according to Mackey-Arens theorem, Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$ is the strongest locally convex topology $S'(\mathbb{R})$ having its topological dual as $S(\mathbb{R})$. So the proposition above implies that the topology $\widetilde{\tau}$ is weaker than the Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$.

Let us review that Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$ is the locally convex topology of uniform convergence on $\sigma(S(\mathbb{R}), S'(\mathbb{R}))$-compact convex sets of $S(\mathbb{R})$, and its generating seminorms can be chosen as $\rho_C(x) = \sup_{y \in C} |y(x)|$ with C running over all the $\sigma(S(\mathbb{R}), S'(\mathbb{R}))$-compact convex subset of $S(\mathbb{R})$; the strong topology $\beta(S'(\mathbb{R}), S(\mathbb{R}))$ is the locally convex topology of uniform convergence on bounded subsets of $S(\mathbb{R})$ with usual topology, and its generating seminorms can be chosen as $\rho_A(f) = \sup_{x \in A} |f(x)|$ with A running over all the bounded sets of $S(\mathbb{R})$. Since any $\sigma(S(\mathbb{R}), S'(\mathbb{R}))$-compact convex subset of $S(\mathbb{R})$ is bounded, thus we have the strong topology $\beta(S'(\mathbb{R}), S(\mathbb{R}))$ is stronger than the Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$. But we have proved that $\widetilde{\tau}$ is stronger than the strong topology, so we finally get such a relationship $\beta(S'(\mathbb{R}), S(\mathbb{R})) \prec \widetilde{\tau} \prec \tau(S'(\mathbb{R}), S(\mathbb{R})) \prec \beta(S'(\mathbb{R}), S(\mathbb{R}))$ (with $\prec$ means "weaker than"), which implies that $\widetilde{\tau}$ coincides with the Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$.and the strong topology $\beta(S'(\mathbb{R}), S(\mathbb{R}))$.

In summary, we have such a theorem for the topology $\widetilde{\tau}$ on $S'(\mathbb{R})$.

Theorem 3.3.23. *If we endow the space $S'(\mathbb{R}) = X^{-\infty}(\mathbb{R})$ with the strongest locally convex topology $\widetilde{\tau}$ such that all the injections $i_s : X^s(\mathbb{R}) \longrightarrow X^{-\infty}(\mathbb{R})$ are continuous, then that topology coincides with the Mackey topology $\tau(S'(\mathbb{R}), S(\mathbb{R}))$ and the strong topology $\beta(S'(\mathbb{R}), S(\mathbb{R}))$. And it has such a property: for any locally convex space F, a linear mapping $g : S'(\mathbb{R}) \longrightarrow F$ is continuous if and only if for any $s \in \mathbb{R}$ the mapping $g|_{X^s} = g_s : X^s(\mathbb{R}) \longrightarrow F$ is continuous.*

3.4 Relationship with Sobolev spaces

In this section, we will discuss the relationship of the function spaces $X^s(\mathbb{R})$ and Sobolev spaces $H^s(\mathbb{R})$. From the characterization of the function space $X^n(\mathbb{R})$ (n nonnegative integer), we can easily deduce that it is a subspace of the Sobolev space $H^n(\mathbb{R})$. Meanwhile, $X^n(\mathbb{R})$ is a space invariant under the Fourier transformation, so a natural question is the relationships between these two spaces and the Fourier transformation.

At first let us compare these two spaces $X^1(\mathbb{R})$ (virial space) and $H^1(\mathbb{R})$. It turns out the former one is a proper subset of the latter one. To see this, consider the function $u(x) = \chi(x)\frac{1}{x}$ where $0 \leq \chi(x) \leq 1$ is a cut off function satisfying $\chi(x) = 0$ for $|x| \leq 1/2$ and $\chi(x) = 1$ for $|x| \geq 1$. This is a smooth function which belongs to L^2 space. By using the Leibniz rule, we can confirm that $u(x)$ is in the space $H^{+\infty}(\mathbb{R}) = \bigcap_{n=0}^{+\infty} H^n(\mathbb{R})$. But on the other hand this is a function not in $X^1(\mathbb{R}) = \{f \in S' | f \in H^1(\mathbb{R}),\ xf \in L^2(\mathbb{R})\}$, since $xu(x) = \chi(x) = 1$ for all $|x| \geq 1$ which is not in L^2 space and Fourier transformation is a unitary mapping on $X^1(\mathbb{R})$.

In fact, we can conclude that $S(\mathbb{R}) = X^{+\infty}(\mathbb{R}) \subsetneq H^{+\infty}(\mathbb{R})$.

From this example, we can see that the Hilbert scale $\{X^s(\mathbb{R})\}$ provide a different way from Sobolev spaces to measure regularities of functions. A function, like $u(x)$ above, may be very smooth in Sobolev meaning but not so regular in the meaning of $X^s(\mathbb{R})$. This difference comes from that the regularity in the meaning of $X^n(\mathbb{R})$ requires not only the information of local regularity—to how many times the function can be taken derivatives to get an L^2 function, it also requires the information of global regularity—how fast it together with its derivatives can fall off comparing the inverse of polynomials.

This viewpoint can bring us some other advantages. As we know, the Sobolev spaces are a good way to understand the regularity of some tempered distribution, but only some of them. This is due to the fact that $H^{-\infty}(\mathbb{R}) = \bigcap_{s\in\mathbb{R}} H^s(\mathbb{R}) \subsetneq S'(\mathbb{R})$. But now the Hilbert scale $\{X^s(\mathbb{R})\}$ enables us to discuss the regularity of every tempered distribution. In particular, let us see such an example $u(x) \equiv (2\pi)^{-1/2}$. This is a function very easy to understand, but we can't discuss its regularity in Sobolev meaning, since we have $\mathcal{F}u(\xi) = \delta(\xi)$ and $\mathcal{F}u(\xi)\langle\xi\rangle^s = \delta(\xi) \notin L^2$ which implies that it is never in $H^{-\infty}(\mathbb{R})$. Now let us reconsider this by using the Hilbert scale $\{X^s(\mathbb{R})\}$. The k-th Hermite coefficient of $\mathcal{F}u(\xi)$ is

$$\begin{aligned} q_k &= \langle \delta(\xi), h_k(\xi)\rangle = h_k(0) \\ &= \begin{cases} 0 & \text{if} \quad k \text{ odd} \\ \pi^{-1/4}(-1)^{k/2}\left(\frac{k!}{2^k(k/2)!(k/2)!}\right)^{1/2} & \text{if} \quad k \text{ even.} \end{cases} \end{aligned} \tag{3.19}$$

From the Sterling's formula, we get that $q_{2k} \approx \pi^{-1/2}(-1)^{k/2}\langle k\rangle^{-1/4}$ and further $\mathcal{F}u(\xi) \in X^s(\mathbb{R})$ for any $s < -1/2$ (denoted as $\mathcal{F}u(\xi) \in X^{(-1/2)^-}(\mathbb{R})$). So we can say all the nonzero constant functions have regularity degrees as $(-1/2)^{-}$. Similarly, all the polynomials of order n can be viewed as tempered distribution and they have the regularity degrees $(-1/2)^{-} - n$.

Although we have mentioned many differences between the function spaces $X^s(\mathbb{R})$ and $H^s(\mathbb{R})$, they do have close relationships. When $s = 0$, they are the same function space $L^2(\mathbb{R})$; when $s = 1$, $X^1(\mathbb{R}) = \{f \in S' | f \in H^1(\mathbb{R}), xf \in L^2(\mathbb{R})\}$ (virial space) in fact equals to $H^1(\mathbb{R}) \cap \mathcal{F}(H^1(\mathbb{R}))$. This proposition can be generalized. It turns out that, we have such a theorem.

Theorem 3.4.1. *For all real number $s \geq 0$, we have $X^s(\mathbb{R}) \subset H^s(\mathbb{R}) \cap \mathcal{F}(H^s(\mathbb{R}))$, and the norm $\|\cdot\|_{X^s}$ is stronger than the norm $\|u\|_s = \left(\|u\|_{H^s}^2 + \|\langle x\rangle^s u(x)\|_{L^2}^2\right)^{1/2}$. In particular, for any nonnegative integer n, we have $X^n(\mathbb{R}) = H^n(\mathbb{R}) \cap \mathcal{F}(H^n(\mathbb{R}))$ and the norm $\|\cdot\|_{X^n}$ is equivalent to the norm $\|u\|_n = \left(\|u\|_{H^n}^2 + \|\langle x\rangle^n u(x)\|_{L^2}^2\right)^{1/2}$. For any real number $-s \leq 0$, we have $H^{-s}(\mathbb{R}) + \mathcal{F}(H^{-s}(\mathbb{R})) \subset X^{-s}(\mathbb{R})$, and it is a dense subset of latter space. In other words, $\overline{H^{-s}(\mathbb{R}) + \mathcal{F}(H^{-s}(\mathbb{R}))} = X^{-s}(\mathbb{R})$. In particular, when $-n$ is a negative integer, we have $H^{-n}(\mathbb{R}) + \mathcal{F}(H^{-n}(\mathbb{R}))$ is in fact the space $X^{-n}(\mathbb{R})$, and the norm $\|\cdot\|_{X^n}$ is equivalent to the norm $\|u\|_n = \inf_{u_1+u_2}(\|u_1\|_{H^{-n}} + \|\langle x\rangle^n u_2(x)\|_{L^2})$.*

This theorem can be generalized to higher dimensional cases. In the following subsections, we will first prove this theorem for the case of nonnegative integers, and then use the method of complex interpolation theory to prove the case of nonnegative real numbers. Finally we will discuss the case of nonpositive real numbers, which is in fact in the dual form of the case $s \geq 0$.

3.4.1 Nonnegative integer cases

In this subsection, we will rewrite the theorem in the case of $n \in \mathbb{Z}_+$ and then prove it.

Theorem 3.4.2. *For any nonnegative integer n, we have $X^n(\mathbb{R}) = H^n(\mathbb{R}) \cap \mathcal{F}(H^n(\mathbb{R}))$, and the norm $\|\cdot\|_{X^n}$ is equivalent to the norm $\|u\|_n = \left(\|u\|_{H^n}^2 + \|\langle x\rangle^n u(x)\|_{L^2}^2\right)^{1/2}$, or written as $\|u\|_n = \left(\|u\|_{H^n}^2 + \|\mathcal{F}u\|_{H^n}^2\right)^{1/2}$.*

Proof. Recall that we have the characterization of the function space $X^n(\mathbb{R})$ as $X^n(\mathbb{R}) = \{u \in S'(\mathbb{R}) |\, x^\alpha \left(\frac{d}{dx}\right)^\beta u \in L^2(\mathbb{R})$ for all indices $0 \leq \alpha + \beta \leq n\}$, which implies that it is a subspace of the function space $H^n(\mathbb{R})$. Meanwhile, since the Fourier transformation is a unitary mapping on the space $X^n(\mathbb{R})$, then it must be true that $X^n(\mathbb{R}) = \mathcal{F}(X^n(\mathbb{R})) \subset \mathcal{F}(H^n(\mathbb{R}))$. Reversely, we have $H^n(\mathbb{R}) \cap \mathcal{F}(H^n(\mathbb{R})) = \{u \in S'(\mathbb{R}) |\, x^\alpha u, \left(\frac{d}{dx}\right)^\alpha u \in L^2(\mathbb{R})$ for all indices $0 \leq \alpha \leq n\}$, then we need to show that it implies that $x^\alpha \left(\frac{d}{dx}\right)^\beta u \in L^2(\mathbb{R})$ for all indices $0 \leq \alpha + \beta \leq n$.

Noticing that $\|\cdot\|_{X^n}$ has an equivalent norm in the form of $\left(\sum_{\alpha+\beta\leq n} \|x^\alpha u^{(\beta)}\|_{L^2}^2\right)^{1/2}$ and $\left(\|u\|_{H^n}^2 + \|\langle x\rangle^n u(x)\|_{L^2}^2\right)^{1/2} \approx \left(\sum_{\alpha\leq n}\left(\|x^\alpha u\|_{L^2}^2 + \|u^{(\alpha)}\|_{L^2}^2\right)\right)^{1/2}$, so we only need to prove that

$$\sum_{\alpha+\beta\leq n} \|x^\alpha u^{(\beta)}\|_{L^2}^2 \approx \sum_{\alpha\leq n}\left(\|x^\alpha u\|_{L^2}^2 + \|u^{(\alpha)}\|_{L^2}^2\right). \tag{3.20}$$

Obviously the left side term is not less than the right side term. Another observation is that the function space $S(\mathbb{R})$ is both a dense subset of the space $X^n(\mathbb{R})$ and $H^n(\mathbb{R}) \cap \mathcal{F}(H^n(\mathbb{R}))$ with the concerned norm respectively. So the only thing left is to prove that

$$\sum_{\alpha\leq n}\left(\|x^\alpha u\|_{L^2}^2 + \|u^{(\alpha)}\|_{L^2}^2\right) \gtrsim \sum_{\alpha+\beta\leq n} \|x^\alpha u^{(\beta)}\|_{L^2}^2 \tag{3.21}$$

is true for every function in $u \in S(\mathbb{R})$.

Below let us see a few cases with n small and then prove it for any positive integers by induction method. When $n = 0, 1$, this inequality are very clear. Now let $n = 2$. There is only one term $\left\|x\frac{d}{dx}u\right\|_{L^2}^2$ at the right side which doesn't appear on the left side. But we have

$$\begin{aligned}
\left\|x\frac{d}{dx}u\right\|_{L^2}^2 &= \left(x\frac{d}{dx}u,\, x\frac{d}{dx}u\right) \quad \text{(where } (\cdot,\cdot) \text{ denote the inner product } (\cdot,\cdot)_{L^2}\text{)} \\
&= \left(-\left(\frac{d}{dx}\right)^2 u,\, x^2u\right) - 2\left(\frac{d}{dx}u,\, xu\right) \\
&\leq \left\|\left(\frac{d}{dx}\right)^2 u\right\|_{L^2} \left\|x^2u\right\|_{L^2} + 2\left\|\frac{d}{dx}u\right\|_{L^2} \|xu\|_{L^2} \\
&\leq \frac{1}{2}\left(\left\|\left(\frac{d}{dx}\right)^2 u\right\|_{L^2}^2 + \left\|x^2u\right\|_{L^2}^2\right) + \left\|\frac{d}{dx}u\right\|_{L^2}^2 + \|xu\|_{L^2}^2,
\end{aligned}$$

which implies the inequality (3.21) is true in the case $n = 2$.

For the case $n = 3$, we only need estimate the term $\left\|x^2\frac{d}{dx}u\right\|_{L^2}^2$ and $\left\|x\left(\frac{d}{dx}\right)^2 u\right\|_{L^2}^2$. We have

$$\begin{aligned}
\left\|x^2\frac{d}{dx}u\right\|_{L^2}^2 &= \left(x^2\frac{d}{dx}u,\, x^2\frac{d}{dx}u\right) \\
&= -4\left(x\frac{d}{dx}u, x^2u\right) - \left(x\left(\frac{d}{dx}\right)^2 u, x^3\frac{d}{dx}u\right) \\
&\leq \frac{1}{2}\left(x\left(\frac{d}{dx}\right)^2 u, x\left(\frac{d}{dx}\right)^2 u\right) + \frac{1}{2}\left(x^3u, x^3u\right) \\
&+ 4\cdot\frac{1}{2}\left(\left(x\frac{d}{dx}u, x\frac{d}{dx}u\right) + \left(x^2u, x^2u\right)\right) \\
&\leq \frac{1}{2}\left(x\left(\frac{d}{dx}\right)^2 u, x\left(\frac{d}{dx}\right)^2 u\right) + \frac{1}{2}\left(x^3u, x^3u\right) \\
&+ C\textstyle\sum_{\alpha\leq 2}\left(\|x^\alpha u\|_{L^2}^2 + \left\|u^{(\alpha)}\right\|_{L^2}^2\right).
\end{aligned}$$

And we also have the estimate for the first term in the last line

$$\begin{aligned}\left(x\left(\frac{d}{dx}\right)^2u, x\left(\frac{d}{dx}\right)^2u\right) &= -\left(\left(\frac{d}{dx}\right)^3u, x^2\frac{d}{dx}u\right) - 2\left(\left(\frac{d}{dx}\right)^2u, x\frac{d}{dx}u\right)\\ &\leq \frac{1}{2}\left(\left(\frac{d}{dx}\right)^3u, \left(\frac{d}{dx}\right)^3u\right) + \frac{1}{2}\left(x^2\frac{d}{dx}u, x^2\frac{d}{dx}u\right)\\ &+ C\textstyle\sum_{\alpha\leq 2}\left(\|x^\alpha u\|_{L^2}^2 + \left\|u^{(\alpha)}\right\|_{L^2}^2\right).\end{aligned}$$

By using the inequalities above we can get

$$\begin{aligned}\left(x^2\frac{d}{dx}u, x^2\frac{d}{dx}u\right) &\leq \frac{1}{4}\left(x^2\frac{d}{dx}u, x^2\frac{d}{dx}u\right) + \frac{1}{4}\left(\left(\frac{d}{dx}\right)^3u, \left(\frac{d}{dx}\right)^3u\right)\\ &+ \frac{1}{2}\left(x^3u, x^3u\right) + C\textstyle\sum_{\alpha\leq 2}\left(\|x^\alpha u\|_{L^2}^2 + \left\|u^{(\alpha)}\right\|_{L^2}^2\right),\end{aligned}$$

which implies that

$$\left\|x^2\frac{d}{dx}u\right\|_{L^2}^2 \lesssim \textstyle\sum_{\alpha\leq 3}\left(\|x^\alpha u\|_{L^2}^2 + \left\|u^{(\alpha)}\right\|_{L^2}^2\right). \tag{3.22}$$

Together with the fact that $\left(\frac{d}{dx}\right)^3 u$ is in the L^2-space, we can deduce from the assumption in the induction method that all the terms $x^\alpha u^{(\beta)}$ with $\alpha+\beta = 3$ and $\beta \geq 1$ are also in the L^2-space and their norms can be controlled by $\left(\sum_{\alpha\leq 3}\left(\|x^\alpha u\|_{L^2}^2 + \left\|u^{(\alpha)}\right\|_{L^2}^2\right)\right)^{\frac{1}{2}}$. Thus the inequality (3.21) is true in the case $n = 3$.

Through observation of the index of the derivative part in the process above, we can find the estimate of the term $\left(x^2\frac{d}{dx}u,\, x^2\frac{d}{dx}u\right)$ has experienced such a chain of processes: $1 \xrightarrow{\times 2} 2 \xrightarrow{\times 2} 1 (\bmod 3)$. This notation will be used in the following proof.

For the case odd $n = 2m+1$, we can obtain such a chain $1 \xrightarrow{\times 2} 2 \xrightarrow{\times 2} \cdots \xrightarrow{\times 2} 1 (\bmod (2m+1))$ in finite steps (assuming the step in the chain is k). Its corresponding process provides us the estimate

$$\left\|x^{2m}\frac{d}{dx}u\right\|_{L^2}^2 \lesssim \textstyle\sum_{\alpha\leq 2m+1}\left(\|x^\alpha u\|_{L^2}^2 + \left\|u^{(\alpha)}\right\|_{L^2}^2\right). \tag{3.23}$$

Then by taking the function $\frac{d}{dx}u$ as a new function, we can deduce from the induction's assumption that all the terms $x^\alpha u^{(\beta)}$ with $\alpha+\beta = 2m+1$ and $\beta \geq 1$ are also in the L^2-space and their norms can be controlled by $\left(\sum_{\alpha\leq 3}\left(\|x^\alpha u\|_{L^2}^2 + \|u^{(\alpha)}\|_{L^2}^2\right)\right)^{1/2}$. So our claim in the theorem is true for $n = 2m+1$.

For the case even $n = 2m$, we can first obtain the estimate of the term $x^m u^{(m)}$

$$\left(x^m u^{(m)}, x^m u^{(m)}\right) \leq \frac{1}{2}\left(x^{2m}u, x^{2m}u\right) + \frac{1}{2}\left(u^{(2m)}, u^{(2m)}\right) + C\sum_{\alpha\leq 2m-1}\left(\|x^\alpha u\|_{L^2}^2 + \|u^{(\alpha)}\|_{L^2}^2\right).$$

Then by taking the function $x^m u$ and $u^{(m)}$ as two new functions, we can deduce from the induction's assumption that all the terms $x^\alpha u^{(\beta)}$ with $\alpha+\beta = 2m$ are also in the L^2-space and their norms can be controlled by $\left(\sum_{\alpha\leq 3}\left(\|x^\alpha u\|_{L^2}^2 + \|u^{(\alpha)}\|_{L^2}^2\right)\right)^{1/2}$. So we have completed the whole proof for our theorem. □

3.4.2 Nonnegative real cases

In this subsection, we will first review some materials in the theory of complex interpolation. Most of these materials come from [R-S2][T], and then prove the theorem for the case of real numbers. Roughly speaking, interpolation theory is such a mathematical scheme: assume that we have a vector space E with two norms $\|\cdot\|^{(0)}$ and $\|\cdot\|^{(1)}$ obeying a consistency condition, then the complex interpolation theory enables us to define a natural family of Banach spaces $\{E_\theta | 0 \leq \theta \leq 1\}$ which interpolate between E_0 and E_1, the completion of E in $\|\cdot\|^{(0)}$ and $\|\cdot\|^{(1)}$; the abstract interpolation theorem then follows easily; namely if $\{E_\theta\}$ interpolates between E_0 and E_1 and $\{F_\theta\}$ interpolates between F_0 and F_1, then any map T which is in $\mathcal{L}(E_0, F_0)$ and in $\mathcal{L}(E_1, F_1)$ extends uniquely to a bounded map of E_θ into F_θ for each θ. Below we will come to more details.

Definition 3.4.3. *Let E be a complex vector space. Two norms $\|\cdot\|^{(0)}$ and $\|\cdot\|^{(1)}$ on E are called* ***consistent*** *if any sequence $\{x_n\}$ that converges to zero in one norm and which is Cauchy in the other norm converges to zero in both norms. If $\|\cdot\|^{(0)}$ and $\|\cdot\|^{(1)}$ are consistent, we define*

$$\|x\|_+ = \inf\{\|y\|^{(0)} + \|z\|^{(1)} \,|x = y + z\}. \tag{3.24}$$

In this way, we can get that $\|\cdot\|_+$ is also a norm on the space E. If E_0, E_1, and E_+ denote the completion of E under $\|\cdot\|^{(0)}$, $\|\cdot\|^{(1)}$ and $\|\cdot\|_+$, then we have continuous imbeddings from E_0 into E_+ and from E_1 into E_+. If Ω is the vertical strip in the complex plane $\Omega = \{z \in \mathbb{C}|0 \leq \operatorname{Re} z \leq 1\}$, and Ω° the interior of Ω, we define $\Gamma(E)$ to be the set of continuous functions f from Ω to E_+ which are analytic in Ω° and satisfy:

(i) if $\operatorname{Re} z = 0$, then $f(z) \in E_0$ and $\theta \to f(i\theta)$ is continuous in $\|\cdot\|^{(0)}$; if $\operatorname{Re} z = 1$, then $f(z) \in E_1$ and $\theta \to f(1 + i\theta)$ is continuous in $\|\cdot\|^{(1)}$;

(ii) $\sup\limits_{z\in\Omega} \|f(z)\|_+ < +\infty$;

(iii) $|||f||| \equiv \sup\limits_{\theta\in\mathbb{R}}\{\|f(i\theta)\|^{(0)}, \|f(1 + i\theta)\|^{(1)}\} < +\infty$.

It turns out that $\Gamma(E)$ with the norm $|||\cdot|||$ is a Banach space and its subspace $K_\theta = \{f \in \Gamma(E)|f(\theta) = 0\}$ $(0 \leq \theta \leq 1)$ is $|||\cdot|||$-closed. So we can define such a quotient space

$$\widetilde{E_\theta} = \Gamma(E)/K_\theta, \quad 0 \leq \theta \leq 1, \tag{3.25}$$

and denote the quotient norm on $\widetilde{E_\theta}$ by $\|\cdot\|^{(\theta)}$. Note that E can be regarded as a subset of $\widetilde{E_\theta}$ under the map $x \longrightarrow [x]$, the constant function with value as x; and $\widetilde{E_\theta}$ can be regarded as a subset of E_+ under the map $[f] \longrightarrow f(\theta)$. We now define E_θ

to be the completion of E in the norm $\|\cdot\|^{(\theta)}$. Thus, the spaces we have defined are related as follows:

$$E \to E_\theta \to \widetilde{E_\theta} \to E_+, \tag{3.26}$$

where each mapping in the chain above is continuous injective map. In special, for $\theta = 0$ (respectively, $\theta = 1$), E_θ is just the space E_0 (respectively, E_1). The spaces E_θ are called **interpolation spaces** between E_0 and E_1 and the norms $\|\cdot\|^{(\theta)}$ are called **interpolation norms** between $\|\cdot\|^{(0)}$ and $\|\cdot\|^{(1)}$. We remark that it is possible to prove that $E_\theta = \{f(\theta) \mid f \in \Gamma(E)\}$.

Theorem 3.4.4. *(Calderón-Lions interpolation theorem) Let E and F be complex vector spaces with given consistent norms $\|\cdot\|_E^{(0)}$ and $\|\cdot\|_E^{(1)}$ on X and $\|\cdot\|_F^{(0)}$ and $\|\cdot\|_F^{(1)}$. Suppose that $T(\cdot)$ is an analytic, uniformly bounded, continuous, $\mathcal{L}(E_+, F_+)$-valued function on the strip Ω with the following properties:*

(i) $T(\theta) : E \to F$ for each $\theta \in (0,1)$;

(ii) For all $y \in \mathbb{R}$, $T(iy) \in \mathcal{L}(E_0, F_0)$ and

$$M_0 = \sup_{y\in\mathbb{R}} \|T(iy)\|_{\mathcal{L}(E_0,F_0)} < +\infty;$$

(iii) For all $y \in \mathbb{R}$, $T(1+iy) \in \mathcal{L}(E_1, F_1)$ and

$$M_1 = \sup_{y\in\mathbb{R}} \|T(1+iy)\|_{\mathcal{L}(E_1,F_1)} < +\infty.$$

Then for any $\theta \in (0,1)$,

$$T(\theta)[E_\theta] \subset F_\theta$$

and

$$\|T(\theta)\|_{\mathcal{L}(E_\theta,F_\theta)} \le M_0^{1-\theta} M_1^\theta.$$

One example that fits in this theoretic scheme is Sobolev spaces. If we have $E_0 = H^s(\mathbb{R})$ with norm $\|\cdot\|_E^{(0)} = \|\cdot\|_{H^s}$ and $E_1 = H^t(\mathbb{R})$ with norm $\|\cdot\|_E^{(1)} = \|\cdot\|_{H^t}$, then we have $E_\theta = H^{(1-\theta)s+\theta t}(\mathbb{R})$ (see [T] vol. I p275-278). As to the case of our functional spaces, they also fit in this scheme.

Proposition 3.4.5. *Let $E = S(\mathbb{R})$ and let $\|\cdot\|^{(0)} = \|\cdot\|_{X^s}$ and norm $\|\cdot\|^{(1)} = \|\cdot\|_{X^t}$ where s and t are any two real numbers. Then we have for any $0 \le \theta \le 1$*

$$E_\theta = X^{(1-\theta)s+\theta t}(\mathbb{R}), \tag{3.27}$$

and the interpolation norms $\|\cdot\|^{(\theta)}$ are just $\|\cdot\|_{X^{(1-\theta)s+\theta t}}$.

Proof. It is easy to verify that the norm $\|\cdot\|^{(0)}$ and $\|\cdot\|^{(1)}$ are consistent and the completion of the space E under these two norms are $X^s(\mathbb{R})$ and $X^t(\mathbb{R})$ respectively. To complete the proof, we only need to show that the norm $\|\cdot\|^{(\theta)}$ is just the norm $\|\cdot\|_{X^{(1-\theta)s+\theta t}}$ on a dense subset of the space E. Let $u(x)$ be a rapid decreasing function with only finitely many nonzero Hermite coefficients, say $u(x) = \sum_{j=0}^N q_{k_j} h_{k_j}(x)$. We define

$$f(z) = \sum_{j=0}^N q_{k_j}(k_j + 1/2)^{\frac{\theta-z}{2}(t-s)} h_{k_j}(x). \tag{3.28}$$

Then this mapping satisfies $f \in \Gamma(E)$ and $f(\theta) = u(x)$. We have $f(iy) = \sum_{j=0}^N q_{k_j}(k_j + 1/2)^{\frac{\theta-iy}{2}(t-s)} h_{k_j}(x)$ and its norm

$$\begin{aligned}\|f(iy)\|_{X^s}^2 &= \sum_{j=0}^N \left|q_{k_j}\right|^2 \left|(k_j + 1/2)^{(\theta-iy)(t-s)}\right| (k_j + 1/2)^s \\ &= \sum_{j=0}^N \left|q_{k_j}\right|^2 (k_j + 1/2)^{(1-\theta)s+\theta t}(k_j + 1/2)^s \\ &= \|u(x)\|_{X^{(1-\theta)s+\theta t}}^2 .\end{aligned}$$

Similarly, $f(1+iy) = \sum_{j=0}^{N} q_{k_j}\ (k_j+1/2)^{\frac{\theta-1-iy}{2}(t-s)}\ h_{k_j}(x)$ and its norm

$$\begin{aligned}\|f(1+iy)\|_{X^t}^2 &= \sum_{j=0}^{N} \left|q_{k_j}\right|^2 \left|(k_j+1/2)^{(\theta-1-iy)(t-s)}\right| (k_j+1/2)^t \\ &= \sum_{j=0}^{N} \left|q_{k_j}\right|^2 (k_j+1/2)^{(1-\theta)s+\theta t}(k_j+1/2)^t \\ &= \|u(x)\|_{X^{(1-\theta)s+\theta t}}^2 .\end{aligned}$$

So we have $|||f||| = \|u(x)\|_{X^{(1-\theta)s+\theta t}}$ and further $\|u\|^{(\theta)} = \|f(\theta)\|^{(\theta)} = \|f\|_{\Gamma(E)/K_\theta} \leq \|u(x)\|_{X^{(1-\theta)s+\theta t}}$.

For the converse, let us assume that $f \in \Gamma(E)$ such that $f(\theta) = u(x)$, then we want to prove that $\|f\|_{\Gamma(E)/K_\theta} \geq \|u(x)\|_{X^{(1-\theta)s+\theta t}}$. Let function v is a rapid decreasing function with only finitely many nonzero Hermite coefficients, say $v(x) = \sum_{j=0}^{N} p_{k_j}(x)$. We define

$$g(z) = \sum_{j=0}^{N} p_{k_j}(k_j+1/2)^{(\theta-z)(-t+s)} h_{k_j}(x). \tag{3.29}$$

Then this mapping is an analytical mapping from Ω to the space $X^{-s} + X^{-t}$ which satisfies $g(\theta) = v(x)$. From the computation made for $f(iy)$ and $f(1+iy)$, we know that $|||g||| = \|v(x)\|_{X^{-(1-\theta)s-\theta t}}$. Let us consider the integration $H(z) = \int_{-\infty}^{+\infty} f(z)$ $g(z)dx$, which is an analytical mapping from Ω to $\mathbb{C}$ satisfying

$$|H(iy)| = \left|\int_{-\infty}^{+\infty} f(iy)g(iy)dx\right| \leq |||f|||\,|||g|||, \tag{3.30}$$

$$|H(1+iy)| = \left|\int_{-\infty}^{+\infty} f(1+iy)g(1+iy)dx\right| \leq |||f|||\,|||g|||. \tag{3.31}$$

From the Hadarmard's three line theorem. we can get $|H(z)| \leq |||f|||\,|||g|||$, and in particular

$$|H(\theta)| = \left|\int_{-\infty}^{+\infty} u(x)v(x)dx\right| \leq |||f|||\,|||g||| = |||f|||\,\|v(x)\|_{X^{-(1-\theta)s-\theta t}}. \tag{3.32}$$

Note that the functions with finitely many nonzero Hermite coefficients are dense in the function space $X^{-(1-\theta)s-\theta t}$, which is the dual space of $X^{(1-\theta)s+\theta t}$. So we can conclude that $|||f||| \geq \|u(x)\|_{X^{(1-\theta)s+\theta t}}$ is true for any $f \in \Gamma(E)$ and further

$$\|u\|^{(\theta)} \geq \|u(x)\|_{X^{(1-\theta)s+\theta t}} . \tag{3.33}$$

In conclusion, we get $E_\theta = X^{(1-\theta)s+\theta t}(\mathbb{R})$ and $\|\cdot\|^{(\theta)} = \|\cdot\|_{X^{(1-\theta)s+\theta t}}$. □

By far we have got two families of function spaces H^s and X^s, which both fit well in the scheme of the complex interpolation theory. Then a simple application of the Calderón-Lions interpolation theorem can provide us the results of the main theorem in this section in case of nonnegative real numbers.

Let $E = F = S(\mathbb{R})$, and $\|\cdot\|_E^{(0)}$, $\|\cdot\|_E^{(1)}$, $\|\cdot\|_F^{(0)}$ and $\|\cdot\|_F^{(1)}$ are respectively $\|\cdot\|_{X^0}$, $\|\cdot\|_{X^n}$, $\|\cdot\|$ and $\|\cdot\|_{H^n}$, where n is an integer greater than the fixed nonnegative real number s. Let $T(\cdot)$ be the mapping from E_+ to F_+ continuously extended from the identity mapping. on $S(\mathbb{R})$. Note that we have already known that the mappings $T : X^0 \to H^0$ and $T : X^n \to H^n$ are both continuous, then the interpolation theorem tells us that T continuously maps X^s into H^s for any real numbers $0 \leq s \leq n$. Since X^s is invariant under the Fourier transformation, it can be concluded that X^s are subsets of $H^s \cap \mathcal{F}(H^s)$ and the natural imbeddings $X^s \hookrightarrow H^s \cap \mathcal{F}(H^s)$ are continuous. Obviously, this result is in fact true for any nonnegative real number s due to the arbitrariness of the choice of integer n.

3.4.3 Negative real cases

In this subsection, we will analyze the case of negative numbers. And as the end of the whole sections, we will provide our conjecture for the characterization of the spaces X^s, which is much more elegant and also requires more careful analysis.

For any real number $-s < 0$, we have that $X^{-s}(\mathbb{R})$ is the dual space of $X^s(\mathbb{R})$ which satisfies $X^s(\mathbb{R}) \subset H^s(\mathbb{R}) \cap \mathcal{F}(H^s(\mathbb{R}))$ and $i : X^s(\mathbb{R}) \to H^s(\mathbb{R})$ (respectively $\mathcal{F}(H^s(\mathbb{R}))$) is continuous. Then its adjoint operator $i^t : H^{-s}(\mathbb{R}) \to X^{-s}(\mathbb{R})$ (respectively,$i^t : \mathcal{F}(H^{-s}(\mathbb{R})) \to X^{-s}(\mathbb{R})$) is also continuous. These mappings are both injective because $S(\mathbb{R})$ and thus $X^s(\mathbb{R})$ are dense subsets of $H^s(\mathbb{R})$ (respectively $\mathcal{F}(H^s(\mathbb{R}))$). So it is easy to conclude that the space $H^{-s}(\mathbb{R})$ can be continuously embedded in the space $X^{-s}(\mathbb{R})$. A little more direct computation can show that the space $H^{-s}(\mathbb{R})$ is really a subspace of the latter one.

Let $u \in H^{-s}(\mathbb{R})$ and $v \in X^s(\mathbb{R}) \subset H^s(\mathbb{R})$ ($s > 0$). We can have

$$\begin{aligned} |\langle u, v\rangle| &= \left|\int_{\mathbb{R}} \mathcal{F}u(\xi)\mathcal{F}^3 v(\xi) d\xi\right| \\ &\leq \|u\|_{H^{-s}} \|v\|_{H^s} \leq \|u\|_{H^{-s}} \|v\|_{X^s}. \end{aligned} \tag{3.34}$$

Writing the function v in the form of Hermite sequence, then the inequality above can be rewritten as $\left|\sum_{k=0}^{+\infty} u_k v_k\right| \leq M \left(\sum_{k=0}^{+\infty} |v_k|^2 (k+1/2)^s\right)^{1/2}$, which implies that $u(x)$ is a function in $X^{-s}(\mathbb{R})$.

Since $X^{-s}(\mathbb{R})$ is invariant under the Fourier transformation, we can further conclude that $H^{-s}(\mathbb{R}) + \mathcal{F}(H^{-s}(\mathbb{R})) \subset X^{-s}(\mathbb{R})$. Noting that $S(\mathbb{R})$ is a dense subset of $X^{-s}(\mathbb{R})$ and obviously a subset of $H^{-s}(\mathbb{R}) + \mathcal{F}(H^{-s}(\mathbb{R}))$, it can be deduced that $\overline{H^{-s}(\mathbb{R}) + \mathcal{F}(H^{-s}(\mathbb{R}))} = X^{-s}(\mathbb{R})$. The space $H^{-s}(\mathbb{R}) + \mathcal{F}(H^{-s}(\mathbb{R}))$ has a natural

Banach structure with norm as

$$\|u\|_{H^{-s}(\mathbb{R})+\mathcal{F}(H^{-s}(\mathbb{R}))} = \inf_{u=u_1+u_2} \left(\|u_1\|_{H^{-s}(\mathbb{R})} + \|u_2\|_{\mathcal{F}(H^{-s}(\mathbb{R}))} \right). \tag{3.35}$$

From the fact that $\|u_1\|_{H^{-s}(\mathbb{R})} \gtrsim \|u_1\|_{X^{-s}(\mathbb{R})}$ and $\|u_2\|_{\mathcal{F}(H^{-s}(\mathbb{R}))} \gtrsim \|u_1\|_{X^{-s}(\mathbb{R})}$, we can easily get $\|u\|_{H^{-s}(\mathbb{R})+\mathcal{F}(H^{-s}(\mathbb{R}))} \gtrsim \|u\|_{X^{-s}(\mathbb{R})}$, which implies that the natural inclusion map $H^{-s}(\mathbb{R}) + \mathcal{F}\left(H^{-s}(\mathbb{R})\right) \hookrightarrow X^{-s}(\mathbb{R})$ is continuous.

We would like to pay special attention to the cases of negative integers. It is already known that $X^n(\mathbb{R}) = H^n(\mathbb{R}) \cap \mathcal{F}\left(H^n(\mathbb{R})\right)$, which is a Hilbert space under the inner product $(u, v) = \sum_{\alpha \le n} \int_{-\infty}^{+\infty} \left(\frac{d}{dx}\right)^\alpha u \, \left(\frac{dv}{dx}\right)^\alpha v + x^{2\alpha} uv \, dx$. Then for any element $l \in X^{-n}(\mathbb{R})$, which is the dual space of $X^n(\mathbb{R})$, there is an element $f \in X^n(\mathbb{R})$ such that for all element $u \in S(\mathbb{R})$

$$\begin{aligned} l(u) = (f, u) &= \sum_{\alpha \le n} \int_{-\infty}^{+\infty} \left(\frac{d}{dx}\right)^\alpha u \left(\frac{d}{dx}\right)^\alpha f + x^{2\alpha} u f \, dx \\ &= \sum_{\alpha \le n} \left\langle \left(-\frac{d}{dx}\right)^\alpha \left(\frac{d}{dx}\right)^\alpha f, u \right\rangle + \left\langle x^{2\alpha} f, u \right\rangle. \end{aligned}$$

So we have such a representation

$$X^{-n}(\mathbb{R}) = \{u \in S'(\mathbb{R}) \,|\, u = \sum_{\alpha \le n} \left(-\frac{d}{dx}\right)^\alpha f_\alpha + \sum_{\beta \le n} x^\beta f_\beta\}, \tag{3.36}$$

where f_α and f_β are all L^2 functions. Together with the fact that $H^{-s}(\mathbb{R}) + \mathcal{F}\left(H^{-s}(\mathbb{R})\right) \subset X^{-s}(\mathbb{R})$, the relation above implies $X^{-n}(\mathbb{R}) = H^{-n}(\mathbb{R}) + \mathcal{F}\left(H^{-n}(\mathbb{R})\right)$.

It is worth to point out that the mapping $i : H^{-n}(\mathbb{R}) + \mathcal{F}\left(H^{-n}(\mathbb{R})\right) \hookrightarrow X^{-n}(\mathbb{R})$ are thus a continuous bijection from Banach space $H^{-n}(\mathbb{R}) + \mathcal{F}\left(H^{-n}(\mathbb{R})\right)$ with norm $\|\cdot\|_{H^{-n}(\mathbb{R})+\mathcal{F}(H^{-n}(\mathbb{R}))}$ onto the Hilbert space $X^{-n}(\mathbb{R}) = H^{-n}(\mathbb{R}) + \mathcal{F}\left(H^{-n}(\mathbb{R})\right)$ with norm $\|\cdot\|_{X^{-n}(\mathbb{R})}$. Then inverse mapping theorem (a special case of open mapping theorem) tell us that $i^{-1} : X^{-n}(\mathbb{R}) \to H^{-n}(\mathbb{R}) + \mathcal{F}\left(H^{-n}(\mathbb{R})\right)$ is also continuous.

Thus we know that these two space are in fact same Banach spaces, and the only difference between them is they have different but mutually equivalent norms.

There is no essential difficulties to generalizing the results above into higher dimensions. When we try to generalize the results in integer cases into the cases of arbitrary real numbers, we have met some difficulties, but we do believe they are in fact also true. Here we write it into the form of a conjecture below to finish this section.

Conjecture 3.4.6. *For any real number $s \geq 0$, we have $X^s(\mathbb{R}^n) = H^s(\mathbb{R}^n) \cap \mathcal{F}(H^s(\mathbb{R}^n))$, and the norm $\|\cdot\|_{X^s}$ is equivalent to the norm $\|u\|_s = \left(\|u\|_{H^s}^2 + \|\langle x\rangle^s u(x)\|_{L^2}^2\right)^{1/2}$. For any real number $-s \leq 0$, we have $H^{-s}(\mathbb{R}^n) + \mathcal{F}(H^{-s}(\mathbb{R}^n)) = X^{-s}(\mathbb{R}^n)$, and the norm $\|\cdot\|_{X^{-s}}$ is equivalent to the norm $\|u\|_{-s} = \inf_{u_1+u_2=u}\left(\|u_1\|_{H^{-s}} + \left\|\langle x\rangle^{-s} u_2(x)\right\|_{L^2}^2\right)$.*

3.5 Other properties

In this section, we will describe some other properties of the function spaces X^s. As we have found, our function spaces X^s have very close relationships with Sobolev function spaces, which is extensively used in the theory of partial differential equations. Since when $s \geq 0$ our function spaces X^s are in fact a subset of the Sobolev spaces H^s, many properties of the function in Sobolev spaces are also true for the function in our spaces. For example, we have Sobolev imbedding theorems:

- If $0 \leq s < n/2$ and $1/p = 1/2 - s/n$. then the inclusion mapping $X^s(\mathbb{R}^n) \hookrightarrow L^p(\mathbb{R}^n)$ is continuous.

- If $s - n/2 = m + \alpha$ where $m \in \mathbb{Z}_+$ and $0 < \alpha < 1$, then all the functions in

$X^s(\mathbb{R}^n)$ have continuous derivatives up to m times, and all of its m times derivatives are Hölder continuous of order α.

Meanwhile our spaces have compact imbedding theorem (proposition 3.3.2), which doesn't require the assumption of "compactly supported" as the Sobolev spaces do.

However, our function spaces also possess some properties different from Sobolev spaces. One of them is its relationship with L^p space.

3.5.1 Relationship with L^p spaces

At first let us analyze the functions in the space $X^s(\mathbb{R}^n)$ with $s > n/2$.

Proposition 3.5.1. *If $s > n/2$, then any function $f = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^s(\mathbb{R}^n)$ is a bounded continuous function, which is tending to zero at infinity and integrable on the whole plane. And the approximation sequence $f_N = \sum_{k=0}^{N} q_k h_k(x)$ is uniformly convergent to f on any bounded sets of $\mathbb{R}^n$.*

Proof. At first, let us see the case of 1-dimension. Since $X^s(\mathbb{R})$ $(s > 1/2)$ is a subset of $H^s(\mathbb{R})$, we can get that $\mathcal{F}f(\xi)$ is an integrable function, which implies that $f(x)$ is a bounded continuous function, which is tending to zero at infinity. Since $X^s(\mathbb{R})$ is a subset of $\mathcal{F}\left(H^s(\mathbb{R})\right)$, we can get

$$\begin{aligned} \int_{-\infty}^{+\infty} |f(x)|\, dx &\leq \left(\int_{-\infty}^{+\infty} |f(x)|^2 \langle x\rangle^{2s}\, dx\right)^{1/2} \left(\int_{-\infty}^{+\infty} \langle x\rangle^{-2s}\, dx\right)^{1/2} \\ &\lesssim \|f\|_{X^s} < +\infty \quad (s), \end{aligned} \tag{3.37}$$

which implies that f is integrable.

As to the uniformly convergence of the sequence, we have for any x in a bounded sets, say $\{|x| \leq M\}$,

$$|f(x) - f_N(x)| = \left|\sum_{k>N}^{+\infty} q_k h_k(x)\right|$$
$$\leq \left(\sum_{k>N}^{+\infty} |q_k|^2 (k+1/2)^s\right)^{1/2} \left(\sum_{k>N}^{+\infty} (k+1/2)^{-s} |h_k(x)|^2\right)^{1/2}. \quad (3.38)$$

We would like to mention that there is such asymptotic representations of the Hermite functions [M-O-S]

$$H_{2n}(x) = (-1)^n 2^n (2n-1)!! e^{x^2/2} \left(\cos\left(\sqrt{4n+1}x\right) + O\left(\frac{1}{\sqrt[4]{n}}\right)\right), \quad (3.39)$$

$$H_{2n+1}(x) = (-1)^n 2^{n+1/2} (2n-1)!! \sqrt{2n+1} e^{x^2/2} \left(\sin\left(\sqrt{4n+3}x\right) + O\left(\frac{1}{\sqrt[4]{n}}\right)\right), \quad (3.40)$$

which imply that on any bounded sets $\{|x| \leq M\}$

$$|h_k(x)| \lesssim \langle k \rangle^{-1/4} \quad (M), \quad (3.41)$$

where the notation (M) means that the constant corresponding to $h_k(x) \lesssim \langle k \rangle^{-1/4}$ depends on M. So the inequality (3.38) can be written as

$$|f(x) - f_N(x)| \lesssim \left(\sum_{k>N}^{+\infty} |q_k|^2 (k+1/2)^s\right)^{1/2} \left(\sum_{k>N}^{+\infty} (k+1/2)^{-s} \langle k \rangle^{-1/2}\right)^{1/2} \quad (M), \quad (3.42)$$

which implies that $f_N = \sum_{k=0}^{N} q_k h_k(x)$ is uniformly convergent to f on the sets $\{|x| \leq M\}$. The proof for higher dimensional case can be obtained similarly. □

Corollary 3.5.2. *For $1 \leq p \leq +\infty$, we have the continuous inclusion $L^p(\mathbb{R}^n) \hookrightarrow X^{(-\frac{n}{2})^-}(\mathbb{R}^n)$, where $X^{(-\frac{n}{2})^-}(\mathbb{R}^n)$ means any space $X^s(\mathbb{R}^n)$ with $s < -\frac{n}{2}$.*

Proof. It suffices to prove the case of $L^1(\mathbb{R}^n)$ and $L^{+\infty}(\mathbb{R}^n)$, since all the other cases can be deduced from the interpolation method. Let f is a function in $L^1(\mathbb{R}^n)$.

Then it can induce a linear mapping T_f defined by $T_f(\varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)dx$ for $\varphi(x) \in X^{\left(\frac{n}{2}\right)^+}$. This operator is well defined and satisfies

$$|T_f(\varphi)| \leq \|f\|_{L^1} \|\varphi\|_{L^{+\infty}} \leq \|f\|_{L^1} \|\varphi\|_{X^{\left(\frac{n}{2}\right)^+}}, \tag{3.43}$$

which means that T_f is an element in $X^{\left(-\frac{n}{2}\right)^-}(\mathbb{R}^n)$, or, for any $s < -\frac{n}{2}$ it is true that $T_f \in X^s(\mathbb{R}^n)$. Thus we get a continuous mapping $T : L^1(\mathbb{R}^n) \to X^s(\mathbb{R}^n)$ by sending f to the element T_f. This mapping T is injective due to the fact that $\int_{-\infty}^{+\infty} f(x)\varphi(x)dx = 0$ for all rapid decreasing functions $\varphi(x)$ implies $f \equiv 0$. So it comes to the conclusion that $L^1(\mathbb{R}^n) \hookrightarrow X^{\left(-\frac{n}{2}\right)^-}(\mathbb{R}^n)$ is a continuous inclusion.

For the case of $L^{+\infty}(\mathbb{R}^n)$ let us consider the mapping $T : L^{+\infty}(\mathbb{R}^n) \to X^s(\mathbb{R}^n)$ which defined by $T_f(\varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)dx$. It is a well defined continuous mapping due to the inequality

$$|T_f(\varphi)| \leq \|f\|_{L^{+\infty}} \|\varphi\|_{L^1} \leq \|f\|_{L^{+\infty}} \|\varphi\|_{X^{\left(\frac{n}{2}\right)^+}}. \tag{3.44}$$

And we can furthermore use similar reasoning to prove the desired results. □

The results of the proposition 3.5.1 and the corollary 3.5.2 can be written together as

$$\forall 1 \leq p \leq +\infty,\ X^{\left(-\frac{n}{2}\right)^+}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \subset X^{\left(\frac{n}{2}\right)^-}(\mathbb{R}^n), \tag{3.45}$$

where each relation "$\subset$" can induce a continuous injection form the former space to latter space. By the imbedding theorem, we can further have such two finer relationships

$$\forall 2 \leq p < +\infty,\ X^{s_{p_+}}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \text{ where } s_{p_+} = n(1/2 - 1/p), \tag{3.46}$$

$$\forall 1 < p \leq 2,\ L^p(\mathbb{R}^n) \subset X^{s_{p_-}}(\mathbb{R}^n) \text{ where } s_{p_-} = n(1/2 - 1/p). \tag{3.47}$$

These facts suggest us to ask such a question:

Problem 3.5.3. *For $1 \leq p \leq +\infty$, what is the best indexes a_p (or $a_p^{\pm}$) and b_p (or $b_p^{\pm}$) so that $X^{a_p}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n) \subset X^{b_p}(\mathbb{R}^n)$, which satisfies that all the related natural inclusions are continuous?*

This question has close relation with the problem of mean convergence of expansions in Hermite series. In 1965, Richard Askey and Stephen Wainger once obtained such a result [A-W].

Theorem 3.5.4. *Let f be in $L^p(-\infty, \infty)$, $4/3 < p < 4$. Define $a_n = \int_0^{+\infty} f(x)h_n(x)dx$ and set $S_n = \sum_{k=0}^{n} a_k h_k(x)$. Then $\|S_n - f\|_{L^p}$ tends to 0 as n approaches infinity. And there is an A_p such that $\|S_n\|_{L^p} \leq A_p \|f\|_{L^p}$ $(n = 0, 1, 2, \cdots)$, where A_p is independent of $f \in L^p$.*

For the case of $1 \leq p \leq 4/3$ and $p \geq 4$, they pointed out that the theorem is false. Now the answers to our question (if available) provide us such inequalities:

$$\|S_n - f\|_{X^{b_p}} \lesssim \|S_n - f\|_{L^p} \lesssim \|S_n - f\|_{X^{a_p}}. \tag{3.48}$$

So if the function f is in the space $X^{a_p}(\mathbb{R}^n)$, then its Hermite expansion converges to f in L^p space; in other cases, it has lower bound of error estimate as $\|S_n - f\|_{X^{b_p}}$.

3.5.2 Product of two X^s functions

In this subsection, we will study the following problem: suppose we have two tempered distribution $u_j \in X^{s_j}(\mathbb{R}^n)$, $j = 1, 2$, then how to define their product u_1u_2 and what is the property of this product. Formally we can do in this way: let $\langle , \rangle$ denote "pair operation" between two elements coming from a locally convex topological space and its dual space, then we define

$$\langle u_1 u_2, \varphi \rangle = \langle u_1, u_2 \varphi \rangle \quad \text{for any } \varphi \in S(\mathbb{R}^n), \tag{3.49}$$

where we need $u_2\varphi \in X^{-s_1}(\mathbb{R}^n)$. Obviously, this definition is compatible with the usual product of two rapid decreasing functions.

The last requirement is not superfluous. In general there is little hope to define the product of two arbitrary tempered distributions. For example one can hardly find a reasonable way to define the square of delta function $\delta(x)$ as a tempered distribution. So restrictions must be posed on the choice of u_1 and u_2. It is well known that if we choose u_2 from the space O_M^n, the set of infinitely differentiable functions on $\mathbb{R}^n$ which together with their derivatives are polynomially bounded, then $u_2\varphi \in S(\mathbb{R}^n) \subset X^{-s_1}(\mathbb{R}^n)$. Here we adopt another restriction, namely that $s_1 + s_2 \geq 0$.

Proposition 3.5.5. *If $u_j \in X^{s_j}(\mathbb{R}^n)$, $j = 1, 2$ and $s_1 + s_2 \geq 0$, then the equality (3.49) provides a well-defined tempered distribution u_1u_2, which is independent of the order of the two elements ($u_1u_2 = u_2u_1$) and satisfies*

$$|\langle u_1 u_2, \varphi \rangle| \lesssim C(\varphi) \, \|u_1\|_{X^{s_1}} \, \|u_2\|_{X^{s_2}} \quad (s_1, s_2), \tag{3.50}$$

where $C(\varphi) \longrightarrow 0$ as $\varphi \longrightarrow 0$ in $S(\mathbb{R}^n)$.

To prove proposition 3.5.5, the essential task is to verify $u_2\varphi \in X^{-s_1}(\mathbb{R}^n)$ and inequality (3.50). This is because we can then deduce that inequality (3.50) defines a continuous functional on $S(\mathbb{R}^n)$ and thus a tempered distribution; meanwhile u_1u_2 can be approximated by the product $\left(\sum_{k=0}^{n} a_k h_k(x)\right) \left(\sum_{l=0}^{n} b_l h_l(x)\right)$, which is of course independent of the order of their product. So we only need prove such a lemma:

Lemma 3.5.6. *For any $\varphi(x) \in S(\mathbb{R}^n)$, we have the mapping $T_\varphi : u \longmapsto \varphi u$ is a morphism of order* 0 *on the Hilbert scale $\{X^s(\mathbb{R}^n)\}$, which satisfies*

$$\|\varphi u\|_{X^s} \lesssim C(\varphi)\|u\|_{X^s} \quad (s). \tag{3.51}$$

Proof. For simplicity of writing we prove it in one dimensional case. The proof for higher dimensional cases are essentially same.

When $s = 0$, we have

$$\|\varphi u\|_{L^2} \leq \|\varphi\|_{L^\infty}\|u\|_{L^2}. \tag{3.52}$$

when $s = n \in \mathbb{N}$, we have $\|\varphi u\|_{X^n} \approx \sum_{0\leq\alpha+\beta\leq n}\left\|x^\alpha(\varphi u)^{(\beta)}\right\|_{L^2}$. By Leibniz's rule, it can be controlled by $\|\varphi u\|_{X^n} \lesssim \sum_{0\leq\alpha+\beta_1+\beta_2\leq n}\left\|\varphi^{(\beta_1)}x^\alpha(u)^{(\beta_2)}\right\|_{L^2}$. Thus we can further have

$$\|\varphi u\|_{X^n} \leq \sup_{0\leq\beta\leq n}\left\|\varphi^{(\beta)}\right\|_{L^\infty}\|u\|_{X^n}. \tag{3.53}$$

when $s = -n$ is a negative integer, we have $\|\varphi u\|_{X^{-n}} \leq \inf_{u_1+u_2=u}\left(\|\varphi u_1\|_{H^{-n}} + \|\varphi u_2\|_{\mathcal{F}(H^{-n})}\right)$, in which each term can be controlled:

$$\|\varphi u_1\|_{H^{-n}} \leq C(\varphi)\|u_1\|_{H^{-n}} \text{ (see theorem 3.5.13)}$$

and

$$\|\varphi u_2\|^2_{\mathcal{F}(H^{-n})} = \int_{-\infty}^{+\infty}\langle x\rangle^{-2n}|\varphi u_2|^2\,dx \leq \|\varphi\|^2_{L^\infty}\|u_2\|^2_{\mathcal{F}(H^{-n})}.$$

Thus we can have

$$\|\varphi u\|_{X^{-n}} \leq C(\varphi)\|u\|_{X^{-n}}.$$

Finally by a simple application of Calderón-Lions interpolation theorem, we can deduce the inequality (3.51). □

Let us recall that for any tempered distribution, there is a unique sequence of Hermite coefficients corresponding to that distribution. We will show next how to compute these coefficients in one dimensional case. Assume that $u = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^{s_1}(\mathbb{R}^1)$, $v = \sum_{l=0}^{+\infty} p_l h_l(x) \in X^{s_2}(\mathbb{R}^1)$ and $s_1+s_2 \geq 0$. Let $c_{klm} = \int_{-\infty}^{+\infty} h_k(x)h_l(x)h_m(x)dx$. Then formally we have

$$(uv)_m = \sum_{k,l=0}^{+\infty} q_k p_l c_{klm}. \tag{3.54}$$

According to proposition 3.5.5, we know that $(uv)_m = \langle uv, h_m(x)\rangle$ is a well defined complex number. But unfortunately, the formula (3.54) is not absolutely convergent in general. So we need make clear the precise meaning of the formula (3.54).

Proposition 3.5.7. *Let u, v, c_{klm} as above. Let $u_{N_1} = \sum_{k=0}^{N_1} q_k h_k(x)$ and $v_{N_2} = \sum_{l=0}^{N_2} p_l h_l(x)$. Then*

$$(uv)_m = \lim_{N_1, N_2 \longrightarrow +\infty} \sum_{k=0}^{N_1}\sum_{l=0}^{N_2} q_k p_l c_{klm}. \tag{3.55}$$

Proof. Noticing that $\sum_{k=0}^{N_1}\sum_{l=0}^{N_2} q_k p_l c_{klm} = (u_{N_1} v_{N_2})_m$, we only need estimate the difference between $(uv)_m$ and $(u_{N_1} v_{N_2})_m$. By the inequality (3.51), we get

$$\begin{aligned} |(uv)_m - (u_{N_1} v_{N_2})_m| &\leq |((u-u_{N_1})\, v)_m| + |(u_{N_1}(v - v_{N_2}))_m| \\ &\lesssim C(h_m(x)) \left\|(u-u_{N_1})\right\|_{X^{s_1}} \left\|u_2\right\|_{X^{s_2}} + \\ &C(h_m(x)) \left\|(u_{N_1})\right\|_{X^{s_1}} \left\|u_2\right\|_{X^{s_2}} \quad (s_1, s_2). \end{aligned}$$

Take the limit $\lim_{N_1,N_2 \longrightarrow +\infty}$ on the right side, then we get the result. □

There are many cases such that the formula (3.54) are in fact absolutely convergent. Here is an example.

Proposition 3.5.8. *Assume that* $u = \sum_{k=0}^{+\infty} q_k h_k(x) \in L^2(\mathbb{R}^1)$*, and* $v = \sum_{l=0}^{+\infty} p_l h_l(x) \in X^s(\mathbb{R}^1)$ *where* $s > 1/2$*. Let* $c_{klm} = \int_{-\infty}^{+\infty} h_k(x) h_l(x) h_m(x) dx$*. Then*

$$(uv)_m = \sum_{k,l=0}^{+\infty} q_k p_l c_{klm}.$$

is absolutely convergent and $uv \in L^2(\mathbb{R}^n)$*.*

Proof. We have

$$\begin{aligned} \sum_{k,l=0}^{+\infty} |q_k p_l c_{klm}| &\leq \sum_{l=0}^{+\infty} \left(\sum_{k=0}^{+\infty} |c_{klm}|^2\right)^{1/2} \left(\sum_{k=0}^{+\infty} |q_k|^2\right)^{1/2} |p_l| \\ &= \left(\sum_{l=0}^{+\infty} |p_l| \, C_{mmll}^{1/2}\right) \|u\|_{L^2}. \end{aligned} \tag{3.56}$$

By the result of the theorem of the last chapter, it is true that $C_{mmll}^{1/2} \lesssim (\max(m,l))^{(-1/2)^-}$. Then the inequality (3.56) can be continued as

$$\begin{aligned} \sum_{k,l=0}^{+\infty} |q_k p_l c_{klm}| &\lesssim \left(\sum_{l=0}^{+\infty} |p_l|^2 \langle l\rangle^s\right)^{1/2} \left(\sum_{l=0}^{+\infty} \langle l\rangle^{-s} \langle l\rangle^{(-1/2)^-}\right)^{1/2} \|u\|_{L^2} \\ &\leq C(s) \|u\|_{L^2} \|v\|_{X^s} \qquad \text{since } s > 1/2. \end{aligned}$$

The statement of $uv \in L^2(\mathbb{R}^n)$ can be deduced like this: $v \in X^s(\mathbb{R}^n)$ where $s > 1/2$ implies $\|v\|_{L^\infty} \lesssim \|v\|_{X^s}$; and the multiplication of a bounded function and an L^2 function must be an L^2 function. □

By using the Fourier Transformation, the definition formula (3.49) can also be written in the form of

$$\langle u_1 u_2, \varphi\rangle = \left\langle \mathcal{F}(u_1 u_2), \mathcal{F}^3(\varphi)\right\rangle = (2\pi)^{-n/2} \left\langle \mathcal{F}(u_1) * \mathcal{F}(u_2), \mathcal{F}^3(\varphi)\right\rangle, \tag{3.57}$$

where $\varphi \in S(\mathbb{R}^n)$ and

$$\mathcal{F}(u_1) * \mathcal{F}(u_2)(\xi) = \langle \mathcal{F}(u_1)(\xi - \cdot), \mathcal{F}(u_2)(\cdot)\rangle. \tag{3.58}$$

It turns out that the function $\mathcal{F}(u_1 u_2)(\xi)$ can be described well. In particular, let us see the case of $u_1 \in X^s(\mathbb{R}^n)$ and $u_2 \in X^{-s}(\mathbb{R}^n)$ where s is nonnegative. The following lemma is crucial.

Lemma 3.5.9. *Let $\tau_a : S'(\mathbb{R}^n) \longrightarrow S'(\mathbb{R}^n)$ is the linear operator mapping $f(\cdot) \mapsto f(a-\cdot)$ for $a \in \mathbb{R}^n$. Then it is a morphism of order 0 on the Hilbert scale $\{X^s(\mathbb{R}^n)\}$ which satisfies*

$$\|\tau_a u\|_{X^s} \lesssim \|u\|_{X^s} \langle a \rangle^{|s|}. \tag{3.59}$$

Proof. For simplicity of writing we prove it in one dimensional case. The proof for higher dimensional cases are essentially same. When $s = 0$, it is easy to see that $\|\tau_a u\|_{L^2} = \|u\|_{L^2}$. When $s = m \in \mathbb{N}$, we have $\|\tau_a u\|_{X^m} \approx \|\tau_a u\|_{H^m} + \|\tau_a u\|_{\mathcal{F}H^m}$. Since $\|\tau_a u\|_{H^m} = \|u\|_{H^m}$ and

$$\begin{aligned} \|\tau_a u\|_{\mathcal{F}H^m} &= \left(\int_{-\infty}^{+\infty} |u(x)|^2 \langle a - x \rangle^{2m} dx\right)^{1/2} \\ &\lesssim \left(\int_{-\infty}^{+\infty} |u(x)|^2 \langle x \rangle^{2m} dx\right)^{1/2} \langle a \rangle^m, \end{aligned}$$

we can get that $\|\tau_a u\|_{X^m} \lesssim \|u\|_{X^m} \langle a \rangle^m$. Then a simple application of Calderón-Lions interpolation theorem provide us that

$$\|\tau_a u\|_{X^s} \lesssim \|u\|_{X^s} \langle a \rangle^s$$

is true for all $s \geq 0$.

For the negative case $-s < 0$, we only need notice that τ_a is symmetric, which can be deduced from the fact that $\langle \tau_a \varphi, \psi \rangle = \langle \varphi, \tau_a \psi \rangle$ is true for all $\varphi, \psi \in S(\mathbb{R}^n)$. In particular, we write for any functions $u \in X^s(\mathbb{R}^1), v \in X^{-s}(\mathbb{R}^1)$

$$|\langle \tau_a u, v \rangle| = |\langle u, \tau_a v \rangle| \lesssim \|u\|_{X^s} \langle a \rangle^s \|v\|_{X^{-s}},$$

which implies that $\|\tau_a v\|_{X^{-s}} \lesssim \|v\|_{X^{-s}} \langle a \rangle^{|-s|}$. ☐

Coming back to the formula (3.58), it is easy to get

$$\begin{aligned}|\mathcal{F}(u_1u_2)(\xi)| &\leq (2\pi)^{-n/2}\|\tau_\xi\mathcal{F}(u_1)\|_{X^s}\|\mathcal{F}(u_2)\|_{X^{-s}}\\ &\lesssim \|u_1\|_{X^s}\|u_2\|_{X^{-s}}\langle\xi\rangle^s,\end{aligned} \tag{3.60}$$

which can be written in the form of the next proposition.

Proposition 3.5.10. *If $u_1 \in X^s(\mathbb{R}^n)$ and $u_2 \in X^{-s}(\mathbb{R}^n)$ where s is nonnegative, then $\mathcal{F}(u_1u_2)(\xi)$ is a function with polynomial growth rate not greater than s.*

It has two corollaries.

Corollary 3.5.11. *Let $s \geq 0$. The multiplication operation is a continuous mapping from $X^s(\mathbb{R}^n) \times X^{-s}(\mathbb{R}^n)$ to the space $X^{(-s-n/2)^-}(\mathbb{R}^n)$. in particular, if $u \in X^s(\mathbb{R}^n)$ and $v \in X^{-s}(\mathbb{R}^n)$, then*

$$\|uv\|_{X^{(-s-n/2)^-}} \lesssim \|u\|_{X^s}\|v\|_{X^{-s}}.$$

Corollary 3.5.12. *Let $s \geq 0$. The multiplication operation is a continuous mapping from $X^{(s+n/2)^+}(\mathbb{R}^n)\times X^s(\mathbb{R}^n)$ to the space $X^s(\mathbb{R}^n)$. in particular, if $u \in X^{(s+n/2)^+}(\mathbb{R}^n)$ and $v \in X^s(\mathbb{R}^n)$, then*

$$\|uv\|_{X^s} \lesssim \|u\|_{X^{(s+n/2)^+}}\|v\|_{X^s}. \tag{3.61}$$

By using the fact that $\|uv\|_{X^{(-s-n/2)^-}} \lesssim \|uv\|_{H^{(-s-n/2)^-}}$, then the corollary 3.5.11 is a direct consequence of the proposition 3.5.10. We will provide the proof of corollary 3.5.12.

Proof. Assume w is an arbitrary function in the space $S(\mathbb{R}^n)$, then from the corollary 3.5.11 we have

$$|\langle uv, w\rangle| = |\langle u, vw\rangle| \lesssim \|u\|_{X^{(s+n/2)^+}}\|v\|_{X^s}\|w\|_{X^{-s}},$$

which implies that uv is a function in the space $X^s(\mathbb{R}^n)$ and satisfies the inequality (3.61). □

One natural question is if we are given the information of the regularities of the functions u and v, then what kind of information can we get for the regularity of their product uv. Recall that in the case of Sobolev spaces, one can also define the product of two functions as a tempered distribution satisfying $\langle u_1 u_2, \varphi\rangle = \langle u_1, u_2\varphi\rangle$ and there is such a result [Q-X-W].

Theorem 3.5.13. *If $u_j \in H^{s_j}(\mathbb{R}^n)$, $j = 1, 2$, and $s_1 + s_2 \geq 0$, then it is true that $u_1 u_2 = u_2 u_1 \in H^s(\mathbb{R}^n)$ and there is a constant C depending on s, s_1, and s_2 satisfying*

$$\|u_1 u_2\|_{H^s} \leq C(s, s_1, s_2)\, \|u_1\|_{H^{s_1}}\, \|u_2\|_{H^{s_2}}, \tag{3.62}$$

where $s \leq s_1$, $s \leq s_2$ and $s \leq s_1 + s_2 - n/2$. When one of the s_j equals $n/2$ or the minimum is just $-n/2$, the last inequality should be changed into strict inequality.

Corollary 3.5.14. *For $s > n/2$, the Sobolev space $H^s(\mathbb{R}^n)$ forms an algebra under the product defined above.*

In some special cases, the theorem 3.5.13 can have analogue in our function spaces X^s. For example, when $u \in X^s(\mathbb{R}^n)$ and $v \in X^{-s}(\mathbb{R}^n)$, we have corollary 3.5.11; when $u \in X^{s_1}(\mathbb{R}^n)$ and $v \in X^{s_2}(\mathbb{R}^n)$ satisfying $0 \leq s_1, s_2$ and $s_1 + s_2 < n/2$, then we have $\|u_1 u_2\|_{X^s} \lesssim \|u_1 u_2\|_{H^s} \leq C(s, s_1, s_2)\, \|u_1\|_{X^{s_1}}\, \|u_2\|_{X^{s_2}}$, where $s = s_1 + s_2 - n/2$. But for the general cases, it turns out to be a more complicated problem.

Here we turn our attention to a related but special question: what is the condition such that the function space $X^s(\mathbb{R}^n)$ is an algebra under the operation of multiplication. If our conjecture ($X^s(\mathbb{R}^n) = H^s(\mathbb{R}^n) \cap \mathcal{F}(H^s(\mathbb{R}^n))$ for all $s \geq 0$)

is true, then it is an easy task to verify that $X^s(\mathbb{R}^n) \subset H^s(\mathbb{R}^n)$ is an algebra for $s > n/2$. Furthermore we can get this subalgebra is an ideal of $H^s(\mathbb{R}^n)$. In other words, if $u \in H^s(\mathbb{R}^n)$ and $v \in X^s(\mathbb{R}^n)$, then $uv \in X^s(\mathbb{R}^n)$. Below we will prove that if m is the smallest integer larger that $n/2$, this property holds. Let us see the case of 1-dimension at first.

Theorem 3.5.15. *For any real number $s \geq 1$, the function space $X^s(\mathbb{R}^1)$ is an ideal of $H^s(\mathbb{R}^1)$ under the multiplication operation. in particular, if $u \in H^s(\mathbb{R}^1)$ and $v \in X^s(\mathbb{R}^1)$ then*

$$\|uv\|_{X^s} \lesssim \|u\|_{H^s} \|v\|_{X^s} \quad (s). \tag{3.63}$$

Proof. When $s = 1$, we have

$$\begin{aligned} \|uv\|_{X^1} &\approx \left\| \frac{d}{dx}(uv) \right\|_{L^2} + \|uv\|_{L^2} + \|uv\|_{\mathcal{F}H^1} \\ &\lesssim \|u\|_{L^\infty} \|v\|_{X^1} + \left\| \frac{d}{dx} u \right\|_{L^2} \|v\|_{L^\infty} + \|u\|_{L^\infty} \|v\|_{\mathcal{F}H^1} \\ &\lesssim \|u\|_{H^{(1/2)+}} \|v\|_{X^1} + \|u\|_{H^1} \|v\|_{X^{(1/2)+}} && (3.64) \\ &\lesssim \|u\|_{H^1} \|v\|_{X^1} . && (3.65) \end{aligned}$$

Similar reasoning can be done for other natural integer numbers.

Assume $s = m \in \mathbb{N}$, then from the characterization of the functions in $X^m(\mathbb{R}^1)$ (theorem 3.2.4) we have $\|uv\|_{X^m} \approx \sum_{0 \leq \alpha + \beta \leq m} \left\| x^\alpha (uv)^{(\beta)} \right\|_{L^2}$, which can be subsequently controlled by

$$\|uv\|_{X^m} \lesssim \sum_{0\le\alpha+\beta_1+\beta_2\le m} \left\|x^\alpha u^{(\beta_1)} v^{(\beta_2)}\right\|_{L^2}$$

$$\lesssim \|u\|_{H^{(1/2)+}} \|v\|_{X^m} + \|u\|_{H^{1+(1/2)+}} \|v\|_{X^{m-1}} + \cdots + \|u\|_{H^{m-1+(1/2)+}} \|v\|_{X^1} + \|u\|_{H^m} \|v\| \tag{3.66}$$

$$\lesssim \|u\|_{H^m} \|v\|_{X^m} \qquad (m). \tag{3.67}$$

We have proved that the theorem is true for all natural integer numbers.

Now assume that $s \in \mathbb{R}$ is between natural numbers 1 and m. Let $u \in H^s(\mathbb{R}^1)$ and $v = \sum_{k=0}^{+\infty} q_k h_k(x) \in X^s(\mathbb{R}^1)$. We define the mapping

$$f(z) = f_1(z) f_2(z) = \mathcal{F}^{-1}\left(\langle\xi\rangle^{-(m-1)(z-\theta)} \mathcal{F}u(\xi)\right) \textstyle\sum_{k=0}^{+\infty} q_k (k+1/2)^{-(m-1)(z-\theta)/2} h_k(x), \tag{3.68}$$

where $0 \le \theta = (s-1)/(m-1) \le 1$ and $z \in \Omega = \{z \in \mathbb{C} | 0 \le \operatorname{Re} z \le 1\}$. The mapping f is an analytical mapping on Ω° with values in $X^1(\mathbb{R}^1) + X^m(\mathbb{R}^1)$. This fact can be justified by these facts: $f_1(z)$ is analytic on Ω° with values in $H^1(\mathbb{R}^1) + H^m(\mathbb{R}^1)$; $f_2(z)$ is analytic on Ω° with values in $X^1(\mathbb{R}^1) + X^m(\mathbb{R}^1)$; together with $H^1 \cdot X^1 \subset X^1$ and $H^m \cdot X^m \subset X^m$.Meanwhile it is easy to verify that $f(\theta) = uv$ and the following two formulae

$$\|f(iy)\|_{X^1} \lesssim \|f_1(iy)\|_{H^1} \|f_2(iy)\|_{X^1} = \|u\|_{H^s} \|v\|_{X^s}$$

$$\|f(1+iy)\|_{X^m} \lesssim \|f_1(1+iy)\|_{H^m} \|f_2(1+iy)\|_{X^m} = \|u\|_{H^s} \|v\|_{X^s} \quad (m), \tag{3.69}$$

which implies that $|||f||| \lesssim \|u\|_{H^s} \|v\|_{X^s}$. Then by the complex interpolation theory and noticing that m can be chose as the smallest integers larger than s, it can be deduced that $uv \in X^s(\mathbb{R}^1)$ and it satisfies the inequality (3.63). $\square$

In the higher dimensional cases, we can still use the same method: first use the characterization of the functions in $X^m(\mathbb{R}^n)$ to prove that for all natural numbers m big enough the space $X^m(\mathbb{R}^n)$ is an ideal of the Sobolev spaces $H^m(\mathbb{R}^n)$; then by using the interpolation method to prove the theorem is true for all real numbers big enough. Since the proof is essentially same, we just skip it and give the statement as follows.

Theorem 3.5.16. *Let m is the smallest integer larger than $n/2$, that is, $m = 1 - \left[-\frac{n}{2}\right]$. For any real number $s \geq m$, the function space $X^s(\mathbb{R}^n)$ is an ideal of $H^s(\mathbb{R}^n)$ under the multiplication operation. Specifically, if $u \in H^s(\mathbb{R}^n)$ and $v \in X^s(\mathbb{R}^n)$ then*

$$\|uv\|_{X^s} \lesssim \|u\|_{H^s} \|v\|_{X^s} \quad (s). \tag{3.70}$$

Remark 3.5.17. *From the observation that $\mathcal{F}(u_1u_2) = (2\pi)^{-n/2}\mathcal{F}(u_1) * \mathcal{F}(u_2)$, there is a similar result stating that in the same conditions as above the function space $X^s(\mathbb{R}^n)$ is an ideal of $\mathcal{F}(H^s(\mathbb{R}^n))$ under the convolution operation.*

3.5.3 Relationship with Bargmann spaces

As the last section of the whole chapter, we would like to point out that our spaces X^s are "essentially" the Bargmann's spaces, which was once studied by V. Bargmann in the 1960's [Bar1][Bar2]. In these works, he first established a kind of integral transformation (the Bargmann transform) and showed that this transformation is a unitary mapping of $L^2(\mathbb{R}^n)$ onto the Fock's space F_n; then in the part two, he found that a family of related function spaces (the Bargmann's spaces) can be used to analyze the properties of tempered distribution. Below is a very rough review on his work and pay our attention to proving that our spaces are isomorphic to the Bargmann's spaces.

Let us start with the definition of Fock's space.

Definition 3.5.18. *Given an identification $\mathbb{R}^{2n} = \mathbb{C}^n$ ($z = (z_1, \cdots, z_n)$, $z_j = x_j + iy_j$), Fock space is the space of entire function on $\mathbb{C}^n$, with finite norm using the inner product*

$$(f_1(z), f_2(z)) = \pi^{-n} \int_{\mathbb{C}^n} \overline{f_1(z)} f_2(z) e^{-|z|^2} d^n z, \tag{3.71}$$

where $d^n z = \prod_{j=1}^n dx_j dy_j$.

Bargmann proved that there is a unitary mapping from complex valued function space $L^2(\mathbb{R}^n)$ onto the the Fock's space F_n:

$$f(z) = B_n(\psi) = \int_{\mathbb{R}^n} B_n(z,q)\psi(q) d^n q, \tag{3.72}$$

$$B_n(z,q) = \pi^{-n/4} \exp\{-\frac{1}{2}(z^2 + q^2) + \sqrt{2} z \cdot q\}, \tag{3.73}$$

where $z^2 = \sum z_j^2, q^2 = \sum q_j^2, z \cdot q = \sum z_j q_j$ and $f \in F_n$ if $\psi \in L^2(\mathbb{R}^n)$. This result can be justified by the fact that this transformation maps an orthonormal basis of $L^2(\mathbb{R}^n)$ as an orthonormal basis of F_n. To see this, let us introduce the some notation here about multi-indices: $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbb{Z}_+^n$, $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$, $z^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n}$ and $\alpha! = \alpha_1! \cdot \alpha_2! \cdots \alpha_n!$. As we have stated the function space $L^2(\mathbb{R}^n)$ has an orthonormal basis consisting of Hermite functions $\{h_\alpha(q) := h_{\alpha_1}(q_1) \otimes h_{\alpha_2}(q_2) \otimes \cdots \otimes h_{\alpha_n}(q_n)\}$. Then the Bargmann transformation maps $h_\alpha(q)$ as the function $u_\alpha(z) = z^\alpha / \sqrt{\alpha!}$, which makes up of an orthonormal basis of the Fock's space F_n when α runs over $\mathbb{Z}_+^n$. In fact, the kernel $B_n(z,q)$ can be read as $B_n(z,q) = \sum_{\alpha \in \mathbb{Z}_+^n} h_\alpha(q) u_\alpha(z)$.

A natural result is that the mapping B_n induces a unitary isomorphism between the linear operators on F_n and those on $L^2(\mathbb{R}^n)$. In the part one of the series of the papers "*On a Hilbert Space of Analytic Functions and an Associated Integral Transform*", *V. Bargmann* studied many operators on F_n and their corresponding operators on $L^2(\mathbb{R}^n)$ such as:

- $z' = c + Uz$ where $c \in \mathbb{C}^n$ and U is a linear unitary transformation; in particular, let $c = 0$ and U be a one parameter subgroup of the multiplier $e^{\imath\tau}$ (τ real numbers), it induces the Fourier transformation on $L^2(\mathbb{R}^n)$ when $\tau = \frac{1}{2}\pi$;

- The operators multiplier z_k and differential operator $\partial/\partial z_k$;

- Linear canonical transformation.

In the part two, *V. Bargmann generalized the definition of the Fock's space.*

Definition 3.5.19. *For every holomorphic function f on $\mathbb{C}^n$, the norm $\|f\|_\rho$ is given by*

$$\|f\|_\rho^2 = \int_{\mathbb{C}^n} |f(z)|^2 \, d\mu_n^\rho(z), \tag{3.74}$$

where $d\mu_n^\rho(z) = \pi^{-n}\theta_{2\rho}^{-2}(z)d^n z$ and $\theta_\rho^\gamma = \left(1+|z|^2\right)^{\rho/2} e^{\gamma|z|^2/2}$. It is said that f is a function in the space F_n^ρ if $\|f\|_\rho < +\infty$.

The space F_n^ρ turns out to be a Hilbert space, with inner product as $(f,g)_\rho = \int_{\mathbb{C}^n} f(z)\overline{g(z)}d\mu_n^\rho(z)$. In particular it has an orthonormal basis (see [Bar2] p36) $\{u_\alpha^\rho(z) = (\eta_{|\alpha|}^\rho)^{-1/2}u_\alpha(z)\}$, where $\eta_{|\alpha|}^\rho = \int_{\mathbb{C}^n} |u_\alpha(z)|^2 \, d\mu_n^\rho(z) = \frac{1}{\Gamma(n+|\alpha|)}\int_0^{+\infty}(1+x)^\rho x^{n+|\alpha|-1}e^{-x}dx$ satisfies $\eta_{|\alpha|}^\rho \approx (n+|\alpha|)^\rho$. Recall that our function space $X^\rho(\mathbb{R}^n)$ has an orthonormal basis $\{h_\alpha^\rho(q) = (|\alpha|+n/2)^{-\rho/2}h_\alpha(q)\}$ where $h_\alpha(q) = h_{\alpha_1}(q_1)\otimes h_{\alpha_2}(q_2)\otimes\cdots\otimes h_{\alpha_n}(q_n)$. Then it is easy to confirm that the linear mapping T from $X^\rho(\mathbb{R}^n)$ to F_n^ρ satisfying $T\left(h_\alpha^\rho(q)\right) = u_\alpha^\rho(z)$ is in fact an isomorphism. So we have such an result.

Theorem 3.5.20. *For each $\rho \in \mathbb{R}$, the function spaces $X^\rho(\mathbb{R}^n)$ is isomorphic to the Bargmann's space F_n^ρ.*

At that time V. Bargmann had realized that there are close relationships between his spaces and tempered distribution. For example it had been already known that the intersection of all the spaces F_n^ρ is "essentially" the rapid decreasing function and their union is "essentially" the tempered distribution. In his work [Bar2], V. Bargmann studied such function spaces: the Bargmann's space F_n^ρ, the function space $E_n = \cap_{k=-\infty}^{+\infty} F_n^k$ (corresponding to $S(\mathbb{R}^n)$), the function space $E_n' = \cup_{k=-\infty}^{+\infty} F_n^k$ (corresponding to $S'(\mathbb{R}^n)$), the function space E_n^ρ (the normed space with $|f|_\rho = \sup_{z\in\mathbb{C}^n} \theta_\rho^{-1}(z)\,|f(z)|$) and their interrelations. The introduction of these function spaces, especially E_n and E_n', provides an auxiliary tool for distribution theory. So the author applied them in the problems in the tempered distribution theory: the convergence in $S'(\mathbb{R}^n)$; the representation in E_n' for basic operations on $S'(\mathbb{R}^n)$, including the partial derivative operator, multiplier operator of the function q_j and the Fourier transformation; the regularity theorem and the kernel theorem of tempered distributions; some special tempered distributions like compactly supported distributions, periodic distributions and their Fourier expansions and homogeneous distributions.

The relationship between the V. Bargmann's spaces and tempered distribution was also noticed by Barry Simon. In the paper [Si], he used the method of the Hermite expansion for tempered distributions to establish the relationship of a sequence space and the tempered distribution. Then this method enabled him to study on an easier target-the sequence space- to get analysis results on the tempered distribution. He also mentioned the relationship of his work and that of V. Bargmann in that paper.

Our work initiates at the analysis on the linear partial differential equation $i\psi_t = \frac{1}{2}\psi_{xx} - \frac{x^2}{2}\psi = A\psi$, where the space X^n is in face the definition domain of the operator $A^{n/2}$ and the Hermite functions arise as the eigenfunctions of the operator A. This is quite different from the work by B. Simon and V. Bargmann, since one of them initiated from the realization of $S(\mathbb{R}^n)$ as a sequence space, which can be traced back to Schwartz's book [Sc]; the other one initiated from the relationship between the Fock space F_n and L^2 space.

Our work also differs from the work by B. Simon and V. Bargmann on contents. We directly fitted the function spaces $X^s(\mathbb{R}^n)$ well in the theoretic structure of Hilbert scales, and most basic operations on those spaces are regarded as the homomorphisms on the Hilbert scale. Like what B. Simon and V. Bargmann have done, we also studied some problems on the topological structure of the spaces $S(\mathbb{R}^n)$ and $S'(\mathbb{R}^n)$, such as locally convex topology, characterization of sequence convergence and characterization of linear operations on $S(\mathbb{R}^n)$ or $S'(\mathbb{R}^n)$ etc. But we restrict these parts in the level of just describing the relationship of the function spaces $X^s(\mathbb{R}^n)$ and the spaces $S(\mathbb{R}^n)$ and $S'(\mathbb{R}^n)$, and many important properties for these two spaces are not covered. Back to B. Simon and V. Bargmann's work, the spaces $S(\mathbb{R}^n)$ and $S'(\mathbb{R}^n)$ themselves were the research target, and the results like the regularity theorem, kernel theorem and sequence completeness were too crucial to be skipped over. In our work, the viewpoint is mostly kept from the analysis, especially PDE, and much concern is paid on the study in the relationship between the Sobolev spaces and the spaces $X^s(\mathbb{R}^n)$ and the problem of how these spaces can be used in nonlinear PDEs (in particular the definition of the product of two functions). In this process, the properties of Hermite functions take some important roles.

Chapter 4

Birkhoff Normal Form

4.1 Introduction

Let us consider the nonlinear Schrödinger equation

$$\begin{cases} i\psi_t = \frac{1}{2}\psi_{xx} - \frac{x^2}{2}\psi - g|\psi|^2\psi & x \in \mathbb{R}^1 \\ \psi(x,0) = \psi_0(x) & \psi \text{ complex valued,} \end{cases} \tag{4.1}$$

where $g = 1$ ($g = -1$) is for defocusing (focusing) cubic nonlinearity. This equation is also known as the Gross-Pitaevskii (GP) equation with a parabolic potential, which was brought up in the theoretical study for Bose-Einstein condensation in 1960's [Gro][Pi]. Since a Bose-Einstein condensate was produced for the first time in the experimental condition in 1995, this partial differential equation has received a lot of attention.

There are already many mathematical papers [O][Z][C1][C2] with emphasis on the local or global well-posedness and blow up conditions of the GP equation in general dimensions. One of the results in this work is that in one dimensional case the solution will always globally exist in the space $X^1(\mathbb{R}) = \Sigma$ with conserved mass and energy. in particular, the equation will always have global solution in the energy space $\Sigma := \{u \in H^1(\mathbb{R}) | xu \in L^2(\mathbb{R})\}$, no matter whether the constant g is negative or nonnegative, and the solution conserves the mass $M = \int_{-\infty}^{+\infty} |\psi|^2 \, dx$ and the energy $E = \frac{1}{2}\int_{-\infty}^{+\infty} |\psi_x|^2 + |x\psi|^2 \, dx + \frac{g}{2}\int_{-\infty}^{+\infty} |\psi|^4 \, dx$.

In this thesis, we will discuss this equation in a wider class of function spaces, which we believe as a very natural choice. One advantage in doing so is that it enables us to consider the equation for the initial data with different regularities, even not within the space Σ. Most part of these results will be provided in the section 4.2.

We comment here that this GP equation has a Hamiltonian structure. Let $H = \frac{1}{2}\int_{-\infty}^{+\infty} |\psi_x|^2 + |x\psi|^2 \, dx + \frac{g}{2}\int_{-\infty}^{+\infty} |\psi|^4 \, dx$, then the right side of the equation (4.1) can be written as

$$\frac{1}{2}\psi_{xx} - \frac{x^2}{2}\psi - g|\psi|^2\psi = -\frac{\partial H}{\partial \overline{\psi}}. \tag{4.2}$$

So the equation (4.1) can be changed in this form

$$\begin{cases} \psi_t = i\frac{\partial H}{\partial \overline{\psi}} & x \in \mathbb{R}^1 \\ \psi(x,0) = \psi_0(x) & \psi \text{ complex valued,} \end{cases} \tag{4.3}$$

which is an infinite dimensional Hamiltonian system. In this thesis, we will make use of our function spaces $X^{2s}(\mathbb{R})$ and take the viewpoint from Hamiltonian PDE to study the GP equation. In this process, the Hilbert scale $X^{2s}(\mathbb{R})$ will provide us the spaces to work in and many technical tools (especially the Birkhoff Normal form) from the theory of Hamiltonian systems will be applied. Below let us have a quick review on the Hamiltonian formalism in infinite dimension [Gre].

Typically a Hamiltonian system in finite dimension reads

$$\begin{cases} \dot{x}_j = \frac{\partial H}{\partial y_j}, & j = 1, \cdots, n \\ \dot{y}_j = -\frac{\partial H}{\partial x_j}, & j = 1, \cdots, n \end{cases} \tag{4.4}$$

where the point (x_j, y_j) is in the **phase space** (or configuration space) M, an open set in $\mathbb{R}^{2n}$ and the **Hamiltonian function** H is a regular real valued function, on the phase space M, that is, $H \in C^\infty(M, \mathbb{R})$. By introducing the canonical **Poisson matrix,** that is,

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

and **Hamiltonian vector field**

$$X_H(x, y) = J\nabla_{x,y} H(x, y),$$

where $\nabla_{x,y} H(x, y)$ denotes the gradient of H with respect to x, y, the Hamiltonian system (4.4) then can be written as

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = X_H(x, y) = \begin{pmatrix} \frac{\partial H}{\partial y_1} \\ \vdots \\ \frac{\partial H}{\partial y_n} \\ -\frac{\partial H}{\partial x_1} \\ \vdots \\ -\frac{\partial H}{\partial x_n} \end{pmatrix}. \tag{4.5}$$

A very basic concept in the theory of Hamiltonian systems is the **Poisson bracket** of two functions defined as: for any two functions F, G, the Poisson bracket of them is a new function $\{F, G\}$ given by

$$\{F, G\}(x, y) = \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(x, y)\frac{\partial G}{\partial y_j}(x, y) - \frac{\partial F}{\partial y_j}(x, y)\frac{\partial G}{\partial x_j}(x, y). \tag{4.6}$$

The Poisson bracket $\{F, H\}$ provide a very good representation of the changing values of the function $F(x, y)$ along the flow of the Hamiltonian system associated to H. In particular, if $t \mapsto (x(t), y(t))$ is a solution of the system (4.4), then

$$\frac{d}{dt}F(x(t), y(t)) = \{F, H\}(x(t), y(t)).$$

Moreover, if $F \in C^{\infty}(M, \mathbb{R})$ satisfies $\{F, H\} = 0$, then it is said that F is **an integral of motion** for H. Obviously, in this setting the function H itself must be an integral of motion for the Hamiltonian system.

Since solutions to a Hamiltonian system will conserve integrals of motion, the initial value problem is solved on the intersection of the level sets of those integral functions, thus a Hamiltonian system becomes easier to study. In some cases, one can find n independent integrals for a given $2n$-dimensional Hamiltonian system. It is said that a $2n$-dimensional Hamiltonian system is **integrable in the sense of Liouville**, If there exist n regular functions $F_1, F_2, \cdots, F_n : M \to \mathbb{R}$ such that: (i) $\{F_j, H\} = 0$ for $j = 1, \cdots, n$; (ii) $\{F_j, F_k\} = 0$ for $j, k = 1, \cdots, n$ (that is, the F_j are in involution); (iii) $(\nabla_{x,y} F_j)_{j=1,\cdots,n}$ are linearly independent. If the last condition is not satisfied on the whole space, but on a dense open subset , it is often called a **Birkhoff integrable Hamiltonian system.**

There is one simple example here for the above definition of integrable Hamiltonian system. Let $M = \mathbb{R}^{2n}$ and

$$H(x, y) = \sum_{j=1}^{n} \omega_j \frac{(x_j^2 + y_j^2)}{2},$$

where $\omega = (\omega_1, \cdots, \omega_n)^t \in \mathbb{R}^n$ is the **frequency vector**. The associated Hamiltonian system is called **harmonic oscillator**, whose solutions are all **quasi-periodic** given by

$$\begin{cases} x_j(t) = x_j(0)\cos\omega_j t + y_j(0)\sin\omega_j t, & j = 1,\cdots,n \\ y_j(t) = -x_j(0)\sin\omega_j t + y_j(0)\cos\omega_j t, & j = 1,\cdots,n. \end{cases} \tag{4.7}$$

It is easy to verify that the functions $I_j = (x_j^2 + y_j^2)/2$ are all integrals of this system and they satisfy the conditions (i), (ii) and they satisfy (iii) on a dense open subset. In other words, this is a Birkhoff integrable Hamiltonian system. It is deserved to mention that by introducing new complex parameter $z_j = (x_j - iy_j)/\sqrt{2}$, then the equation (4.4) can be written as

$$\frac{d}{dt}z_j = i\frac{\partial H}{\partial \overline{z_j}}, \quad j = 1,\cdots,n \tag{4.8}$$

and its solution have a clear form as

$$z_j(t) = z_j(0)e^{-i\omega_j t}. \quad j = 1,\cdots,n. \tag{4.9}$$

In many important physics models, the corresponding Hamiltonian systems may be not integrable themselves, but they can be regarded as a perturbation of an integrable system. A general philosophy in this situation is to transform the Hamiltonian in a way such that the new Hamiltonian system is closer to an integrable one. In particular, if we have a Hamiltonian function $H = H_0 + P$ where H_0 is integrable and P is a perturbation term, then we want to find a transformation Γ on the phase space such that $H \circ \Gamma = \widetilde{H_0} + \widetilde{P}$ with $\widetilde{H_0}$ still integrable and $\widetilde{P} \ll P$. Since we need such a transformation to conserve the Hamiltonian structure, it is a natural thing to restrict our consideration within this class of transformations.

Definition 4.1.1. *A map $\Gamma : M \ni (x,y) \mapsto (\xi,\eta) \in M$ is a* ***canonical transformation (or symplectic transformation)*** *if it satisfies: (i) Γ is a diffeomor-*

phism; (ii) Γ *preserves the Poisson bracket, that is,* $\{F,H\}(x,y)=\{\widetilde{F},\widetilde{H}\}(\xi,\eta)$ *with* $\widetilde{F}=F\circ\Gamma^{-1}$.

This definition has an equivalent characterization: the diffeomorphism Γ preserves the 2-form $\omega^2=\sum_{j=1}^n dx_j\wedge dy_j$ on the space M, that is $\Gamma^*\omega^2=\omega^2$. Then under such a transformation, the Hamiltonian system appears in the new variables (ξ,η) as

$$\dot{\xi}_j=\frac{\partial\widetilde{H}}{\partial\eta_j},\eta_j=-\frac{\partial\widetilde{H}}{\partial\xi_j},\ j=1,\cdots,n, \tag{4.10}$$

which is essentially same as (4.4). One easy way to construct canonical transformations is by a Lie transform.

Definition 4.1.2. *Let* $\chi: M\to\mathbb{R}$ *be a regular function and denote* $\Phi(t,x,y)$ *is the flow generated by Hamiltonian vector field* X_χ *with initial data* (x,y)*. Then the map* $\Gamma(x,y):=\Phi(1,x,y)$ *(if available) is called the* ***Lie transform*** *associated to* χ.

A Lie transformation may not be well-defined for every point $(x,y)\in M$, but for any regular function it is must be locally well defined in a neighborhood of the zero point, and more important it is always canonical. By making use of these canonical transformations, many Hamiltonian systems can be changed into the form of $\widetilde{H}=H_0+P+R$, where $H_0=\sum_{j=1}^n\omega_jI_j$, P satisfying $\{P,H_0\}=0$ (in **normal form**) is at least cubic, and R is a higher order term than P. Moreover, under some kind of nonresonant conditions, the term P can be chosen to depend only on the parameter I_j, which implies that the truncated system $\widetilde{H}_{tr}=H_0+P$ can be completely solved.

The Hamiltonian formalism introduced above can be generalized into infinite dimensional case. In this thesis, we will use this idea to deal with the 1-dimensional

Gross-Pitaevskii equation. One basic difficulty here is that a good understanding for the Hermite functions, especially of the asymptotic behavior of integrals of products of Hermite functions, is required in this process. Meanwhile, the Hamiltonian system corresponding to the GP equation is completely resonant, hence we can't expect to have the term P in integrable form. Here we provide the main result of us in this Hamiltonian formalism.

Theorem 4.1.3. *For the Hamiltonian $H = H_2 + H_4$ corresponding to the 1-dimensional GP equation, there exists a real analytic, symplectic change of coordinates Γ in a neighborhood of the origin $V \subset l_2^s(\mathbb{Z}_+;\mathbb{C}) \cong X^{2s}(\mathbb{R})$ with $2s \geq 1$, that takes H into its Birkhoff normal form up to order 4. That is $H \circ \Gamma = H_2 + G + R$ with the following properties:*

(i) $G(p) = \frac{g}{2} \sum_{k+l=m+n} C_{klmn} \overline{p_k} \overline{p_l} p_m p_n$ is a continuous polynomial of degree four with a regular vector field.

(ii) $R \in \aleph^s(V)$ and $\| X_R(p) \|_{l_2^s} \leqslant C_s \| p \|_{l_2^s}^5$ for all $p \in V$.

(iii) Γ is close to the identity: $\| \Gamma(p) - Id(p) \|_{l_2^s} \leqslant C_s \|p\|_{l_2^s}^3$ for all $p \in V$.

This theorem says, as in the finite dimensional case, that we can change the coordinates in a neighborhood of the origin in such a way that the Hamiltonian is in normal form up to order four. This does not mean that the Birkhoff normal form Hamiltonian is integrable, as there are nontrivial resonances. Its proof consists of several part of analysis and will be provided from the section 4.3 to the section 4.4. In the subsequent sections, we will study some truncated Hamiltonian systems and discuss on the impact of the perturbation term to a class of special solutions to the truncated system. Now let us begin with section 4.2, devoted to the application of the function spaces X^{2s} into the GP equation.

4.2 The equation in the function space X^{2s}

In this section, we will use the traditional way to study the local (global) well-posedness of the PDE (4.1) in our function spaces X^{2s}. Then we will write the PDE in the function space $l_2^s(\mathbb{Z}_+)$, which is equivalent to the space $X^{2s}(\mathbb{R})$. In this way, we can adapt the system into the Hamiltonian scheme, and the discussion for the solution with rougher initial data in $X^1(\mathbb{R})$ can be made much easier. Some basic properties of the solution like mass conservation and energy conservation will also be discussed.

4.2.1 Local and global existence

Let us consider the 1-dimensional GP equation in the form

$$\begin{cases} i\psi_t = -A\psi - g|\psi|^2\psi & x \in \mathbb{R}^1 \\ \psi(x,0) = \psi_0(x) & \psi \text{ complex valued.} \end{cases} \tag{4.11}$$

We say that a continuous curve $\psi : t \ni I \mapsto \psi(t) \in X^{2s}(\mathbb{R})$ in the function space $X^{2s}(\mathbb{R})$ (that is, $\psi \in C(I, X^{2s}(\mathbb{R}))$) is a solution of the equation (4.11), if it satisfies for any time $t \in I$

$$\psi(t) = e^{iAt}\psi_0 + ig\int_0^t e^{iA(t-\tau)}|\psi(\tau)|^2\psi(\tau)d\tau, \tag{4.12}$$

where I is a neighborhood of the origin point. In the following section, we will discuss on the local existence of the solution and show that the solution will globally exist for any initial data smooth enough.

Theorem 4.2.1. *For any real number $2s \geq 1$ and any function $\psi_0 \in X^{2s}(\mathbb{R})$, the initial value problem (4.11) has a unique solution $\psi(\cdot)$ defined in a time interval $[0,T]$, $T = T(\|\psi_0\|_{X^{2s}}) > 0$, satisfying*

$$\psi \in C\left([0,T], X^{2s}(\mathbb{R})\right) \cap C^1\left([0,T], X^{2s-2}(\mathbb{R})\right).$$

Proof. Consider a nonlinear mapping

$$F\psi(t) := e^{iAt}\psi_0 + ig\textstyle\int_0^t e^{iA(t-\tau)}|\psi(\tau)|^2\psi(\tau)d\tau. \tag{4.13}$$

Then the initial value problem can be reduced to a problem of finding a fixed point of the nonlinear mapping F, which can be achieved by showing that the mapping F is a contraction mapping.

Since e^{iAt} is always a unitary mapping on the space $X^{2s}(\mathbb{R})$ and for $2s \geq 1$ the space $X^{2s}(\mathbb{R})$ is an algebra under the multiplication operation, we can have

$$\begin{aligned}
\|F\psi(t)\|_{X^{2s}} &\leq \|\psi_0\|_{X^{2s}} + |g| \textstyle\int_0^t \left\|e^{iA(t-\tau)}|\psi(\tau)|^2\psi(\tau)\right\|_{X^{2s}} d\tau \\
&\leq \|\psi_0\|_{X^{2s}} + \textstyle\int_0^t \|\psi(\tau)\|_{X^{2s}}^3 \, d\tau \\
&\leq \|\psi_0\|_{X^{2s}} + T \max_{t\in[0,T]} \|\psi(\tau)\|_{X^{2s}}^3 \, d\tau.
\end{aligned}$$

So there is a constant $M = M(\|\psi_0\|_{X^{2s}}) \geq 2\,\|\psi_0\|_{X^{2s}}$ such that for all $T \leq \|\psi_0\|_{X^{2s}}/M^3$, the operation F is a continuous mapping on the closed subset of $C([0,T], X^{2s}(\mathbb{R}))$, $B = \{\psi \in C([0,T], X^{2s}(\mathbb{R})) \mid \|\psi\|_{X^{2s}} \leq M\}$. Meanwhile, for any two functions

$\psi_1, \psi_2 \in B$, we have

$$\begin{aligned}\|F\psi_1 - F\psi_2\|_{X^{2s}} &= \left\| ig\textstyle\int_0^t e^{\imath A(t-\tau)}|\psi_1(\tau)|^2\psi_1(\tau) - e^{\imath A(t-\tau)}|\psi_2(\tau)|^2\psi_2(\tau)d\tau \right\|_{X^{2s}} \\ &\leq \textstyle\int_0^t \left\| (\psi_1(\tau) - \psi_2(\tau))\overline{\psi_1(\tau)}\psi_1(\tau) \right\|_{X^{2s}} + \left\| \psi_2(\tau)\overline{(\psi_1(\tau) - \psi_2(\tau))}\psi_1(\tau) \right\|_{X^{2s}} \\ &+ \left\| \psi_2(\tau)\overline{\psi_2(\tau)}\,(\psi_1(\tau) - \psi_2(\tau)) \right\|_{X^{2s}} d\tau \\ &\leq \max_{t\in[0,T]} \|(\psi_1(\tau) - \psi_2(\tau)\|_{X^{2s}} M^2 T \leq 1/2 \max_{t\in[0,T]} \|(\psi_1(\tau) - \psi_2(\tau)\|_{X^{2s}}.\end{aligned}$$

So the mapping F is really a contraction mapping and thus it can be concluded that there is locally a unique fixed point for the operation defined in (4.13). Note that $e^{\imath At}\psi_0, \int_0^t e^{\imath A(t-\tau)}|\psi(\tau)|^2\psi(\tau)d\tau$ are in $C^1([0,T], X^{2s-2}(\mathbb{R}))$, then all the conclusions in the theorem can be obtained. □

Remark 4.2.2. *This proof has essentially used the fact that the space $X^{2s}(\mathbb{R})$ is an algebra for any $2s \geq 1$. If our conjecture 3.4.6, which is for the characterization of the space $X^{2s}(\mathbb{R})$, is right, then the space can be guaranteed to be an algebra when $2s$ is bigger than $1/2$. Thus the local existence theorem can be easily generalized to the case $2s > 1/2$.*

After knowing that $\psi \in C([0,T], X^{2s}(\mathbb{R})) \cap C^1([0,T], X^{2s-2}(\mathbb{R}))$ is the unique solution with initial data $\psi(x,0) = \psi_0(x)$, we can then consider the same PDE with new iniatial data $\widetilde{\psi}(x,0) = \psi(x,T)$. Then the solution can be uniquely extended for a little more time $\widetilde{T} = \widetilde{T}(\|\psi(T)\|_{X^{2s}}) > 0$. This process can be done repeatedly. Since for any initial data the solution exists locally and uniquely, it can be easily deduced that there exists a maximal interval $I = [0, T^*)$ such that $\psi \in C(I, X^{2s}(\mathbb{R})) \cap C^1(I, X^{2s-2}(\mathbb{R}))$ satisfies the integral equation (4.12). Here there are only two possibilities for the maximal time T^*: either $T^* = +\infty$ or $0 < T^* < +\infty$

and $\lim_{t \to (T^*)^-} \|\psi(t)\|_{X^{2s}} = +\infty$. Using the same reasoning for the negative time direction, we can know that there exists a minimal time $T_* < 0$ with similar properties.

Note that at the endpoint of the maximal solution interval (if finite) the X^{2s} norm of the solution must tend to infinity, so if at any finite time this phenomenon will never happen, we can then conclude that the solution must exist globally. Below we will use this method to get the global existence. During that process, these facts will be needed: the solution will conserve the mass $M = \int_{-\infty}^{+\infty} |\psi(t,x)|^2 \, dx$ and the energy (Hamiltonian) function $H = \frac{1}{2}\int_{-\infty}^{+\infty} |\psi_x|^2 + |x\psi(t,x)|^2 \, dx + \frac{g}{2}\int_{-\infty}^{+\infty} |\psi|^4 \, dx$. These facts will be proved in a latter part of this section.

Theorem 4.2.3. *For any $n \in \mathbb{N}$ and any function $\psi_0 \in X^n(\mathbb{R})$, the initial value problem (4.11) has a unique global solution $\psi(\cdot)$ defined in time interval $(-\infty, +\infty)$. The solution ψ is in the space $C(\mathbb{R}_t, X^n(\mathbb{R})) \cap C^1(\mathbb{R}_t, X^{n-2}(\mathbb{R}))$.*

Proof. Let us first consider the case of $n = 1$. If $g = 1$, the proof is simple, since the conserved energy function $H = H(\psi_0) \geq \frac{1}{2} \|\psi(t,x)\|_{X_x^1}^2$, so we can deduce that $\|\psi(t,x)\|_{X_x^1}$ is always bounded. By the above reasoning about the necessary condition on the blowing up condition at a finite time T^*, we know that this condition will not be satisfied, thus the solution must exist globally. If $g = -1$, we claim that the X^1 norm of the solution also remains bounded, which can result in the global existence of the solution. From the continuous imbedding of the space $X^{1/4}(\mathbb{R})$ into the space $L^4(\mathbb{R})$, we get $\|\psi(t)\|_{L_x^4} \lesssim \|\psi(t)\|_{X_x^{1/4}}$; meanwhile, from the complex interpolation theory on the Hilbert scale $X^{2s}(\mathbb{R})$, we get $\|\psi(t)\|_{X_x^{1/4}} \leq \|\psi(t)\|_{X_x^0}^{3/4} \|\psi(t)\|_{X_x^1}^{1/4}$. Then by using the conservation laws, we get

$$\begin{aligned} H(\psi_0) &= \|\psi(t)\|^2_{X^1_x} - \frac{1}{2}\|\psi(t)\|^4_{L^4_x} \\ &\geq \|\psi(t)\|^2_{X^1_x} - \frac{C}{2}\|\psi(t)\|^3_{X^0_x}\|\psi(t)\|_{X^1_x} \\ &\geq \|\psi(t)\|^2_{X^1_x} - \frac{C}{2}M^{3/2}\|\psi(t)\|_{X^1_x}. \end{aligned}$$

So it can be concluded that the X^1 norm of the solution remains bounded.

Now let us see the case of integers $n \geq 2$. Since the function space $X^n(\mathbb{R}) = H^n(\mathbb{R}) \cap \mathcal{F}H^n(\mathbb{R})$ and their norms are also equivalent, we turn to estimate the growth rate of the $H^n(\mathbb{R})$ and $\mathcal{F}H^n(\mathbb{R})$ norm of the solutions. In formal computation we have

$$\begin{aligned} \frac{d}{dt}\int_{-\infty}^{+\infty}\partial_x^n\psi\overline{\partial_x^n\psi}dx &= \operatorname{Re}\int_{-\infty}^{+\infty}\partial_x^n\left(\imath A\psi + ig|\psi|^2\psi\right)\overline{\partial_x^n\psi}dx \\ &= \operatorname{Im}\int_{-\infty}^{+\infty}\partial_x^n(A\psi)\overline{\partial_x^n\psi}dx + \operatorname{Im}\int_{-\infty}^{+\infty}\partial_x^n\left(g|\psi|^2\psi\right)\overline{\partial_x^n\psi}dx \\ &= \operatorname{Im}\int_{-\infty}^{+\infty}[\partial_x^n, A]\psi\overline{\partial_x^n\psi}dx + g\operatorname{Im}\int_{-\infty}^{+\infty}\partial_x^n\left(|\psi|^2\psi\right)\overline{\partial_x^n\psi}dx \qquad (4.14) \\ &= I + II, \end{aligned}$$

since the term $\int_{-\infty}^{+\infty}A(\partial_x^n\psi)\overline{\partial_x^n\psi}dx$ must be real and thus its imaginary part is zero. This computation is said to be a formal one for $\psi \in C(I, X^n(\mathbb{R}))$ because the term $\int_{-\infty}^{+\infty}\partial_x^n(A\psi)\overline{\partial_x^n\psi}dx$ requires the function $\psi(t) \in X^{n+1}(\mathbb{R})$, otherwise it may not exist.

But observing that the right side term of equality (4.14) is well-defined for $\psi \in C(I, X^n(\mathbb{R}))$, we claim that that equality is in fact true for any solution $\psi \in C(I, X^n(\mathbb{R}))$. The justification is like this; the whole computation holds for any function in $C(I, X^{n+1}(\mathbb{R}))$; for any time $t \in I$ and $|\Delta t| \ll 1$ fixed, we can choose $\widetilde{\psi}(t) \in X^{n+1}(\mathbb{R})$ which is arbitrarily close to the function $\psi(t)$ in a period $[t, t+\Delta t]$; taking the derivative on parameter t on the term $\int_{-\infty}^{+\infty}\partial_x^n\widetilde{\psi}\overline{\partial_x^n\widetilde{\psi}}dx$, we get term in the right side of equality (4.14) for function $\widetilde{\psi}$; since the result is well defined for $\psi \in C(I, X^n(\mathbb{R}))$, so let $\widetilde{\psi}(t)$ tend to $\psi(t)$ then we can get what we have claimed.

We can now continue to estimate the term I and II. For the first term, since $[\partial_x^n, A] = nx\partial_x^{n-1} + \frac{n(n-1)}{2}\partial_x^{n-2}$, it proceeds as

$$\begin{aligned}
|I| &= \left|\operatorname{Im}\int_{-\infty}^{+\infty} nx\partial_x^{n-1}\psi\overline{\partial_x^n\psi}dx - \operatorname{Im}\int_{-\infty}^{+\infty}\frac{n(n-1)}{2}\partial_x^{n-1}\psi\overline{\partial_x^{n-1}\psi}dx\right| \\
&\lesssim \left\|x\partial_x^{n-1}\psi\right\|_{L^2}\left\|\partial_x^n\psi\right\|_{L^2} \\
&\lesssim \|\psi\|_{X^n}^2 .
\end{aligned}$$

For the second term, we have

$$\begin{aligned}
|II| &= \left|\operatorname{Im}\int_{-\infty}^{+\infty}\sum_{n_1+n_2+n_3=n}\frac{n!}{n_1!n_2!n_3!}\partial_x^{n_1}\psi\partial_x^{n_2}\overline{\psi}\partial_x^{n_3}\psi\overline{\partial_x^n\psi}dx\right| \\
&\lesssim \sum_{n_1+n_2+n_3=n}\frac{n!}{n_1!n_2!n_3!}\left\|\partial_x^{n_1}\psi\right\|_{L^\infty}\left\|\partial_x^{n_2}\psi\right\|_{L^\infty}\left\|\partial_x^{n_3}\psi\right\|_{L^2}\left\|\partial_x^{n}\psi\right\|_{L^2} .
\end{aligned}$$

Without loss of generality we can assume that $n_3 \geq n_2 \geq n_1$, which guarantees that $n_1, n_2 < n$. The term $\left\|\partial_x^{n_j}\psi\right\|_{L^\infty}$ $(j = 1, 2)$ can be controlled by

$$\begin{aligned}
\left\|\partial_x^{n_j}\psi\right\|_{L^\infty} &\lesssim \|\psi\|_{X^{n_j+(1/2)^+}} \\
&\lesssim \|\psi\|_{X^1}^{\frac{n-(n_j+(1/2)^+)}{n-1}}\|\psi\|_{X^n}^{\frac{(n_j+(1/2)^+)-1}{n-1}} ;
\end{aligned}$$

the term $\left\|\partial_x^{n_3}\psi\right\|_{L^2}\left\|\partial_x^{n}\psi\right\|_{L^2}$ can be controlled by

$$\left\|\partial_x^{n_3}\psi\right\|_{L^2}\left\|\partial_x^{n}\psi\right\|_{L^2} \lesssim \|\psi\|_{X^1}^{\frac{n-n_3}{n-1}}\|\psi\|_{X^n}^{\frac{n_3-1}{n-1}}\|\psi\|_{X^n} .$$

Thus their product has a power of the term $\|\psi\|_{X^n}$ as $1+\frac{n_1+n_2+n_3+1^+-3}{n-1} = 2-\frac{1^-}{n-1} < 2$. From the estimate of the solution for the case $n = 1$, we know $\|\psi\|_{X^1}$ will always be bounded, so we can have $|II| \lesssim \|\psi\|_{X^n}^{2-\frac{1^-}{n-1}}$. The only exception is the case of $n_1 = 0$ or $n_1 = n_2 = 0$, where the term $\left\|\partial_x^{n_j}\psi\right\|_{L^\infty} = \|\psi\|_{L^\infty}$ can be controlled instead by

$\|\psi\|_{X^1}$ and thus keep bounded. This will result in a little rise on the power of the term $\|\psi\|_{X^n}$, but anyway it can be controlled as

$$|II| \lesssim \|\psi\|_{X^n}^2 .$$

Now it can be concluded that

$$\left|\frac{d}{dt}\int_{-\infty}^{+\infty}\partial_x^n\psi\overline{\partial_x^n\psi}dx\right| \lesssim \|\psi\|_{X^n}^2 .$$

Similarly, we can get the estimate

$$\left|\frac{d}{dt}\int_{-\infty}^{+\infty}x^n\psi\overline{x^n\psi}dx\right| \lesssim \|\psi\|_{X^n}^2 .$$

Putting these estimate together, it can be deduced that

$$\frac{d}{dt}\|\psi\|_{X^n}^2 \lesssim \|\psi\|_{X^n}^2 .$$

So the X^n norm of the solution can have at most an exponential growth. In particular, the initial value problem must have a unique global solution on time. □

4.2.2 The equation in q coordinates

Since the function space $X^{2s}(\mathbb{R})$ is equivalent to the space $l_2^s(\mathbb{Z}_+)$, the nonlinear Schrödinger equation (4.1) can be written in the q coordinate, where $q = (q_0, q_1, \cdots)$ is a point in the space $l_2^s(\mathbb{Z}_+)$ satisfying that $\psi(t,x) = \sum_{k=0}^{+\infty} q_k(t)h_k(x) \in X^{2s}(\mathbb{R})$. It reads as

$$\begin{cases} -i\frac{d}{dt}q = I_w q + gJq, \quad q \in l_2^s(\mathbb{Z}_+) \\ \qquad q(0) = q_0, \end{cases} \tag{4.15}$$

where $g = \pm 1$, the operator I_w is the mapping corresponding to the operator A and the operator J is the mapping corresponding to the operator $|\psi|^2\psi$. in particular,

I_w is the mapping from l_2^s to l_2^{s-1} defined by $(I_w q)_k = \omega_k q_k = (k+1/2)q_k$; J is the mapping formally defined by $\sum_{l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_l}q_m q_n$. We say that a continuous curve $q : t \ni I \mapsto q(t) \in l_2^s(\mathbb{Z}_+)$ in the function space $l_2^s(\mathbb{Z}_+)$ (that is, $q \in C(I, l_2^s(\mathbb{Z}_+))$) is a solution of the equation (4.15), if it satisfies

$$q(t) = e^{iI_\omega t}q_0 + ig\textstyle\int_0^t e^{iI_\omega(t-\tau)}Jq(\tau)d\tau. \tag{4.16}$$

It is easy to verify that $q(t) \in C(I, l_2^s(\mathbb{Z}_+))$ is a solution of the integral equation (4.16) if and only if $\psi(t,x) = \sum_{k=0}^{+\infty} q_k(t)h_k(x)$ belongs to the space $C(I, X^{2s}(\mathbb{R}))$ and satisfies the equation (4.12). Then from the results in the theorem 4.2.1 and theorem 4.2.3, we can know that in the cases of both $g=1$ and $g=-1$, for any real number $2s \geq 1$ and any initial data $q_0 \in l_2^s(\mathbb{Z}_+)$, the solution $q(t)$ will always locally exist in the space $l_2^s(\mathbb{Z}_+)$; if $2s = n \in \mathbb{N}$, then the solution will globally exists.

In the equality (4.16), the operator I_w is clearly defined on the space $l_2^s(\mathbb{Z}_+)$ for any $s \in \mathbb{R}$, but the situation is very different for the operator J. This is because to define the product of two tempered distribution we need a condition on their regularities, namely $s_1 + s_2 \geq 0$. So we want to clarify here what is the condition to guarantee the operator J to be well defined and how it looks like.

Proposition 4.2.4. *For $2s \geq 1/6$, the operator J is well defined from the space $l_2^s(\mathbb{Z}_+)$ into the space $l_2^{-\infty}(\mathbb{Z}_+)$, and for any $k \in \mathbb{Z}_+$*

$$(Jq)_k = \lim_{N_1,N_2,N_3\to+\infty} \sum_{l=0}^{N_1}\sum_{m=0}^{N_2}\sum_{n=0}^{N_3} C_{klmn}\overline{q_l}q_m q_n. \tag{4.17}$$

For $2s > 3/4$ (or $2s = (3/4)^+$), then for any $k \in \mathbb{Z}_+$

$$(Jq)_k = \textstyle\sum_{l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_l}q_m q_n \tag{4.18}$$

is absolutely convergent.

Proof. Note that

$$\sum_{l=0}^{N_1}\sum_{m=0}^{N_2}\sum_{l=0}^{N_3}C_{klmn}\overline{q_l}q_mq_n = \int_{-\infty}^{+\infty}h_k(x)\overline{u}_{N_1}u_{N_2}u_{N_3}dx = \left(\overline{u}_{N_1}u_{N_2}u_{N_3}\right)_k, \quad (4.19)$$

where $u_{N_j} = \sum_{l=0}^{N_j} q_l h_l(x)$. By the imbedding theorem, $X^{2s}(\mathbb{R}) \hookrightarrow L^3(\mathbb{R})$, we have $X^{2s}(\mathbb{R})\cdot X^{2s}(\mathbb{R})\cdot X^{2s}(\mathbb{R}) \subset L^1(\mathbb{R})$ and

$$\begin{aligned}
\left|(\overline{u}_{N_1}u_{N_2}u_{N_3})_k - (\overline{u}uu)_k\right| &\le \|h_k(x)\|_\infty \big(\|\overline{u}-\overline{u}_{N_1}\|_{L^3}\|u_{N_2}\|_{L^3}\|u_{N_3}\|_{L^3} \\
&\quad + \|\overline{u}_{N_1}\|_{L^3}\|u_{N_2}-u\|_{L^3}\|u_{N_3}\|_{L^3} + \|\overline{u}_{N_1}\|_{L^3}\|u_{N_2}\|_{L^3}\|u_{N_3}-u\|_{L^3}\big) \\
&\lesssim \langle k\rangle^{-1/12}\big(\|\overline{u}-\overline{u}_{N_1}\|_{X^{2s}}\|u\|_{X^{2s}}\|u\|_{X^{2s}} \\
&\quad + \|\overline{u}\|_{X^{2s}}\|u_{N_2}-u\|_{X^{2s}}\|u\|_{X^{2s}} + \|\overline{u}\|_{X^{2s}}\|u\|_{X^{2s}}\|u_{N_3}-u\|_{X^{2s}}\big),
\end{aligned}$$

which implies the equality (4.17).

When $2s=(3/4)^+$, by recalling that $|C_{klmn}| \lesssim \langle k\rangle^{(-1/8)+\varepsilon}\langle l\rangle^{(-1/8)+\epsilon}\langle m\rangle^{(-1/8)+\varepsilon}\langle n\rangle^{(-1/8)+\varepsilon}$, we have

$$\begin{aligned}
\sum_{l,m,n\in\mathbb{Z}_+}\left|C_{klmn}\overline{q_l}q_mq_n\right| &\lesssim \langle k\rangle^{(-1/8)+\epsilon}\left(\sum_{l=0}^{+\infty}\langle l\rangle^{(-1/8)+\epsilon}|q_l|\right)^3 \\
&\lesssim \langle k\rangle^{(-1/8)+\epsilon}\left(\sum_{l=0}^{+\infty}\langle l\rangle^{2s}|q_l|^2\right)^3\left(\sum_{l=0}^{+\infty}\langle l\rangle^{2(-s-1/8+\epsilon)}\right)^3.
\end{aligned}$$

Since $2s>3/4$ and $\epsilon>0$ is arbitrarily small, we can get the desired result. □

Remark 4.2.5. *In fact, we can get that when $2s \ge 1/6$ the operator J is a bounded operator from $X^{2s}(\mathbb{R})$ to $X^{(-2/3)^-}(\mathbb{R})$. This is due to these facts that if $u\in X^{2s_1}(\mathbb{R})$, $v\in X^{2s_2}(\mathbb{R})$ and $0\le 2s_1, 2s_2, 2s_1+2s_2 < 1/2$, then $uv\in X^{2s}(\mathbb{R})$ for $2s = \min(2s_1, 2s_2, 2s_1+2s_2-1/2)$; if $u\in X^{2\sigma}(\mathbb{R})$, $v\in X^{-2\sigma}(\mathbb{R})$ for $\sigma>0$, then $uv\in X^{(-2\sigma-1/2)^-}(\mathbb{R})$. So $|u|^2\in X^{1/6}(\mathbb{R})\cdot X^{1/6}(\mathbb{R})\subset X^{-1/6}(\mathbb{R})$ and further $|u|^2u\in X^{(-2/3)^-}(\mathbb{R})$. All the mappings (multiplications) arising here are continuous.*

One advantage to write the equation (4.1) in the q coordinate is that this method enables us to study the behavior of each mode of the solution, and thus many functions of these modes with some physics or mathematics meaning such as mass, energy and X^{2s} norm. The first thing we want to point out is that although the solution $q(t)$ can only be expected as a continuous curve in a given space $l_2^s(\mathbb{Z}_+)$, its k-th mode function $q_k(t)$ will be at least continuously differentiable.

It is not hard to deduce from the equation (4.15) that the k-th mode function $q_k(t)$ should satisfy the following equation

$$-i\dot{q}_k = \omega_k q_k + g\,(Jq)_k\,, \quad \forall k \in \mathbb{Z}_+. \tag{4.20}$$

Note that Jq is well defined for all $q \in l_2^s(\mathbb{Z}_+)$ with $2s \geq 1/6$, then it is easy to know that if $q(t)$ is a continuous solution curve to the equation (4.15) in this space, the function $q_k(t)$ must have continuous derivative $i(\omega_k q_k + g\,(Jq)_k)$. So we have proved the following proposition.

Proposition 4.2.6. *For $2s \geq 1/6$, if $q(t) \in C\left(I, l_2^s(\mathbb{Z}_+)\right)$ is a solution to the equation (4.15), then for any $k \in \mathbb{Z}_+$ its k-th mode function $q_k(t)$ is at least a C^1 function on the interval I.*

Since each q_k is continuous differentiable, it is natural to expect those functions depending on these $q\,(t)$ are also differentiable if $q(t)$ is regular enough. In particular, we can obtain such two conservation laws.

Proposition 4.2.7. *For $2s \geq 1/4$, if $q(t) \in C\left(I, l_2^s(\mathbb{Z}_+)\right)$ is a solution to the equation (4.15), then it conserves the mass function $M(t) = \sum_{k=0}^{+\infty} |q_k(t)|^2$.*

For $2s \geq 1$, if $q(t) \in C\left(I, l_2^s(\mathbb{Z}_+)\right)$ is a solution to the equation (4.15), then it *conserves the energy function*

$$E(t) = \sum_{k=0}^{+\infty}\omega_k \left|q_k(t)\right|^2 + \frac{g}{2}\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_k(t)q_l(t)}q_m(t)q_n(t). \tag{4.21}$$

Proof. For any $k \in \mathbb{Z}_+$, we have

$$\begin{aligned}\frac{d}{dt}\left|q_k(t)\right|^2 &= 2\,\mathrm{Re}\left(i\left(\omega_k q_k(t) + g\,(Jq)_k\,(t)\right)\overline{q_k(t)}\right)\\ &= 2\,\mathrm{Im}\,\omega_k\left|q_k(t)\right|^2 + 2g\,\mathrm{Im}\,(Jq)_k\,(t)\overline{q_k(t)}\\ &= 2g\,\mathrm{Im}\,(Jq)_k\,(t)\overline{q_k(t)}.\end{aligned}$$

When $2s \geq 1/4$, we claim that $Jq(t) \in C\left(I, l_2^{-1/8}(\mathbb{Z}_+)\right)$.

Let $\psi(x,t) = \sum_{k=0}^{+\infty}q_k(t)h_k(x)$, then it belongs to $C\left(I, X^{2s}(\mathbb{R})\right)$. Due to the following facts: $X^{1/4}(\mathbb{R}) \hookrightarrow L^4(\mathbb{R})$ and

$$L^4(\mathbb{R})\cdot L^4(\mathbb{R})\cdot L^4(\mathbb{R}) \subset L^{4/3}(\mathbb{R}) \subset X^{-1/4}(\mathbb{R}),$$

our claim can be easily verified. Thus the summation $\sum_{k=0}^{+\infty}(Jq)_k\,(t)\overline{q_k(t)}$ will be locally uniformly, absolutely convergent for time t. So we have the following equality

$$\begin{aligned}\frac{d}{dt}M(t) &= \frac{d}{dt}\sum_{k=0}^{+\infty}\left|q_k(t)\right|^2\\ &= 2g\,\mathrm{Im}\sum_{k=0}^{+\infty}(Jq)_k\,(t)\overline{q_k(t)}\\ &= 2g\,\mathrm{Im}\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_k(t)q_l(t)}q_m(t)q_n(t)\\ &= 2g\,\mathrm{Im}\int_{-\infty}^{+\infty}\left|\psi(x,t)\right|^4 dx = 0,\end{aligned}$$

which implies the conservation law of mass.

As to the energy function $E(t)$ in the form of equality (4.21), it is well defined for any $q(t) \in C\left(I, l_2^s(\mathbb{Z}_+)\right)$ with $2s \geq 1$. To get the conservation law of energy, it suffices to prove that $E(t)$ is continuously differentiable and its derivative is always zero. Let us see how this comes about.

For any $k \in \mathbb{Z}_+$, we have

$$\frac{d}{dt}\omega_k \left|q_k(t)\right|^2 = 2g \operatorname{Im} (Jq)_k (t)\overline{(I_\omega q)_k (t)}.$$

Since the space $X^1(\mathbb{R})$ is an algebra under multiplication, we can have $Jq \in C\left(I, l_2^{1/2}(\mathbb{Z}_+)\right)$ and $I_\omega q \in C\left(I, l_2^{-1/2}(\mathbb{Z}_+)\right)$. Then it is true that

$$\frac{d}{dt}\sum_{k=0}^{+\infty}\omega_k \left|q_k(t)\right|^2 = 2g\sum_{k=0}^{+\infty} \operatorname{Im} (Jq)_k (t)\overline{(I_\omega q)_k (t)}. \tag{4.22}$$

For the other part, let $k \in \mathbb{Z}_+$ fixed, then we get

$$\begin{aligned}
&\frac{d}{dt}\frac{g}{2}\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_k(t)q_l(t)}q_m(t)q_n(t) \\
&= g\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\left(\overline{\frac{d}{dt}q_k(t)}\overline{q_l(t)}q_m(t)q_n(t) + \overline{q_k(t)q_l(t)}\frac{d}{dt}q_m(t)q_n(t)\right).
\end{aligned}$$

From the symmetry of the tensor C_{klmn}, it can be continued as

$$\begin{aligned}
&\frac{d}{dt}\frac{g}{2}\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_k(t)q_l(t)}q_m(t)q_n(t) \\
&= 2g\sum_{k=0}^{+\infty} \operatorname{Re} \overline{i\left(\omega_k q_k(t) + g\,(Jq)_k\,(t)\right)}\,(Jq)_k \\
&= -2g\sum_{k=0}^{+\infty} \operatorname{Im} (Jq)_k (t)\overline{(I_\omega q)_k (t)} - 2g^2\sum_{k=0}^{+\infty} \operatorname{Im} \overline{(Jq)_k}\,(Jq)_k \\
&= -2g\sum_{k=0}^{+\infty} \operatorname{Im} (Jq)_k (t)\overline{(I_\omega q)_k (t)}.
\end{aligned}$$

Together with the equality (4.22), we can conclude that the function $E(t)$ always have zero derivative, which implies the conservation law of the energy. □

4.3 Birkhoff normal form in the case of $5/4 > 2s > 1/2$

In this section, we will adapt the equation (4.1) into the scheme of the infinite dimensional Hamiltonian system, then many concepts appeared in section 4.1 will be generalized into the infinite dimensional cases. After these jobs are done, the Birkhoff normal form theorem corresponding to this system will be proved for the case of $2s > 1/2$. Now let us begin with the writing of the PDE into the Hamiltonian form.

The Hamiltonian of the nonlinear Schrödinger equation (4.1) is

$$H(\psi, \overline{\psi}) = \int_{-\infty}^{+\infty} A\psi \cdot \overline{\psi} dx + \frac{g}{2} \int_{-\infty}^{+\infty} |\psi|^4 \, dx.$$

One observation is that the equation (4.1) can be written in the Hamiltonian form

$$\psi_t = i \frac{\partial H}{\partial \overline{\psi}}, \tag{4.23}$$

where $\frac{\partial H}{\partial \overline{\psi}}$ is the gradient of H with respect to $\overline{\psi}$.

We rewrite H as a Hamiltonian in infinitely many coordinates by making the ansatz

$$\psi(x, t) = \sum_{j \geq 0} q_j(t) h_j(x).$$

Let us take the coordinates from the Hilbert space $l_2^s(\mathbb{Z}_+)$ of all complex-valued sequences $q = (q_0, q_1, \cdots)$ with

$$\|q\|_{l_2^s}^2 = \sum_{j \geq 0} |q_j|^2 \langle j \rangle^{2s} < +\infty.$$

In this way we obtain the Hamiltonian on the phase space $\wp_s \equiv \wp_s(\mathbb{R}) := l_2^s(\mathbb{Z}_+; \mathbb{C})$ (recall that all notations of $l_2^s(\mathbb{Z}_+)$ in this thesis is in fact $l_2^s(\mathbb{Z}_+; \mathbb{C})$) with only real values

$$H(q,\overline{q}) = H_2 + H_4$$
$$= \sum_{j\geq 0}\omega_j\left|q_j\right|^2 + \frac{g}{2}\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{q_k}\overline{q_l}q_m q_n, \tag{4.24}$$

where $\omega_j = j + 1/2$. In general, we will say that a function F defined in the variable $(q,\overline{q})$ is **real** when $F(q,\overline{q})$ is always real.

Its equation of motion is

$$q_t = \imath\frac{\partial H}{\partial\overline{q}}. \tag{4.25}$$

In particular, for each component q_j it reads as

$$\frac{d}{dt}q_j = \imath\frac{\partial H}{\partial\overline{q_j}}. \tag{4.26}$$

This is a classical Hamiltonian equation of motion written in complex notation.

The quadratic term H_2 describes the linear integrable Schrödinger equation and gives rise to a linear Hamiltonian vector field which is unbounded of order 2. The fourth order term H_4 is not integrable, but gives rise to a bounded vector field of order 0. Of course, note that these Hamiltonians and their derivatives are well defined only for those points in the phase space with sufficient regularities, so we need first specify what are the regularity conditions.

4.3.1 Regularity of the Hamiltonian functions

Definition 4.3.1. *Let E and F be Banach spaces on a field $\Bbbk$, and $U \subset E$ be an open subset of E. It is said that a function $f : U \to F$ is **Fréchet differentiable** at $x \in U$ if there exists a linear operator $A_x \in \mathcal{L}(E,F)$ such that*

$\|f(x+y)-f(x)-A_x(y)\|_F = o(\|y\|_E)$. *It is said that the mapping* $f : U \to F$ *is a* ***continuously differentiable function****, or* C^1 ***function****, if it is Fréchet differentiable at each point in* U*, and the mapping* $f' : U \to \mathcal{L}(E, F)$ *defined by* $x \longmapsto A_x$ *is continuous.*

This definition can be generalized into higher order derivative cases.

Definition 4.3.2. *For* $n \in \mathbb{Z}_+$ *it is said that the mapping* $f : U \to F$ *is a* C^{n+1} ***function****, if it is a* C^n ***function*** *in* U*, and the mapping*

$$f^{(n+1)} : U \to \mathcal{L}\left(\underbrace{E, (\cdots (E, \mathcal{L}(E, F)))}_{n \text{ times } E}\right) \cong \mathcal{L}(\underbrace{E \times \cdots \times E}_{n \text{ times}}, F) = \mathcal{L}^n(E, F) \text{ defined by } x \longmapsto f \tag{4.27}$$

is continuous. In particular, it is said that $f : U \to F$ *is a* C^0 ***function*** *if the mapping is continuous; and it is a* C^∞ *function if the mapping is* C^n *for every* $n \in \mathbb{Z}_+$*.*

The Fréchet derivatives are a natural generalization of the conception of derivatives of the real valued functions on $\mathbb{R}^m$, and they have many basic properties familiar to us. For example, they satisfy the Fundamental theorem of calculus (the Newton-Leibniz Formula) and the chain rule for taking derivatives. Here without proof we want to mention the following two properties.

Proposition 4.3.3. *Suppose that the line segment between* $x \in U$ *and* $x + h \in U$ *lies entirely within* U*. If* $f : E \to F$ *is* C^k *then*

$$f(x+h) = f(x) + f'(x)\,(h) + \frac{1}{2!} f^{(2)}(x)\,(h, h) + \cdots + \frac{1}{(k-1)!} f^{(k-1)}(x)\,(h, h, \cdots, h) + R_k, \tag{4.28}$$

where the remainder term is given by

$$R_k(x)\{h\} = \frac{1}{(k-1)!}\int_0^1 (1-t)^{k-1} f^{(k)}(h, h, \cdots h)\, dt. \tag{4.29}$$

Proposition 4.3.4. *If the function f is C^k, then*

$$f^{(k)}(x)(h_1, \cdots, h_k) = f^{(k)}\left(h_{\sigma(1)}, \cdots, h_{\sigma(k)}\right)$$

for every permutation σ of $\{1, 2, ..., k\}$.

Now we will prove that those Hamiltonian functions defined above are in fact smooth (C^∞) on the spaces $X^{2s}(\mathbb{R}) \cong l_2^s(\mathbb{Z}_+)$.

Proposition 4.3.5. *Consider the Hamiltonian function $H = H_2 + H_4$ as the mapping from real vector space $X^{2s}(\mathbb{R})$ ($l_2^s(\mathbb{Z}_+)$) to the space of real numbers, then it is C^∞. in particular, H_2 is C^∞ on the space $X^{2s}(\mathbb{R})$ for $2s \geq 1$; H_4 is C^∞ on the space $X^{2s}(\mathbb{R})$ ($l_2^s(\mathbb{Z}_+)$) for $2s \geq 1/4$.*

Proof. First, we claim that the function H_2 is well defined on the space $X^{2s}(\mathbb{R})$ for $2s \geq 1$; and the function H_4 is well defined on the space $X^{2s}(\mathbb{R})$ for $2s \geq 1/4$. The former claim comes from the fact that $H_2(u) = \|u\|_{X^1}^2$. The latter one is true due to the fact that $X^{2s}(\mathbb{R})$ with $2s \geq 1/4$ can be continuously embedded in the space $L^4(\mathbb{R})$ and thus $|u|^2 u \in L^{4/3}(\mathbb{R}) \subset X^{-1/4}(\mathbb{R})$. So we have for $u = \sum_{j \geq 0} q_j h_j(x)$

$$\begin{aligned} 2H_4(u) &= \left| \sum_{k \in \mathbb{Z}_+} \overline{q_k} \, (Jq)_k \right| \\ &\lesssim \|q\|_{X^{2s}} \|Jq\|_{X^{-1/4}} \lesssim \|q\|_{X^{2s}}^4 . \end{aligned} \tag{4.30}$$

Meanwhile, it can be deduced that

$$2H_4(u) = \lim_{N_1, N_2, N_3, N_4 \to +\infty} \sum_{k=0}^{N_1} \sum_{l=0}^{N_2} \sum_{m=0}^{N_3} \sum_{n=0}^{N_4} C_{klmn} \overline{q_k}\,\overline{q_l} q_m q_n. \tag{4.31}$$

To prove the functions are smooth, we need to determine what are the Fréchet derivatives of them. For the function H_2, we have

$$\begin{aligned} H_2(u+\delta u) &= \sum_{k\in\mathbb{Z}_+} \omega_k \left(q_k + \delta q_k\right)\left(\overline{q_k} + \overline{\delta q_k}\right) \\ &= H_2(u) + \sum_{k\in\mathbb{Z}_+} \left(\omega_k q_k \overline{\delta q_k} + \omega_k \delta q_k \overline{q_k}\right) + \sum_{k\in\mathbb{Z}_+} \omega_k \delta q_k \overline{\delta q_k}, \end{aligned}$$

which implies that $H_2'(u)(\delta u) = \langle Au, \delta\overline{u}\rangle + \langle A\overline{u}, \delta u\rangle$. Since the operator A continuously maps $X^{2s}(\mathbb{R})$ onto $X^{2s-2}(\mathbb{R}) \cong \mathcal{L}(X^{2-2s}(\mathbb{R}), \mathbb{C})$ and $2s \geq 1$, it can be deduced that the first order derivative $H_2'(u) \in \mathcal{L}(X^{2s}(\mathbb{R}), \mathbb{R})$ and continuously depends on the function u. Further we have

$$\begin{aligned} \left(H_2'(u+\delta u^{(2)}) - H_2'(u)\right)(\delta u^{(1)}) &= \sum_{k\in\mathbb{Z}_+} \left(\omega_k \left(q_k + \delta q_k^{(2)} - q_k\right) \overline{\delta q_k^{(1)}} + \omega_k \delta q_k^{(1)} \left(\overline{q_k} + \overline{q_k^{(2)}} - \overline{q_k}\right)\right) \\ &= \sum_{k\in\mathbb{Z}_+} \left(\omega_k \delta q_k^{(2)} \overline{\delta q_k^{(1)}} + \omega_k \delta q_k^{(1)} \overline{q_k^{(2)}}\right), \end{aligned}$$

which implies $H_2''(u)(\delta u^{(1)}, \delta u^{(2)}) \equiv \left\langle A\delta u^{(1)}, \overline{\delta u^{(2)}}\right\rangle + \left\langle \overline{A\delta u^{(1)}}, \delta u^{(2)}\right\rangle$. It obviously belongs to $\mathcal{L}^2(X^{2s}(\mathbb{R}), \mathbb{R})$ and doesn't depend on the function u. Consequently, all the derivatives of higher orders of the function H_2 are always zero. Next let us see the case of the function H_4.

We have

$$\begin{aligned} 2H_4(u+\delta u) &= \sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn} \left(\overline{q_k} + \overline{\delta q_k}\right)\left(\overline{q_l} + \overline{\delta q_l}\right)\left(\overline{q_m} + \overline{\delta q_m}\right)\left(\overline{q_n} + \overline{\delta q_n}\right) \\ &= \sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn} \overline{\xi_k \xi_l} \xi_m \xi_n, \end{aligned} \tag{4.32}$$

where ξ is either q or δq. Let us count the numbers of the appearance of δq in the term $\overline{\xi_k \xi_l}\xi_m\xi_n$. If it is zero, then those terms give $H_4(u)$; if it is one, then they form the linear part of the difference of $H_4(u+\delta u) - H_4(u)$. Then we know that $2H_4'(u)$ is

$$2H_4'(u)(\delta u) = \sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\left(\overline{\delta q_k}\overline{q_l}q_mq_n + \overline{q_k}\overline{\delta q_l}q_mq_n + \overline{q_k}\overline{q_l}\delta q_mq_n + \overline{q_k}\overline{q_l}q_m\delta q_n\right). \tag{4.33}$$

By continuing in this way, we can get that

$$2H_4^{(j)}(u)(\delta u^{(1)},\cdots,\delta u^{(j)}) = \sum_{\overline{\xi_k\xi_l}\xi_m\xi_n\in W_j} C_{klmn}\overline{\xi_k\xi_l}\xi_m\xi_n \quad (j=1,2,3,4),$$

where ξ is q or $\delta q^{(1)}$ or $\cdots$ or $\delta q^{(j)}$ and W_j = {all the terms $\overline{\xi_k\xi_l}\xi_m\xi_n$ that $\delta q^{(1)}$, $\delta q^{(2)}, \cdots$,and $\delta q^{(j)}$ appear for exactly one time}. Due to the inequality of (4.30), it is easy to verify that all these Fréchet derivatives are bounded multilinear forms on $X^{2s}(\mathbb{R})$ with $2s \geq 1/4$ and continuously depend on the function $u \in X^{2s}(\mathbb{R})$. Meanwhile, the derivative of fourth order can be written as

$$2H_4^{(4)}(u)(\delta u^{(1)},\cdots,\delta u^{(4)}) = \textstyle\sum_{\sigma\in S_4}\sum_{k,l,m,n\in\mathbb{Z}_+} C_{klmn}\overline{\delta u_k^{(\sigma 1)}\delta u_l^{(\sigma 2)}}\delta u_m^{(\sigma 3)}\delta u_n^{(\sigma 4)},$$

where σ runs over all the permutations on $\{1,2,3,4\}$, that is the permutation group S_4. This derivative doesn't depend on the function u, thus all higher order derivatives are zero operator. So we have completed the proof. □

In fact these Hamiltonians have better smoothness property: they are real analytic. It is natural to have this property considering all these Hamiltonians are real valued, continuous polynomials on the phase space. For the sake of the completeness of this thesis, we also provide the definition of the real analyticity of a mapping [B-D][K-M1].

Definition 4.3.6. *A curve in a sequentially complete locally convex space E over the field $\mathbb{C}$ as $f:\mathbb{R}\to E$ is called* ***(weakly) real analytic*** *if $u\circ f \in A(\mathbb{R})$, the* *space of scalar-valued real analytic functions, for every $u \in E'$, and it is denoted as $f \in A(\mathbb{R}, E)$.*

Definition 4.3.7. *A map $f : \mathbb{R} \to E$ is called **topologically real analytic**, and it is denoted as $f \in A_t(\mathbb{R}, E)$, if for every $t \in \mathbb{R}$ there are $\varepsilon > 0$ and $a_i \in E$ such that $f(x) = \sum_{j=0}^{+\infty} a_j (x-t)^j$ for all $x \in (t-\varepsilon, t+\varepsilon)$ and the series converges in E.*

When the space E is a Banach space, then these two concepts are equivalent to each other. We are interested in the description of the real analyticity of a mapping from one space to another space. Here is a definition from the paper [K-M1].

Definition 4.3.8. *Let E, F be Banach space over the field $\mathbb{C}$, and U be an open set of E. A mapping $f : U \to F$ is called real analytic if it maps smooth curves to smooth curves and real analytic curves to real analytic curves.*

According to the theorem 3.4 in the paper [K-M1], the mapping $f : U \to F$ in this setting is real analytic if and only if it is smooth and is real analytic along each affine line in E. In particular, multilinear mappings are real analytic if and only if they are bounded. A more dedicated description on the real analyticity of mappings between convex topological spaces can be found in the paper [K-M2]. So the results in the proposition (4.3.5) are not only true for the smoothness but also true for real analyticity. In other words, H_2 is real analytic on the space $X^{2s}(\mathbb{R})$ for $2s \geq 1$ and H_4 is real analytic on the space $X^{2s}(\mathbb{R})$ for $2s \geq 1/4$.

4.3.2 Symplectic transformations

After obtaining the PDE in the form of a Hamiltonian system, we want to investigate how the Hamiltonian looks like if the coordinate is changed by a symplectic transformation, especially one induced by Lie transform. To do this, a basic concept known as Poisson bracket should be generalized into the case of infinite dimension. Note that the definition of the Poisson bracket in the formula (4.6) can be rewritten as

$$\{F,G\} = i\sum_{j=1}^{n}\left(\frac{\partial F}{\partial z_j}\frac{\partial G}{\partial \overline{z_j}} - \frac{\partial F}{\partial \overline{z_j}}\frac{\partial G}{\partial z_j}\right),$$

if we introducing the new complex variables $z_j = (x_j - iy_j)/\sqrt{2}$. So it is a natural thing to generalize the concept into the infinite dimensional case in this way

$$\{F,G\} = i\sum_{j\geq 0}\left(\frac{\partial F}{\partial q_j}\frac{\partial G}{\partial \overline{q_j}} - \frac{\partial F}{\partial \overline{q_j}}\frac{\partial G}{\partial q_j}\right). \tag{4.34}$$

When F, G are defined on an open set U of the phase space $\wp_s = l_2^s(\mathbb{Z}_+)$ such that $F, G \in C^1(U, \mathbb{R})$ and $X_F \in C(U, \wp_s)$, then the formula (4.34) is a well defined real valued function and the Poisson bracket $\{F, G\}$ is continuous on the set U. In this thesis, we are particularly interested in the following class of Hamiltonian functions.

Definition 4.3.9. *Let $s \geq 0$, we denote by $\aleph^s$ the space of real valued functions F defined on an open set U of the phase space $\wp_s$ and satisfying*

$$F \in C^\infty(U, \mathbb{R}) \text{ and } X_F \in C^\infty(U, \wp_s).$$

It is not hard to verify that if F in $\aleph^s$ is defined on an open set $U \subset \wp_s$ and $G \in C^\infty(U, \mathbb{R})$, then $\{F, G\}$ is also a smooth real valued function on U. In some cases, if G is also in $\aleph^s$, then $\{F, G\}$ is in $\aleph^s$ too, and it depends on F and G continuously. We are particularly interested in the case when F and G are both continuous homogeneous polynomials on the space $l_2^s(\mathbb{Z}_+)$ $(X^{2s}(\mathbb{R}))$, which is in the form of

$$\sum_{j\in\mathbb{Z}_+^{2n}} C_{j_1 j_2 \cdots j_{2n}} q_{j_1}\overline{q_{j_2}}\cdots\overline{q_{j_{2n}}}. \tag{4.35}$$

We can always assume that every coefficient $C_{j_1j_2\cdots j_{2n}}$ is symmetric under any permutation on the set $\{j_1, j_2, \cdots, j_{2n}\}$ keeping the subsets $\{j_1, j_3, \cdots, j_{2n-1}\}$ and $\{j_2, j_4, \cdots, j_{2n}\}$ invariant, since all these permutations form a group acting on $j \in \mathbb{Z}_+^{2n}$ and $C_{j_1j_2\cdots j_{2n}}$ can be redefined as the mean value of the orbit of the group action. Meanwhile, there is a natural way to define a norm for this kind of continuous homogeneous polynomials

$$|||F||| = \sup_{\forall j, \|q_j\|_{l_2^s}=1} |F(q_1, q_2, \cdots, q_n)|, \tag{4.36}$$

and for F in $\aleph^s$

$$\left|\left|\left|\frac{\partial F}{\partial \overline{q}}\right|\right|\right| = \sup_{\forall j, \|q_j\|_{l_2^s}=1} \left\|\frac{\partial F}{\partial \overline{q}}(q_1, q_2, \cdots, q_{n-1})\right\|_{l_2^s}. \tag{4.37}$$

Proposition 4.3.10. *Assume that F, G are two real valued functions in the class of $\aleph^s$, and they are in the form of (4.35) respectively of order $2n$ and $2m$. Then their Poisson bracket is also a real valued function in the class of $\aleph^s$, which is in the form of (4.35) of order $2n + 2m - 2$ and continuously depends on the F and G*

$$|\{F, G\}| \leq 2\min(m, n)\, |||F|||\, |||G|||\, \|q\|_{l_2^s}^{2n+2m-2} \tag{4.38}$$

$$\left\|\frac{\partial}{\partial \overline{q}}\{F, G\}\right\|_{l_2^s} \leq 4\max(m, n) \left|\left|\left|\frac{\partial F}{\partial \overline{q}}\right|\right|\right| \left|\left|\left|\frac{\partial G}{\partial \overline{q}}\right|\right|\right| \|q\|_{l_2^s}^{2n+2m-3}. \tag{4.39}$$

Proof. Since we have

$$\frac{\partial F}{\partial q_l} = n \sum_{(j_2, j_3, \cdots, j_{2n}) \in \mathbb{Z}_+^{2n-1}} C_{l j_2 \cdots j_{2n}} \overline{q_{j_2}} \cdots \overline{q_{j_{2n}}},$$

$$\frac{\partial G}{\partial \overline{q_l}} = m \sum_{(k_1, k_2, \cdots, k_{2m-1}) \in \mathbb{Z}_+^{2m-1}} D_{k_1 k_2 \cdots k_{2m-1} l}\, q_{k_1} \overline{q_{k_2}} \cdots q_{k_{2m-1}},$$

then under the introduction of the notation $j' = (j_2, j_3, \cdots, j_{2n-1})$ and $k' = (k_2, k_3, \cdots, k_{2m-1})$ their Poisson bracket can be written as

$$\begin{aligned}\{F, G\} &= nmi \sum_{(k_1,k',j',j_{2n}) \in \mathbb{Z}_+^{2n+2m-2}} \sum_{l \in \mathbb{Z}_+} (D_{k_1 k' l} C_{l j' j_{2n}} q_{k_1} \overline{q_{k_2}} \cdots q_{k_{2m-1}} \overline{q_{j_2}} \cdots \overline{q_{j_{2n}}} \\ &\quad - nmi \sum_{(j_1,j',k',k_{2m}) \in \mathbb{Z}_+^{2n+2m-2}} \sum_{l \in \mathbb{Z}_+} D_{l k' k_{2m}} C_{j_1 j' l} \overline{q_{k_2}} \cdots q_{k_{2m-1}} \overline{q_{k_{2m}}} q_{j_1} \overline{q_{j_2}} \cdots q_{j_{2n-1}}) \\ &= nmi \sum_{(k_1,k',j',j_{2n}) \in \mathbb{Z}_+^{2n+2m-2}} E_{j_1 j' k' k_{2m}} q_{j_1} \overline{q_{j_2}} \cdots q_{j_{2n-1}} \overline{q_{k_2}} \cdots q_{k_{2m-1}} \overline{q_{k_{2m}}},\end{aligned}$$

where

$$E_{j_1 j' k' k_{2m}} = \sum_{l \in \mathbb{Z}_+} (D_{j_1 k' l} C_{l j' k_{2m}} - D_{l k' k_{2m}} C_{j_1 j' l}).$$

So we can conclude that $\{F, G\}$ is also a real valued continuous homogeneous polynomials on the space $l_2^s(\mathbb{Z}_+)$ $(X^{2s}(\mathbb{R}))$ in the form of.(4.35).

As to the estimate of the Poisson bracket, we can proceed like this. Without loss of generality, let us assume that $n \leq m$, then we can have

$$\{F, G\} = -2n \operatorname{Im} \sum_{l \in \mathbb{Z}_+} \sum_{(j_2,j_3,\cdots,j_{2n}) \in \mathbb{Z}_+^{2n-1}} C_{l j_2 \cdots j_{2n}} \frac{\partial G}{\partial \overline{q_l}} \overline{q_{j_2}} \cdots \overline{q_{j_{2n}}}.$$

It can be regarded as the value of multilinear operator $\frac{\partial F}{\partial \overline{q}}$ at point $(\frac{\partial G}{\partial \overline{q}}, q, \cdots, q) \in l_2^s(\mathbb{Z}_+)^n$, so the inequality (4.38) follows immediately. Similarly, by taking the partial derivatives, we get

$$\begin{aligned}&\frac{\partial}{\partial q_r}\{F, G\} \\ &= i \sum_{l \geq 0} \left(\left(\frac{\partial^2 F}{\partial q_r \partial q_l} \frac{\partial G}{\partial \overline{q_l}} + \frac{\partial F}{\partial q_l} \frac{\partial^2 G}{\partial q_r \partial \overline{q_l}} \right) - \left(\frac{\partial^2 F}{\partial q_r \partial \overline{q_l}} \frac{\partial G}{\partial q_l} + \frac{\partial F}{\partial \overline{q_l}} \frac{\partial^2 G}{\partial q_r \partial q_l} \right) \right). \qquad (4.40)\end{aligned}$$

The first term is

$$\sum_{l \geq 0} \frac{\partial^2 F}{\partial q_r \partial q_l} \frac{\partial G}{\partial \overline{q_l}} = n(n-1) \sum_{(l,j_2,j_3,\cdots,j_{2n-2},j_{2n}) \in \mathbb{Z}_+^{2n-1}} C_{l j_2 \cdots j_{2n-2} r j_{2n}} \overline{q_{j_2}} \cdots \overline{q_{j_{2n-2}}} \frac{\partial G}{\partial \overline{q_l}} \overline{q_{j_{2n}}},$$

and it can be regarded as $(n-1)$ times the value of the multilinear functional $\frac{\partial F}{\partial q}$ at the point of $(q, q, \cdots, \frac{\partial G}{\partial \overline{q}}) \in l_2^s(\mathbb{Z}_+)^{n-1}$. Note that F, G are in the class of functions of $\aleph^s$, and thus $\frac{\partial F}{\partial \overline{q}} = \overline{\frac{\partial F}{\partial q}}$ and $\frac{\partial G}{\partial \overline{q}}$ are both continuous multilinear functionals on the space $l_2^s(\mathbb{Z}_+)$, we can deduce that

$$\left\| \sum_{l \geq 0} \frac{\partial^2 F}{\partial q \partial q_l} \frac{\partial G}{\partial \overline{q_l}} \right\|_{l_2^s} \leq (n-1) \left|\left|\left| \frac{\partial F}{\partial \overline{q}} \right|\right|\right| \left|\left|\left| \frac{\partial G}{\partial \overline{q}} \right|\right|\right| \|q\|_{l_2^s}^{2n+2m-3}.$$

Repeat this analysis for every term in the equality (4.40), it can be deduced that $\frac{\partial}{\partial \overline{q}}\{F, G\}$ is a C^∞ mapping on the phase space $\wp_s$ and it satisfies

$$\left\| \frac{\partial}{\partial \overline{q}}\{F, G\} \right\|_{l_2^s} \leq 4 \max(m, n) \left|\left|\left| \frac{\partial F}{\partial \overline{q}} \right|\right|\right| \left|\left|\left| \frac{\partial G}{\partial \overline{q}} \right|\right|\right| \|q\|_{l_2^s}^{2n+2m-3}.$$

□

The concept of Poisson bracket can help much in investigating the change of a Hamiltonian system under symplectic transformation. Suppose that, G is a C^∞ real valued Hamiltonian functional of the class $\aleph^s$ on the space $l_2^s(\mathbb{Z}_+)$ $(s \geq 0)$, then the Hamiltonian system induced by this Hamiltonian function

$$\frac{d}{dt} q = i \frac{\partial G}{\partial \overline{q}}$$

will provide a flow mapping $\Phi(t, q)$, which depends smoothly on the parameter q and depends on the parameter t in this way: belonging to the class of $C^\infty(I, l_2^s(\mathbb{Z}_+))$ for some time interval I depending on the point q. If the Hamiltonian function G equals to zero at the origin point, then the time-1 flow mapping $\Phi(1, q)$ will be well defined for a neighbourhood of the origin point. This mapping is called as the Lie transformation generated by the functional G.

In additional, if F is a C^∞ real valued Hamiltonian functional on the same space, then like in the finite dimensional cases it is true that

$$\frac{d}{dt}F(\Phi(t,q)) = \{F, G\}(\Phi(t,q)). \tag{4.41}$$

Now for any given position q, the curve $F(\Phi(t,q))$ is smooth in the space $l_2^s(\mathbb{Z}_+)$. By applying the above equality repeatedly, the Taylor expansion series between time $t=0$ and $t=1$ provide us

$$\begin{aligned} F(\Phi(1,q)) &= F(q) + \{F,G\}(q) + \frac{1}{2!}\{\{F,G\},G\}(q) + \cdots \\ &+ \frac{1}{m!}F^{(m)}(q) + \frac{1}{m!}\int_0^1 (1-t)^m \{F^{(m)}, G\}(\Phi(t,q))dt, \end{aligned} \tag{4.42}$$

where $F^{(m+1)} = \{F^{(m)}, G\}$ and $F^{(0)} = F$.

As to our problem, the Hamiltonian function $H = H_2 + H_4$ is C^∞ and real valued on the space $l_2^s(\mathbb{Z}_+) \cong X^{2s}(\mathbb{R})$ with $2s \geq 1$. If a continuous homogeneous polynomial F is in the form of (4.35) and belongs to the class of $\aleph^s$, then under the Lie transformation Γ generated by this functional F, the Hamiltonian function H can be written as

$$\begin{aligned} H \circ \Gamma = H_2 + H_4 + \{H_2, F\} + \{H_4, F\} + \tfrac{1}{2!}\{\{H_2, F\}, F\} \\ + \tfrac{1}{2!}\{\{H_4, F\}, F\} + \cdots + \tfrac{1}{m!}H_2^{(m)} + \tfrac{1}{m!}H_4^{(m)} \\ + \tfrac{1}{m!}\int_0^1 (1-t)^m \left(H_2^{(m+1)} + H_4^{(m+1)}\right)(\Phi(t,\cdot))dt. \end{aligned} \tag{4.43}$$

In the above formula, each term before $\frac{1}{m!}H_4^{(m)}$ is a continuous homogeneous polynomial and its order can be computed out in a straightforward way.

For the sake of simplifying the function $H \circ \Gamma$, we choose

$$F_4(q) := \frac{gi}{2}\sum_{k,l,m,n\in\mathbb{Z}_+} \frac{C_{klmn}}{D(k+l-m-n)}\overline{q_k}\overline{q_l}q_m q_n, \tag{4.44}$$

where

$$C_{klmn} = \int_{-\infty}^{+\infty} h_k(x)h_l(x)h_m(x)h_n(x)dx,$$

$$D(k+l-m-n) = \begin{cases} k+l-m-n & \text{if} \quad k+l-m-n \neq 0 \\ \iota & \text{if} \quad k+l-m-n = 0. \end{cases} \tag{4.45}$$

Then plugging it into the formula (4.43), we can formally get

$$\begin{aligned} H \circ \Gamma &= H_2 + H_4 + \{H_2, F_4\} + R \\ &= H_2 + H_4 - \frac{g}{2} \sum_{k+l-m-n \neq 0} C_{klmn} \overline{q_k q_l} q_m q_n + R \\ &= H_2 + \frac{g}{2} \sum_{k+l-m-n=0} C_{klmn} \overline{q_k q_l} q_m q_n + R, \end{aligned} \tag{4.46}$$

where R includes all the terms in the right side of (4.43) except $H_2+H_4+\{H_2, F\}$. To make this formula meaningful rigorously, we need to investigate when the definition formula (4.44) is well defined. In doing so, the convolution operation of two sequences will arise and its properties must be studied. And then we will provide the proof of the main theorem.

4.3.3 Regularity of the convolution operation

We will study the regularity property of the convolution operation. Formally we always have the convolution of $p \in l_b^{s_1,2}$ and $q \in l_b^{s_2,2}$ as

$$(p * q)_l = \sum_{k \in \mathbb{Z}} p_k q_{l-k} = \sum_{k \in \mathbb{Z}} p_{l-k} q_k, \tag{4.47}$$

which suggests the convolution admits the commutative rule. The only problem that we need be careful on is the problem of convergence. It is a natural choice to ask $s_1 + s_2 \geq 0$ to guarantee the equality (4.47) well defined. Under this condition we have

$$\left|\sum_{k\in\mathbb{Z}} p_k q_{l-k}\right| \le \left(\sum_{k\in\mathbb{Z}} |p_k|^2 \langle k\rangle^{2s_1}\right)^{1/2} \left(\sum_{k\in\mathbb{Z}} |q_{l-k}|^2 \langle k\rangle^{2s_2}\right)^{1/2}$$
$$\le \|p\|_{l_b^{s_1,2}} \|q\|_{l_b^{s_2,2}} \langle l\rangle^{|s_2|},$$

which implies the summation is absolutely convergent and independent of the order of the p and q. Note that $l_b^{s_j,2} \cong H^{s_j}(T)$ $(j = 1,2)$ and $p * q$ corresponds to the product of two functions on $H^{s_j}(T)$, so it is understandable to get such a result on the regularity, which is similar to the case of $H^{s_j}(\mathbb{R})$.

Theorem 4.3.11. *If $p \in l_b^{s_1,2}$, $q \in l_b^{s_2,2}$ and $s_1 + s_2 \ge 0$, then it is true that $p * q$ in (4.47) is well defined and independent of the order of the p and q. Meanwhile $p * q \in l_b^{s,2}$, and there is a constant C depending on $s, s_1,$ and s_2 satisfying*

$$\|p * q\|_{l_b^{s,2}} \le C(s, s_1.s_2) \|p\|_{l_b^{s_1,2}} \|q\|_{l_b^{s_2,2}}, \tag{4.48}$$

where $s \le s_1$, $s \le s_2$ and $s \le s_1 + s_2 - 1/2$. When one of the s_j is equal to $1/2$, or $s_1 + s_2 = 0$, then the last inequality should be changed as strict inequality.

Here we would like to provide a proof based on the following method. First we write the convolution in the form of integral with respect to some kernel function and then the problem can be transferred into the estimate for certain L^2 norm of the kernel function. That estimate is very complicated if one need consider all the cases of the real numbers s_1 and s_2 like what is done in the textbook [Q-X-W]. We would like to mention that by using duality, one need only consider much simpler cases.

Let us write $\langle l\rangle^s (p * q)_l = \sum_{k\in\mathbb{Z}} p_k q_{l-k} = \sum_{k\in\mathbb{Z}} (p_k \langle k\rangle^{s_1}) (q_{l-k} \langle l-k\rangle^{s_2}) F(k,l)$, where

$$F(k,l) = \langle k\rangle^{-s_1} \langle l-k\rangle^{-s_2} \langle l\rangle^s. \tag{4.49}$$

Note that $p_k \langle k \rangle^{s_1}$ and $(q_{l-k} \langle l-k \rangle^{s_2})$ corresponding to elements in l^2, then the following lemma changes the problem into getting an estimate of the kernel function $F(k,l)$.

Lemma 4.3.12. *Let the operator T_F on the space $l^2 \times l^2$ defined as $T_F(p,q) = \sum_{k\in\mathbb{Z}} p_k q_{l-k} F(k,l)$, where the kernel function satisfies that*

$$\textstyle\sum_{k\in\mathbb{Z}} |F(k,l)|^2 \leq M^2 \quad M \text{ independent of } l, \text{ or} \tag{4.50}$$

$$\textstyle\sum_{l\in\mathbb{Z}} |F(k,l)|^2 \leq M^2 \quad M \text{ independent of } k, \tag{4.51}$$

Then it is true that $T_F(p,q) \in l^2$ and

$$\|T_F(p,q)\|_{l^2} \leq M \|p\|_{l^2} \|q\|_{l^2}.$$

Proof. If the inequality (4.50) is true, then the Schwartz inequality gives

$$|(T_F(p,q))_l|^2 \leq M^2 \textstyle\sum_{k\in\mathbb{Z}} |p_k q_{l-k}|^2,$$

which provides the result immediately. If the inequality (4.51) is true, then let us denote r as an arbitrary element in the space l^2, then

$$\begin{aligned} \left|\textstyle\sum_{l\in\mathbb{Z}} (T_F(p,q))_l r_l\right| &= \left|\textstyle\sum_{k\in\mathbb{Z}} p_k \left(\sum_{l\in\mathbb{Z}} q_{l-k} F(k,l) r_l\right)\right| \\ &\leq M \|p\|_{l^2} \|q\|_{l^2} \|r\|_{l^2}, \end{aligned}$$

which provides the desired result. □

So we only need to divide the kernel function $F(k,l)$ in several parts and for each of them get the estimate of the L^2 norm either on the parameter k or on the parameter l for all possible cases of s_1 and s_2. By duality, we can reduced the problems into simpler cases.

Lemma 4.3.13. *The theorem 4.3.11 is true if and only if it is true for the cases of* $0 \le s_2 \le s_1$.

Proof. It suffices to prove that the theorem is true in the case of $0 \le s_2 \le s_1$ are enough to deduce that the theorem is true for all the possible cases. Since the convolution $p * q$ is independent of the order of the p and q, so without loss of the generality, we can assume that $s_1 \ge s_2$. Notice that the condition $s_1 + s_2 \ge 0$ thus implies $s_1 \ge 0$, then only case we need treat on is that $s_2 < 0 \le s_1$. Here let us introduce the "pair operator" $\langle , \rangle$, which is defined between a locally convex vector space and its topological dual space. In particular, for $p \in l_b^{\sigma,2}$, $q \in l_b^{-\sigma,2}$, the pair operation gives

$$\langle p, q \rangle = \sum_{k \in \mathbb{Z}} p_k q_k.$$

Now if $s_1 > 1/2$, then it is true that $s \le \min(s_1, s_2, s_2 + (s_1 - 1/2)) = s_2$. Let us define the convolution in this way

$$\langle p * q, r \rangle := \langle q, \tau p * r \rangle, \tag{4.52}$$

where τ is the mapping such that $(\tau q)_k = q_{-k}$ and r is any element in the space $l_b^{-s,2}$. According to the assumption, it is true that $\tau p * r \in l_b^{-s_2,2}$ and $\|\tau p * r\|_{l_b^{-s_2,2}} \lesssim \|p\|_{l_b^{s_1,2}} \|r\|_{l_b^{-s,2}} \quad (s, s_1, s_2)$, since $\min(s_1, -s, -s + (s_1 - 1/2)) = -s \ge -s_2$. So the right side of the equality (4.52) is well defined for any $r \in l_b^{-s,2}$, which induces a well defined convolution $p * q \in l_b^{s,2}$ satisfying $\|p * q\|_{l_b^{s,2}} \lesssim \|p\|_{l_b^{s_1,2}} \|q\|_{l_b^{-s_2,2}} \quad (s, s_1, s_2)$. Since all the summations appearing in the equality (4.52) are absolutely convergent, it is easy to confirm that such a defined convolution $p * q$ is same as in the formerly defined convolution in (4.47).

If $0 \le s_1 < 1/2$, then it is true that $s \le \min(s_1, s_2, s_2+(s_1-1/2)) = s_2+(s_1-1/2)$ or $-s \ge -s_2+(1/2-s_1)$. According to the assumption, we have $\tau p * r \in l_b^{-s_2,2}$ and $\|\tau p * r\|_{l_b^{-s_2,2}} \lesssim \|p\|_{l_b^{s_1,2}} \|r\|_{l_b^{-s,2}} \quad (s, s_1, s_2)$, since $\min(s_1, -s, -s+(s_1-1/2)) = -s+(s_1-1/2) \ge -s_2$. Then the equality (4.52) is well defined for all $r \in l_b^{-s,2}$, which means $p * q \in l_b^{s,2}$ satisfying the estimate (4.48).

If $s_1 = 1/2$, then we can further consider it in two subcases. When $s_2 = -1/2$ and thus $s < s_2 + s_1 - 1/2 = -1/2$, it is true that the right side of equality (4.52) is still well defined since in this case $\tau p * r \subset l_b^{-s,2} * l_b^{1/2,2} \subset l_b^{1/2,2}$. So this equality induces a well defined function $p * q$ in the space $l_b^{-s,2}$. When $-s_1 < s_2 < 0$ and thus $s < s_2 + s_1 - 1/2 = s_2$, when can consider $p \in l_b^{s_1-\delta,2}$, $q \in l_b^{s_2-\delta,2}$ for $\delta < (s_2 - s)/2$ and $s_1 + s_2 - 2\delta \ge 0$. Therefore it is reduced to the case of $s_1 < 1/2$ and we can deduce that $p * q \in l_b^{s,2}$ satisfying the inequality (4.48). So we have analyzed all the possible cases and the lemma has thus been proved. □

We are now in a very good position to finally prove the theorem 4.3.11, which has been reduced to the estimate of certain L^2 norms of the kernel function in the case of $0 \le s_2 \le s_1$.

Proof. Let us divide the kernel function $F(k,l)$ according to the position of the points (k,l) into such three mutually disjointed parts: $\{\langle k\rangle \le \langle l\rangle/2\}$, $\{\langle l-k\rangle \le \langle l\rangle/2\}$ and $\{\langle k\rangle > \langle l\rangle/2, \langle l-k\rangle > \langle l\rangle/2\}$ and denote them respectively as $F_j(k,l), j = 1,2,3$. According to the lemma (4.3.12), it suffices to get estimates for these $F_j(k,l)$ in the form of inequality (4.50) and (4.51).

For the first part $F_1(k,l)$, since all the points are in $\{\langle k\rangle \le \langle l\rangle/2\}$, it can be deduced that $\langle k\rangle < \langle l/2\rangle$ and further $|k| < |l/2|$. Thus the term $\langle l-k\rangle^{-s_2}$ can be controlled between $\langle l/2\rangle^{-s_2}$ and $\langle 3l/2\rangle^{-s_2}$, that is to say, $\langle l-k\rangle^{-s_2} \approx \langle l\rangle^{-s_2} \quad (s_2)$. So we have the estimate

$$\sum_{k\in\mathbb{Z}} |F_1(k,l)|^2 = \sum_{k\in\mathbb{Z},\langle k\rangle\le\langle l\rangle/2} |F(k,l)|^2 \le C(s_2)\,\langle l\rangle^{2s-2s_2} \sum_{k\in\mathbb{Z},\langle k\rangle\le\langle l\rangle/2} \langle k\rangle^{-2s_1}. \qquad (4.53)$$

If $s_1 > 1/2$, then the above inequality can proceed as $\sum_{k\in\mathbb{Z}} |F_1(k,l)|^2 \le C(s_2)\,\langle l\rangle^{2s-2s_2} \sum_{k\in\mathbb{Z}} \langle k\rangle^{-2s_1} = O(1) \quad (s_1,s_2)$; if $s_1 = 1/2$, then it continues as $\sum_{k\in\mathbb{Z}} |F_1(k,l)|^2 \lesssim C(s_2)\,\langle l\rangle^{2s-2s_2} \ln\langle l\rangle = O(1) \quad (s,s_2)$ since in this case $s < s_2$; if $s_1 < 1/2$, then it continues as $\sum_{k\in\mathbb{Z}} |F_1(k,l)|^2 \lesssim C(s_2)\,\langle l\rangle^{2s-2s_2}\,\langle l\rangle^{-2s_1+1} = O(1) \quad (s_1,s_2)$. For the second part $F_2(k,l)$, by the parameter change we get

$$\begin{aligned}\sum_{k\in\mathbb{Z}} |F_2(k,l)|^2 &= \sum_{k\in\mathbb{Z},\langle l-k\rangle\le\langle l\rangle/2} \langle k\rangle^{-2s_1}\,\langle l-k\rangle^{-2s_2}\,\langle l\rangle^{2s} \\ &= \sum_{k'\in\mathbb{Z},\langle k'\rangle\le\langle l\rangle/2} \langle k'\rangle^{-2s_2}\,\langle l-k'\rangle^{-s_1}\,\langle l\rangle^{2s}.\end{aligned}$$

A similar analysis for the case of $F_1(k,l)$ is still true, where the only difference is that we need compare s_2 and $1/2$ at this time.

For the third part $F_3(k,l)$, we turn to the estimate for $\sum_{l\in\mathbb{Z}} |F_3(k,l)|^2$. It is easy to confirm the result in the theorem if $s < s_1 + s_2 - 1/2$, since in this case $\sum_{l\in\mathbb{Z}} |F_3(k,l)|^2 \lesssim \sum_{l\in\mathbb{Z}} \langle l\rangle^{-2s_1-2s_2+2s} = O(1) \quad (s,s_1,s_2)$. That means we have proved such cases like $s_1 \ge 1/2$ and $s_1 = s_2 = 0$. For the remainder cases ($0 \le s_2 \le s_1 < 1/2$ and $s_1 > 0$), since $\langle l-k\rangle^{-2s_2} \le C(s_2)\,\langle l\rangle^{-2s_2}$ and $s = s_1 + s_2 - 1/2$, we have

$$\begin{aligned}\sum_{l\in\mathbb{Z}} |F_3(k,l)|^2 &\le C(s_2)\,\langle k\rangle^{-2s_1} \sum_{l\in\mathbb{Z},\langle l\rangle<2\langle k\rangle} \langle l\rangle^{2s_1-1} \\ &\le C(s_1,s_2)\,\langle k\rangle^{-2s_1+2s_1} = O(1) \quad (s_1,s_2).\end{aligned}$$

Thus we have proved the theorem 4.3.11. $\square$

4.3.4 Proof of the main theorem 4.1.3 (in the case of $5/4 > 2s > 1/2$)

Now we are at a good position to prove the main theorem. Just like what we have done in the formal computation, we define a real valued function

$$F_4(q) = \frac{gi}{2} \sum_{k,l,m,n \in \mathbb{Z}_+} \frac{C_{klmn}}{D(k+l-m-n)} \overline{q_k q_l} q_m q_n, \tag{4.54}$$

where

$$D(k+l-m-n) = \begin{cases} k+l-m-n & \text{if} \quad k+l-m-n \neq 0 \\ i & \text{if} \quad k+l-m-n = 0. \end{cases}$$

Then we will see that this function is well defined on the space $l_2^s(\mathbb{Z}_+)$ with $5/4 > 2s > 1/2$, and thus its Lie transform will also exist in the same space.

Lemma 4.3.14. *For $5/4 > 2s > 1/2$, the Hamiltonian vector field X_{F_4} is real analytic, and it forms a map from some neighbourhood of the origin in X^{2s} ($l_2^s(\mathbb{Z}_+)$) into X^{2s} ($l_2^s(\mathbb{Z}_+)$), with*

$$\|X_{F_4}(q)\|_{l^s} = O\left(\|q\|_{l^s}^3\right) \quad (s). \tag{4.55}$$

Proof. We have

$$\begin{aligned} |(X_{F_4}(q))_k| &= \left| \sum_{l,m,n \in \mathbb{Z}_+} \frac{C_{klmn}}{D(k+l-m-n)} \overline{q_l} q_m q_n \right| \\ &\lesssim \sum_{l,m,n \in \mathbb{Z}_+} \frac{|C_{klmn}|}{\langle k+l-m-n \rangle} \left| \overline{q_l} q_m q_n \right|. \end{aligned} \tag{4.56}$$

Recall that we have $|C_{klmn}| \lesssim \langle k \rangle^{-1/8+\varepsilon} \langle l \rangle^{-1/8+\varepsilon} \langle m \rangle^{-1/8+\varepsilon} \langle n \rangle^{-1/8+\varepsilon}$ $(0 < \varepsilon \ll 1)$, the inequality above can be continued as

$$
\begin{aligned}
|(X_{F_4}(q))_k| &\lesssim \langle k\rangle^{-1/8+\varepsilon} \sum_{l,m,n\in\mathbb{Z}_+} \frac{\langle l\rangle^{-1/8+\varepsilon}\langle m\rangle^{-1/8+\varepsilon}\langle n\rangle^{-1/8+\varepsilon}}{\langle k+l-m-n\rangle} |\overline{q_l} q_m q_n| \\
&\lesssim \langle k\rangle^{-1/8+\varepsilon} \sum_{d\in\mathbb{Z}} \sum_{m+n-l=k-d} \frac{\langle l\rangle^{-1/8+\varepsilon}\langle m\rangle^{-1/8+\varepsilon}\langle n\rangle^{-1/8+\varepsilon}}{\langle k+l-m-n\rangle} |\overline{q_l} q_m q_n| .
\end{aligned}
$$

Let τ be the mapping from $l_2^s(\mathbb{Z}_+)$ to $l_b^{s,2}$ defined by

$$
(\tau q)_k = \begin{cases} q_{k-1} & \text{if} \quad k\in\mathbb{N} \\ \overline{q_0} & \text{if} \quad k=0 \\ \overline{q_{-k}} & \text{if} \quad -k\in\mathbb{N}. \end{cases} \tag{4.57}
$$

Obviously, this mapping τ is a bounded operator. So we have

$$
\begin{aligned}
|(X_{F_4}(q))_k| &\lesssim \langle k\rangle^{-1/8+\varepsilon} \sum_{d\in\mathbb{Z}} \sum_{m+1+n+1-l=k+2-d} \frac{1}{\langle d\rangle} \left|(\tau\widetilde{q})_{-l}(\tau\widetilde{q})_{m+1}(\tau\widetilde{q})_{n+1}\right| \\
&\lesssim \langle k\rangle^{-1/8+\varepsilon} \left(\tau\widetilde{q} * \tau\widetilde{q} * \tau\widetilde{q} * \frac{1}{\langle d\rangle}\right)_{k+2} \qquad \text{for } k\geq 0,
\end{aligned} \tag{4.58}
$$

where $\widetilde{q}_j = \langle j\rangle^{-1/8+\varepsilon}|q_j|$.

For $5/8-2\varepsilon > s > 3/8+2\varepsilon$, it is true that $\tau\widetilde{q}$ belongs to the space $l_b^{1/2+\varepsilon,2}$, which is an algebra under the convolution. Note that $\frac{1}{\langle d\rangle} \in l_b^{1/2-\varepsilon,2}$, we can get that $X_{F_4}(q) \in l_2^{5/8-2\varepsilon}(\mathbb{Z}_+)$ and

$$
\begin{aligned}
\|X_{F_4}(q)\|_{l_2^{5/8-2\varepsilon}} &\lesssim \|\tau\widetilde{q}\|^3_{l_b^{s+1/8-\varepsilon,2}} \left\|\frac{1}{\langle d\rangle}\right\|_{l_b^{1/2-\varepsilon,2}} \quad (s,\varepsilon) \\
&\lesssim \|q\|^3_{l_2^s} \quad (s,\varepsilon).
\end{aligned} \tag{4.59}
$$

Since $\varepsilon>0$ can be arbitrarily small, we can deduce that if $5/4 > 2s > 3/4$ then $X_{F_4}(q)$ is in the space $l_2^{(5/8)^-}$, which is in general a little more regular than q. In particular, the inequality (4.55) is satisfied.

For $1/4+5/2\varepsilon < s \leq 3/8$, it is true that $\tau\widetilde{q}$ belongs to the space $l_b^{1/8+s-\varepsilon,2}$, which can not be an algebra. But in this time we have $\tau\widetilde{q} * \tau\widetilde{q} * \tau\widetilde{q} \in l_2^{3s-5/8-3\varepsilon}(\mathbb{Z}_+)$

and further $\tau\widetilde{q}*\tau\widetilde{q}*\tau\widetilde{q}*1/\langle d\rangle \in l_2^{3s-5/8-4\varepsilon}(\mathbb{Z}_+)$. Since it is true that $3s-1/2-5\varepsilon > s$ when $s > 1/4+5/2\varepsilon$, the inequality (4.58) implies that

$$\begin{aligned}\|X_{F_4}(q)\|_{l_2^s} &\leq \|X_{F_4}(q)\|_{l_2^{3s-1/2-5\varepsilon}} \\ &\lesssim \|\tau\widetilde{q}\|^3_{l_b^{s+1/8-\varepsilon,2}} \left\|\frac{1}{\langle d\rangle}\right\|_{l_b^{1/2-\varepsilon,2}} \quad (s,\varepsilon) \\ &\lesssim \|q\|^3_{l_2^s} \quad (s,\varepsilon).\end{aligned}$$

Note that $\varepsilon > 0$ can be arbitrarily small, we can deduce that if $1/2 < 2s \leq 3/4$, then $X_{F_4}(q)$ is in the space $l_2^s(\mathbb{Z}_+)$ for all the q in the same space, and it satisfies the inequality (4.55). □

According to this lemma (4.3.14), the Hamiltonian function F_4 belongs to the class of $\aleph^s$ with $5/4 > 2s > 1/2$ and its associated Lie transformation Γ is well defined in a neighbourhood of the origin point in the space $l_2^s(\mathbb{Z}_+)$. Meanwhile, like the Hamiltonian vector field X_{F_4}, this transformation is also real analytic. Note that the Hamiltonian function associated to our PDE is regular on the space $l_2^s(\mathbb{Z}_+)$ with $2s \geq 1$, we can write the new Hamiltonian function $H\circ\Gamma$ on the space $l_2^s(\mathbb{Z}_+)$ with $5/4 > 2s \geq 1$ as

$$\begin{aligned}H\circ\Gamma &= H_2 + H_4 + \{H_2. F_4\} + \{H_4, F_4\} + \frac{1}{2!}\{\{H_2, F_4\}, F_4\} \\ &+ \frac{1}{2!}\{\{H_4, F_4\}, F_4\} + \cdots + \frac{1}{m!}H_2^{(m)} + \frac{1}{m!}H_4^{(m)} \\ &+ \frac{1}{m!}\int_0^1 (1-t)^m \left(H_2^{(m+1)} + H_4^{(m+1)}\right)(\Phi(t,\cdot))dt \\ &= H_2 + \frac{g}{2}\sum_{k+l-m-n=0} C_{klmn}\overline{q_k q_l} q_m q_n + R,\end{aligned}$$

where $\Phi(t,\cdot)$ is the flow associated to the Hamiltonian function F_4 and the remainder term R includes all the terms except $H_2 + H_4 + \{H_2, F_4\}$. We claim that the function R is in the class of $\aleph^s$ when s satisfies $5/4 > 2s \geq 1$.

By the proposition 4.3.10, we can easily see that $H_4^{(m)}$ are all in the class of $\aleph^s$. Note that

$$\{H_2, F_4\} = -H_4 + \frac{g}{2} \sum_{k+l-m-n=0} C_{klmn}\overline{q_k}\overline{q_l}q_m q_n$$

and

$$\begin{aligned} &\left| \frac{\partial}{\partial \overline{q_j}} \sum_{k+l-m-n=0} C_{klmn}\overline{q_k}\overline{q_l}q_m q_n \right| \\ &= 2 \left| \sum_{m+n-l=j} C_{jlmn}\overline{q_l}q_m q_n \right| \\ &\lesssim \langle j \rangle^{-1/8+\varepsilon} \left(\tau\widetilde{q} * \tau\widetilde{q} * \tau\widetilde{q} \right)_j , \end{aligned}$$

which implies the function $G(p) = \frac{g}{2}\sum_{k+l=m+n} C_{klmn}\overline{p_k}\overline{p_l}p_m p_n$ is also in the class of $\aleph^s$ and so does the function $H_2^{(m)}$ with $m \geq 1$. So we can use the inequalities (4.39)(4.38) to deduce that the remainder term R is a well defined function in the class of $\aleph^s$, which is at least of order 6 at the origin and satisfies

$$|| X_R(p) ||_{l_2^s} \leqslant C_s || p ||_{l_2^s}^5 .$$

As to the difference between the mapping Γ and the identity, we have

$$\Gamma(p) - p = \Phi(1, p) - \Phi(0, p) ,$$

where $\Phi(t, \cdot)$ denotes the flow associated the Hamiltonian function F_4. Then it can proceed as

$$\begin{aligned} || \Gamma(p) - p ||_{l_2^s} &\leq \int_0^1 || \frac{d}{dt}\Phi(t, p) ||_{l_2^s} dt \\ &\leq \int_0^1 || X_{F_4} \left(\Phi(t, p) \right) ||_{l_2^s} dt \\ &= O\left(||p||_{l_2^s}^3 \right) . \end{aligned}$$

So we have completed the proof of the main theorem in the case of $5/4 > 2s \geq 1$.

In fact the main theorem is also true in the case of $2s \geq 5/4$. The key fact needed there is that the Hamiltonian vector field X_{F_4} is also real analytic in that case. In the following section we will concentrate on proving this key fact in this new case, since all the other claims in the main theorem will follow by essentially the same technique.

4.4 Birkhoff normal form in the case of $2s \geqslant 5/4$

The essential task of this section is to prove the following lemma.

Lemma 4.4.1. *For $2s \geq 5/4$, the hamiltonian vector field X_{F_4} is real analytic, and it forms a map from some neighbourhood of the origin in X^{2s} ($l_2^s(\mathbb{Z}_+)$) into X^{2s} ($l_2^s(\mathbb{Z}_+)$), with*

$$\|X_{F_4}(q)\|_{l^s} = O\left(\|q\|_{l^s}^3\right) \quad (s). \tag{4.60}$$

To do this, we need a better understanding on the tensor C_{klmn}, especially when one of the indices is much bigger than all the others.

4.4.1 Proof of the lemma 4.4.1 in the case of $2s \geqslant 5/4$

After obtaining the estimate of the coefficient C_{klmn} as in the corollary 2.4.15, we are in a good position to prove the main lemma 4.4.1 and thus the Birkhoff normal form theorem for our PDE in the function space $l_2^s(\mathbb{Z}_+)$ ($X^{2s}(\mathbb{R})$) with $2s \geqslant 5/4$.

Proof. The Hamiltonian vector field X_{F_4} satisfies the estimate

$$\begin{aligned}
|(X_{F_4}(q))_k| &= \left| \sum_{l,m,n\in\mathbb{Z}_+} \frac{C_{klmn}}{D(k+l-m-n)} \overline{q_l} q_m q_n \right| \\
&\le \left| \sum_{\langle l\rangle\langle m\rangle\langle n\rangle<\langle k\rangle/M} \frac{C_{klmn}}{D(k+l-m-n)} \overline{q_l} q_m q_n \right| \\
&+ \left| \sum_{\langle l\rangle\langle m\rangle\langle n\rangle\ge\langle k\rangle/M} \frac{C_{klmn}}{D(k+l-m-n)} \overline{q_l} q_m q_n \right| \\
&= I + II, \qquad (4.61)
\end{aligned}$$

where M is a number to guarantee that the inequality $\langle l\rangle \langle m\rangle \langle n\rangle < \langle k\rangle /M$ implies that $l+m+n < k/N$, which is the required condition in the corollary 2.4.15. This M exists since $\langle k\rangle \approx k+1$ for all $k \in \mathbb{Z}_+$ and thus $\langle l\rangle \langle m\rangle \langle n\rangle \approx (l+1)(m+1)(n+1) \ge l+m+n$.

Recall that $|C_{klmn}| \lesssim a^{-k}$ is true for all the terms in part I of the sum (4.61). Therefore for $2s \geqslant 5/4$,

$$\begin{aligned}
I &\lesssim \sum_{l+m+n<k/N} \frac{a^{-k}}{\langle k\rangle} |q_l|\,|q_m|\,|q_n| \\
&\lesssim \frac{a^{-k}}{\langle k\rangle} \left(\sum_{l\in\mathbb{Z}_+} \langle l\rangle^{2s} |q_l|^2 \right)^{3/2} \left(\sum_{l\in\mathbb{Z}_+} \langle l\rangle^{-2s} \right)^{3/2} \\
&\lesssim \frac{a^{-k}}{\langle k\rangle} \|q\|^3_{l_2^s}.
\end{aligned}$$

Thus the component I is exponentially decreasing in k large, and it must belong to the class of $l_2^{+\infty}$.

For terms in part II, we use the same method as in the proof of the lemma 4.3.14. Using the fact that $|C_{klmn}| \lesssim (\langle k\rangle \langle l\rangle \langle m\rangle \langle n\rangle)^{-1/8+\varepsilon}$ $(0 < \varepsilon \ll 1)$, part two can be controlled by

$$
\begin{aligned}
II &\lesssim \langle k\rangle^{-1/8+\varepsilon} \sum_{\langle l\rangle\langle m\rangle\langle n\rangle \geq \langle k\rangle/M} \frac{\langle l\rangle^{-1/8+\varepsilon}\langle m\rangle^{-1/8+\varepsilon}\langle n\rangle^{-1/8+\varepsilon}}{\langle k+l-m-n\rangle} |\overline{q_l} q_m q_n| \\
&\lesssim \langle k\rangle^{-1/8+\varepsilon}\langle k\rangle^{-1/8+\varepsilon} \sum_{\langle l\rangle\langle m\rangle\langle n\rangle \geq \langle k\rangle/M} \frac{(\langle l\rangle\langle m\rangle\langle n\rangle)^{-s+1/2+\varepsilon}}{\langle k+l-m-n\rangle} \widetilde{q}_l \widetilde{q}_m \widetilde{q}_n \quad (M) \\
&\lesssim \langle k\rangle^{-1/4+2\varepsilon}\langle k\rangle^{-s+1/2+\varepsilon} \sum_{d\in\mathbb{Z}} \sum_{m+n-l=k-d} \frac{1}{\langle k+l-m-n\rangle} \widetilde{q}_l \widetilde{q}_m \widetilde{q}_n \quad (M),
\end{aligned}
$$

where $\widetilde{q}_j = \langle j\rangle^{s-1/2-\varepsilon}|q_j|$. By introducing the mapping τ as defined in (4.57), which is a bounded mapping from $l_2^s(\mathbb{Z}_+)$ to $l_b^{s,2}$, the above inequality can be continued for any $k \geq 0$

$$
\begin{aligned}
II &\lesssim \langle k\rangle^{-1/4+2\varepsilon}\langle k\rangle^{-s+1/2+\varepsilon} \sum_{d\in\mathbb{Z}} \sum_{m+1+n+1-l=k+2-d} \frac{1}{\langle d\rangle} \left|(\tau\widetilde{q})_{-l}(\tau\widetilde{q})_{m+1}(\tau\widetilde{q})_{n+1}\right| \quad (M) \\
&\lesssim \langle k\rangle^{-1/4+2\varepsilon}\langle k\rangle^{-s+1/2+\varepsilon} \left(\tau\widetilde{q} * \tau\widetilde{q} * \tau\widetilde{q} * \frac{1}{\langle d\rangle}\right)_{k+2} \quad (M). \qquad (4.62)
\end{aligned}
$$

Note that $\widetilde{q} \in l_2^{1/2+\varepsilon}(\mathbb{Z}_+)$ and $\frac{1}{\langle d\rangle} \in l_b^{1/2-\varepsilon,2}$, the property of the convolution (see theorem 4.3.11) enables us to deduce that the part II is in the space of $l_2^{s+1/4-4\varepsilon}(\mathbb{Z}_+)$ and its norm can be controlled by a constant (depending on s, ε and M) times $\|q\|_{l_2^s}^3$.

Since the positive real number ε can be chosen arbitrarily small and M is a universal constant in the estimate of the part II, together with the result of the estimate on the part I, we have proved that the Hamiltonian vector fields X_{F_4} is a continuous mapping from $l_2^s(\mathbb{Z}_+)$ $(X^{2s}(\mathbb{R}))$ into a smoother space $l_2^{s+1/4^-}(\mathbb{Z}_+)$ $(X^{2s+1/2^-}(\mathbb{R}))$ satisfying

$$
\|X_{F_4}(q)\|_{l^s} = O\left(\|q\|^3_{l_2^{s+1/4-}}\right) \quad (s).
$$

Of course the inequality (4.60) is also true. Note that F_4 is a continuous polynomial on $l_2^s(\mathbb{Z}_+)$, the C^∞ smoothness and real analyticity of the Hamiltonian vector fields X_{F_4} follows at once. □

After showing that the Hamiltonian vector fields X_{F_4} is smooth, we can construct the Lie transform generated by the flow of the Hamiltonian vector field of the function F_4 in the space $l_2^s(\mathbb{Z}_+)$ $(X^{2s}(\mathbb{R}))$ with $2s \geqslant 5/4$, and then consider the original Hamiltonian function in the new coordinates. But all these discussions become straightforward after we have proved lemma 4.4.1, since they are in fact the same as in the case of $5/4 > 2s > 1/2$. So we omit the argument here, and consider the proof of the main theorem 4.1 complete.

4.5 Application of the Birkhoff normal form

In this section, we will discuss some applications of the Birkhoff normal form theorem, which can be considered to have transformed the original system into two parts: one is the principal part, which has resonant nonlinear terms of order four; the other one is a small perturbation with order of at least six. Unlike the near integrable case, the principal system can not be solved out explicitly. Here we will focus on a study of properties of solutions of the principal system.

The Birkhoff normal form helps to introduce new coordinates to study the original PDE. Let $p = \Gamma^{-1}q$, then the Hamiltonian function H in the new coordinates will be $H(q) = H \circ \Gamma(p)$. Since the transformation Γ is symplectic, and thus preserves the Hamiltonian structure, we can write the system in the new coordinates as

$$\begin{aligned} \frac{dp}{dt} &= i\frac{\partial H \circ \Gamma}{\partial \overline{p}} \\ &= iI_\omega p + ig\widetilde{J}p + i\frac{\partial R}{\partial \overline{p}}, \end{aligned} \tag{4.63}$$

where $(I_\omega p)_k = (k + 1/2)\, p_k$ and $\left(\widetilde{J}p\right)_k = \sum_{k+l=m+n} C_{klmn}\overline{p_l}p_m p_n$. In doing this, it is natural to ask the transformation to be performed in the space of $l_2^s(\mathbb{Z}_+)$ $(X^{2s}(\mathbb{R}))$ with $2s \geq 1$, in which the Hamiltonian function H is finite.

The equation (4.63) can be regarded as a perturbation of the following system (let us call it the truncated system)

$$\frac{dp}{dt} = iI_\omega p + ig\widetilde{J}p, \tag{4.64}$$

whose associated Hamiltonian function is

$$\begin{aligned} H_{tr} &= H_2(p) + G(p) \\ &= \sum_{k=0}^{+\infty} \omega_k \left|p_k\right|^2 + \frac{g}{2} \sum_{k+l=m+n} C_{klmn} \overline{p_k}\overline{p_l} p_m p_n. \end{aligned} \tag{4.65}$$

It is reasonable to believe that solutions of the original system with small amplitude will have similar properties to those of the truncated system, at least over finite interval intervals of time.

The truncated Hamiltonian system turns out to be interesting. First of all, it has local well-posedness in a space consisting of rougher functions. In the proof of the Birkhoff normal form theorem (case of $2s < 5/4$), we showed that if s satisfies $2s > 1/2$ then the Hamiltonian function G is in the class of $\aleph^s$. As a consequence, the operator $\widetilde{J}$ in (4.64) is continuous on the space $l_2^s(\mathbb{Z}_+)$ $(X^{2s}(\mathbb{R}))$. Then through a study of its associated integral equation, we get the local well-posedness in the space $l_2^s(\mathbb{Z}_+)$ $(X^{2s}(\mathbb{R}))$ with $2s > 1/2$. Meanwhile, this system admits several conservation laws: the L^2 norm ($M = \sum_{k=0}^{+\infty} |p_k|^2$) and X^1 norm, the Hamiltonian function $G(p)$ and H_{tr} itself will be conserved along the flow. These conservation laws can be deduced from the facts

$$\{H_{tr}, M\} = 0,\ \{H_{tr}, H_2\} = 0 \text{ and } \{H_{tr}, H_{tr}\} = 0.$$

In particular, any initial data in the space $l_2^{1/2}(\mathbb{Z}_+)$ $(X^1(\mathbb{R}))$ will result in a global flow in time.

In general, solutions of the Hamiltonian system (4.64) cannot be solved out explicitly, but there exist particular single-mode activated solutions, which are explicit, namely

$$\begin{cases} p_{k_0}(t) = p_{k_0}(0)\exp\{iw_{k_0}t + iC_{k_0k_0k_0k_0}\left|p_{k_0}\right|^2 t\}, \\ \qquad p_k(t) = 0, \qquad k \neq k_0. \end{cases}$$

For other solutions, some interesting symmetry properties will arise.

The first one concerns the difference of the sign of the constant g. Recall that in the original system, $g = 1$ $(g = -1)$ represents the defocusing case (respectively, focusing case). Given a fixed datum at time zero, let us denote the k-th action function of the solution of the truncated system (4.64) by $I_k^{(de)}(t) = \left|p_k^{(de)}(t)\right|^2$ and $I_k^{(f)}(t) = \left|p_k^{(f)}(t)\right|^2$ respectively for $g = 1$ and $g = -1$. We claim that these action functions only differ in the time direction, that is, $I_j^{(f)}(t) = I_j^{(de)}(-t)$.

This property is due to the following observation: if we introduce a time-dependent coordinate change

$$\widetilde{p}(t) = e^{-iL_\omega t}p(t), \tag{4.66}$$

or in each direction of the eigenfunction

$$\widetilde{p_k}(t) = e^{-i\omega_k t}p_k(t),$$

then the system (4.64) will be changed into the form of

$$\frac{d\widetilde{p}}{dt} = ig\widetilde{J}\widetilde{p}, \tag{4.67}$$

whose associated Hamiltonian function is

$$G(\widetilde{p}) = \frac{g}{2}\sum_{k+l=m+n} C_{klmn}\overline{\widetilde{p}_k}\overline{\widetilde{p}_l}\widetilde{p}_m\widetilde{p}_n.$$

Note that the difference of the signature of the constant g in the system (4.67) will only result in the reverse of the time direction, and the coordinate change (4.66) will not affect on its action functions, we can get the claim.

Meanwhile, since the right side of the system (4.67) is homogeneous of order three, another symmetry property will arise here: if $\widetilde{p}(t)$ is a solution of the system (4.67) then $\lambda\widetilde{p}(\lambda^2 t)$ (λ any positive real number) is also a solution.

Back to the original system (4.63), a simple result on the large time evolution can be easily deduced. Let us denote $N_s(p) := \|p\|^2_{X^s} = \sum_{k=0}^{+\infty}(k+\frac{1}{2})^s p_k\overline{p_k}$, which equals to M if $s=0$ and $H_2(p)$ if $s=1$. Using that N_1 Poisson commutes with itself and the function $G(p)$, we have

$$\begin{aligned}\frac{d}{dt}N_1(p(t)) &= \{N_1, H\circ\Gamma\}\\ &= \{N_1, R\}.\end{aligned}$$

According to the Birkhoff normal form theorem, the function R in is in the class of $\aleph^s$. Then from the property of the Poisson bracket, it follows that

$$\begin{aligned}\left|\frac{d}{dt}N_1(p(t))\right| &= |\{N_1, R\}|\\ &\le CN_1^3.\end{aligned}$$

We deduce that there exists $\varepsilon_0 > 0$ and $C > 0$, such that if the initial data $\|p_0\|_{X^1} = \varepsilon < \varepsilon_0$ the solution $p(t)$ of the Hamiltonian system associated to H which takes value (p_0) at $t=0$ satisfies

$$\|p(t)\|_{X^1} \le 2\varepsilon \quad \text{for } |t| \le \frac{C}{\varepsilon^4}, \tag{4.68}$$

and for $r \geq 2$

$$|N_1(p(t)) - N_1(p(0))| \leq \varepsilon^r \quad \text{for } |t| \leq \frac{C}{\varepsilon^{6-r}}. \tag{4.69}$$

In the rest of this section, we will study another truncated Hamiltonian system, which only has one nonintegrable resonant term. In this special case, the oscillations of the action functions of its solutions can be understood very precisely.

4.5.1 The {0,1,2} system

Let us consider the following truncated Hamiltonian system

$$h = \sum_{k=0}^{2} \omega_k |p_k|^2 + \frac{g}{2} \sum_{k+l=m+n,\ k,l,m,n \in \{0,1,2\}} C_{klmn} \overline{p_k}\,\overline{p_l} p_m p_n, \tag{4.70}$$

where $\omega_k = k + 1/2$, $C_{klmn} = \int h_k(x) h_l(x) h_m(x) h_n(x) dx$ and $g = \pm 1$ (defocusing and focusing case). It is an approximation to the truncated system (4.65).

We are interested in the following question: how will the action functions associated with its eigenmodes behave as time evolves? This problem has a close relationship with the oscillation of the X^{2s} norm of the solutions. Let $I_k(t) = |p_k|^2$ denote the action function of the k-th eigenmode. There are such two integrals in this system: one is the mass $M = \sum_{k=0}^{2} I_k$ and the other one is the X^1 norm defined as $\|p(t)\|_{X^1}^2 = \sum_{k=0}^{2} \omega_k I_k$. From these two conservation law, it can be deduced that if the initial data $p(0) = (p_0^*, p_1^*, p_2^*)$ is given, the values of the action functions $I_k(t)$ must lie on a straight line (in fact a line segment. since action functions are nonnegative) passing through the initial state

$$\frac{I_0(t) - I_0^*}{1} = \frac{I_1(t) - I_1^*}{-2} = \frac{I_2(t) - I_2^*}{1}.$$

We are interested that how the X^{2s} $(2s > 1)$ norm of the solutions will behave as time evolves. For a general initial datum (I_0^*, I_1^*, I_2^*), if every state in the whole line segment can be reached, then the minimum and maximum values of the X^{2s} $(2s > 1)$ norm will happen at the two end points of the line segment: one is corresponding to concentrate energy of the system to the eigenmode I_1 until one of the modes I_0 and I_2 becomes zero, which gives the minimum value of the X^{2s} $(2s > 1)$ norm; the other one is corresponding to transfer the energy in the eigenmode I_1 to the other two modes until I_1 becomes zero, which gives maximum value of the X^{2s} $(2s > 1)$ norm. So we want to investigate the following questions: (i) For a general initial state, can these two extreme states really appear as time evolves? (ii) What is the behavior of the X^{2s} $(2s > 1)$ norm of the solutions, whether or not it is monotone increasing, decreasing or oscillating between the two extreme states?

These questions can be answered through finding out all the possible phase portraits for this system. First, let us reduce the problem to a planary Hamiltonian system depending on two parameters. The basic idea is to choose a good set of action-angle variables to simplify the Hamiltonian system.

From the study on the infinite dimensional system (4.65), we know that the constant $g = 1$ and $g = -1$, i.e. the focusing case and the defocusing case, only differs by a choice of the time direction in rotating coordinates. So without loss of generality, we always set $g = 1$ in our analysis below. This special truncated system leads to the following ODE system

$$\begin{cases} \dot{p_0} = i\omega_0 p_0 + 2i\sum_{k=0}^{2} C_{00kk} \left|p_k\right|^2 p_0 - \imath C_{0000} \left|p_0\right|^2 p_0 + iC_{0211}\overline{p_2}p_1p_1 \\ \dot{p_1} = i\omega_1 p_1 + 2i\sum_{k=0}^{2} C_{11kk} \left|p_k\right|^2 p_1 - iC_{1111} \left|p_1\right|^2 p_1 + i2C_{0211}\overline{p_1}p_0p_2 \\ \dot{p_2} = i\omega_2 p_2 + 2i\sum_{k=0}^{2} C_{22kk} \left|p_k\right|^2 p_2 - \imath C_{2222} \left|p_2\right|^2 p_2 + iC_{0211}\overline{p_0}p_1p_1, \end{cases} \tag{4.71}$$

where in particular, the coefficient C_{0112} is $1/8\sqrt{\pi}$.

Let us introduce the following action-angle variables (I, φ) defined by

$$p_k = \sqrt{I_k} e^{\imath\varphi_k}, \quad k = 0, 1, 2.$$

This is a symplectic transformation. In the domain of $\{I_k > 0 \text{ and } \varphi_k \in T = \mathbb{R}/2\pi\}$, this transformation is symplectic. The Hamiltonian function will be in the following form

$$h = \langle \omega, I \rangle + \langle I, BI \rangle + 2C_{0112} I_1 \sqrt{I_0 I_2} \cos(\varphi_0 + \varphi_2 - 2\varphi_1), \tag{4.72}$$

where $\omega = (\omega_0, \omega_1, \omega_2)^t$ and $I = (I_0, I_1, I_2)^t$ are vectors in $\mathbb{R}^3$, $\langle \cdot, \cdot \rangle$ denotes the usual inner product of the two vectors, and $B = (b_{kl})$ is the 3×3 coefficient matrix defined by

$$b_{kl} = \begin{cases} \frac{1}{2} C_{kkkk} & if \quad k = l \\ C_{kkll} & if \quad k \neq l. \end{cases}$$

Referring to the explicit value of the coefficients C_{klmn} in the chapter two, the matrix B is given by

$$B = \frac{1}{\sqrt{2\pi}} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{3}{8} \\ \frac{1}{2} & \frac{3}{8} & \frac{7}{16} \\ \frac{3}{8} & \frac{7}{16} & \frac{41}{128} \end{pmatrix}.$$

Noting that there are two integrals of motion (mass and X^1 norm) for this system, we can make a further symplectic transformation to simplify the system in the domain under consideration. Set the matrix

$$R = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{pmatrix}, \tag{4.73}$$

and introduce the new parameters $(J.\theta)$ as follows

$$J = RI. \text{ and } \theta = R\varphi,$$

or in other words,

$$\begin{cases} J_0 = \frac{1}{\sqrt{3}}(I_0 + I_1 + I_2) \\ J_1 = \frac{1}{\sqrt{2}}(I_0 - I_2) \\ J_2 = \frac{1}{\sqrt{6}}(I_0 - 2I_1 + I_2) \end{cases} \tag{4.74}$$

and

$$\begin{cases} \theta_0 = \frac{1}{\sqrt{3}}(\varphi_0 + \varphi_1 + \varphi_2) \\ \theta_1 = \frac{1}{\sqrt{2}}(\varphi_0 - \varphi_2) \\ \theta_2 = \frac{1}{\sqrt{6}}(\varphi_0 - 2\varphi_1 + \varphi_2). \end{cases} \tag{4.75}$$

In this new coordinates (J, θ), the Hamiltonian function h can be written as

$$h = \langle R\omega, J\rangle + \langle J, RBR^t J\rangle + 2C_{0112}I_1\sqrt{I_0 I_2}\cos(\sqrt{6}\theta_2), \tag{4.76}$$

where I_k are all the linear combinations of the variables J_0, J_1, J_2

$$\begin{cases} I_0 &= \frac{1}{\sqrt{3}}J_0 + \frac{1}{\sqrt{2}}J_1 + \frac{1}{\sqrt{6}}J_2 \\ I_1 &= \frac{1}{\sqrt{3}}J_0 + 0 - \frac{2}{\sqrt{6}}J_2 \\ I_2 &= \frac{1}{\sqrt{3}}J_0 - \frac{1}{\sqrt{2}}J_1 + \frac{1}{\sqrt{6}}J_2. \end{cases} \tag{4.77}$$

The matrix RBR^t is computed to be

$$RBR^t = \frac{1}{\sqrt{2\pi}} \begin{pmatrix} \frac{163}{128} & \frac{31}{128\sqrt{6}} & -\frac{5}{128\sqrt{2}} \\ \frac{31}{128\sqrt{6}} & \frac{9}{256} & \frac{7}{256\sqrt{3}} \\ -\frac{5}{128\sqrt{2}} & \frac{7}{256\sqrt{3}} & -\frac{29}{256} \end{pmatrix}. \tag{4.78}$$

The point of this coordinate transformation is that the Hamiltonian function h does not depend on the parameter θ_0 and θ_1, and therefore the quantities J_0 and J_1 are both integrals of motion of the system. From the relationship (4.77), it follows that the action functions $\{I_k(t)\}$ in fact only have one remaining degree freedom. In particular, from an initial point with action functions (I_0^*, I_1^*, I_2^*), the solution satisfies

$$(I_0(t), I_1(t), I_2(t)) = (I_0^*, I_1^*, I_2^*) + \frac{1}{\sqrt{6}}(J_2(t) - J_2^*)(1, -2, 1).$$

It means that the understanding on the behavior of the function $J_2(t)$ provides every detail for the oscillations of action functions $\{I_k(t)\}$.

Focus on the oscillation of the function $J_2(t)$. The Hamiltonian system (4.76) is written as a system of two ODEs

$$\begin{cases} \frac{d}{dt}J_2 = -\frac{\partial h}{\partial \theta_2} = \frac{\sqrt{6}}{4\sqrt{\pi}} I_1 \sqrt{I_0 I_2} \sin(\sqrt{6}\theta_2) \\ \frac{d}{dt}\theta_2 = \frac{\partial h}{\partial J_2} = \frac{1}{\sqrt{2\pi}} \left(\frac{-5}{64\sqrt{2}} J_0 + \frac{7}{128\sqrt{3}} J_1 + \frac{-29}{128} J_2 \right) \\ \qquad\qquad + \frac{1}{4\sqrt{\pi}} \frac{d}{dJ_2} \left(\sqrt{I_0 I_2} I_1 \right) \cos(\sqrt{6}\theta_2). \end{cases}$$

By introducing two functions of J_2

$$\begin{cases} a(J_2) = \frac{1}{\sqrt{2\pi}} \left(\frac{5}{64\sqrt{2}} J_0 - \frac{7}{128\sqrt{3}} J_1 + \frac{29}{128} J_2 \right) \\ b(J_2) = \frac{1}{4\sqrt{\pi}} \frac{d}{dJ_2} \left(\sqrt{I_0 I_2} I_1 \right), \end{cases} \tag{4.79}$$

the ODE system can be rewritten as

$$\begin{cases} \frac{d}{dt}J_2 = \frac{\sqrt{6}}{4\sqrt{\pi}} I_1(J) \sqrt{I_0(J) I_2(J)} \sin(\sqrt{6}\theta_2) \\ \frac{d}{dt}\theta_2 = b(J_2)\cos(\sqrt{6}\theta_2) - a(J_2). \end{cases} \tag{4.80}$$

In the above system, J_0 and J_1 play the role of two parameters, which provide restrictions on the range of the variable J_2. Since all the action functions $\{I_k(t)\}$ must be nonnegative (including the singular points of the symplectic transformations), we deduce the following from the relationship (4.74) and (4.77)

$$\begin{cases} J_0 \geq 0 \\ |J_1| \leq \sqrt{\frac{3}{2}} J_0 \\ J_2^- \leq J_2 \leq J_2^+, \end{cases} \tag{4.81}$$

where $J_2^+ = J_0/\sqrt{2}$ and $J_2^- = -\sqrt{2}J_0 + \sqrt{3}\,|J_1|$. We say that a pair of real number (J_0, J_1) is admissible, if it satisfies the first two inequalities in (4.81). Once an admissible pair of integrals (J_0, J_1) is given, the solutions $J_2(t)$ must fall into the region of $\left\{J_2^- \leq J_2 \leq J_2^+\right\}$. In particular, $J_2^- < J_2 < J_2^+$ corresponds to all the regular points of the symplectic transformations, $J_2 = J_2^-$ corresponds to all the singular points with q_0 or q_2 equal to zero, and $J_2 = J_2^+$ corresponds to all the singular points with q_1 equal to zero. Besides that, it is true that

$$\begin{cases} \text{if } J_1 \geq 0, \quad J_2^- = J_2 \Leftrightarrow I_2 = 0 \\ \text{if } J_1 \leq 0. \quad J_2^- = J_2 \Leftrightarrow I_0 = 0. \end{cases}$$

For any given initial data (I_0^*, I_1^*, I_2^*) (or (J_0^*, J_1^*, J_2^*)), if the solution $J_2(t)$ approaches point $J_2^-(J_0^*, J_1^*)$, it means that action is being transferred to the 1-th eigenmode, and the X^{2s} $(2s > 1)$ norm of the solution is becoming smaller; on the contrary, if the solution $J_2(t)$ approaches the point $J_2^+(J_0^*, J_1^*)$, it means the action of the 1-th eigenmode is being transferred into the other two modes, and the X^{2s} $(2s > 1)$ norm of the solution is becoming larger. So our task has been reduced to understand the oscillations of the solution $J_2(t)$ of the Hamiltonian ODE system (4.80) given two admissible parameters (J_0, J_1). The domain under consideration is the set $(J_2^-, J_2^+) \times \mathbb{R}/2\sqrt{6}\pi$.

In the following paragraphs, we will classify and describe all the possible phase portraits. When $J_1 = 0$, the function $b(J_2)$ has very different smoothness properties at the point J_2^- than in the other cases, so we first discuss this case. We find that there is only one fixed point within the domain being considered. From the first equality of system (4.80), it follows that all the possible fixed points must fall into the vertical lines $\theta_2 = 0, \pi/\sqrt{6}$; from the second equality, we can further determine that there is only one fixed point on the vertical line $\theta_2 = 0$, and there are no fixed point on the vertical line $\theta_2 = \pi/\sqrt{6}$. The fact needed here is that the functions $a(J_2)$ and $b(J_2)$ both depend linearly on J_2 and they satisfy

$$\begin{cases} a(J_2^-) = \frac{-3}{16\sqrt{\pi}} J_0 & a(J_2^+) = \frac{39}{256\sqrt{\pi}} J_0 \\ b(J_2^-) = \frac{\sqrt{2}}{8\sqrt{\pi}} J_0 & b(J_2^+) = -\frac{\sqrt{2}}{8\sqrt{\pi}} J_0. \end{cases}$$

It is also easy to conclude that the unique fixed point is elliptic by studying its local linearized system.

The following picture (4.1) is for the case of $J_0 = 1$ and $J_1 = 0$. In the picture the x coordinate denotes $\sqrt{6}\theta_2$ and the y coordinate denotes J_2. All the solutions passing only through the regular points are periodic, which implies that along that orbit the behavior of the parameters $(J_2(t), \theta_2(t))$ is periodic. On the horizontal line $J_2 = J_2^+$, that is, $I_1 = 0$. system (4.71) reduces to an integrable system and all solutions will remain on that line. On the horizontal line $J_2 = J_2^-$, that is, $I_0 = I_2 = 0$, all solutions of system (4.71) reduce to the single mode activated solution; they will also remain on that line. There is one orbit that remains in the considered domain but has two "ending" points at point $(J_2^+, \pm\theta^*)$, where θ^* satisfies $b(J_2^+)\cos(\sqrt{6}\theta^*) - a(J_2^+) = 0$. This is a separative solution, and needs infinite time

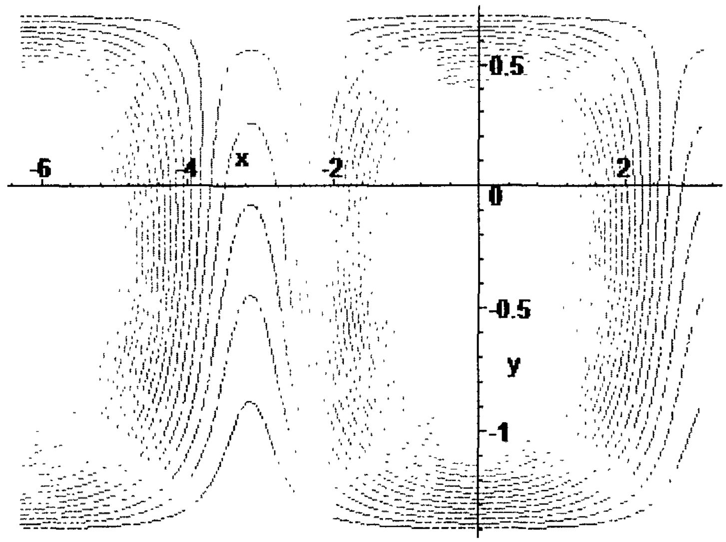

Figure 4.1: $J_0 = 1$ and $J_1 = 0$

to finish its tour along the whole orbit.

When $J_1 \neq 0$, the phase portrait will be more complicated. The numbers of the fixed points and the topological structure of the system can both vary according to the different choices of the admissible integrals pair (J_0, J_1).

At first let us study flows near or within the singular points. On the horizontal line $J_2 = J_2^+$, that is, $I_1 = 0$, the system (4.71) reduces to an integrable system and all the flows there will remain on that line. In other words, if a flow starts from a regular point, it can at most approach to this horizontal line, but never reach it. As to the case of the horizontal line $J_2 = J_2^-$, things are different. Suppose a flow reach a point on that line at some time t_0, then it is true that one and only one of the components I_0 and I_2 is zero. In the case of $q_0 = 0$ (respectively $q_2 = 0$), we deduce from the system (4.71) that $\dot{q}_0$ (respectively, $\dot{q}_2$) will be nonzero. Thus we know there exists an interval of time containing t_0, such that $q_0(t)$ (respectively, $q_2(t)$) is nonzero except the time point t_0. It means that if a flow evolves into a singular point on the horizontal line $J_2 = J_2^-$, it will "leave" it at once and be back into regular points state again. So there are two kinds of solutions that have its points or its limiting points on the horizontal lines consisting of singular points: if its limiting points are on the horizontal line $J_2 = J_2^+$, the flow can be arbitrarily close to the point as time evolves but never reach it; if one of its points is on the horizontal line $J_2 = J_2^-$, then the flow can really reach that point, but will also leave that line at once and come back into regular points state again.

Secondly, let us consider the numbers of the fixed points. The fixed points must satisfy

$$\begin{cases} \sqrt{6}\theta_2 = 0, \pi \\ b(J_2)\cos(\sqrt{6}\theta_2) - a(J_2) = 0. \end{cases}$$

In the case of $\sqrt{6}\theta_2 = 0$, there is always a unique solution. This is because the function $b(J_2)$ is decreasing and $a(J_2)$ is increasing, and they satisfy

$$\lim_{J_2 \to (J_2^-)^+} b(J_2) = +\infty > a(J_2^-) \text{ and } b(J_2^+) < 0 < a(J_2^+).$$

The case of $\sqrt{6}\theta_2 = \pi$ is more complicated. Any fixed point on it must satisfy

$$-b(J_2) = a(J_2).$$

Here let us introduce parameter $s \neq 0 \in [-1, 1]$, which is defined by the relationship $J_1 = \sqrt{3/2}J_0 s$. It can be verified that the function $-b(J_2)$ is increasing and concave. Meanwhile we have

$$\left.\frac{d}{dJ_2}\left(-b(J_2)\right)\right|_{J_2=J_2^+} = \frac{1}{\sqrt{\pi}}\frac{1}{6\sqrt{1-s^2}}$$

and

$$\frac{d}{dJ_2}\left(a(J_2)\right)\Big|_{J_2=J_2^+} = \frac{1}{\sqrt{\pi}}\frac{29}{128\sqrt{2}},$$

which means $-b'(J_2^+) > a'(J_2^+)$. So there are only these three possibilities:

- If $-b(J_2^+) > a(J_2^+)$ then there is exactly one fixed point in the interval of (J_2^-, J_2^+);
- If $-b(J_2^+) = a(J_2^+)$ then there is exactly one fixed point at J_2^+;
- If $-b(J_2^+) < a(J_2^+)$ then there are no fixed points on the interval $[J_2^-, J_2^+]$.

The equation for the parameter s at the transition points is that

$$-b(J_2^+) = a(J_2^+),$$

or in the form of

$$\frac{J_0}{4\sqrt{2\pi}}\sqrt{1-s^2} = \frac{J_0}{256\sqrt{\pi}}(39-7s).$$

There are two solutions $s_1 = -0.3877520341\ldots$ and $s_2 = 0.6481239941\ldots$. Thus the three possibilities mentioned above correspond respectively to the following cases $s \in (s_1, 0) \cup (0, s_2)$ and $s = s_1$ or s_2 and $s \in [-1, s_1) \cup (s_2, 1]$.

Finally we will provide the phase portraits, mentioning these two simple properties of them: (1) all the fixed points which lie out side of the line of $J_\pm$ are elliptic points, (2) the phase portrait should be symmetric respect to the vertical line $\theta_2 = 0$. In the following phase portraits, we always choose $J_0 = 1$. They are arranged in the decreasing order of the parameter s.

When $J_0 = 1$ and $J_1 = 1$ ($s = 0.816\ldots$), the phase portrait is the picture (4.2);

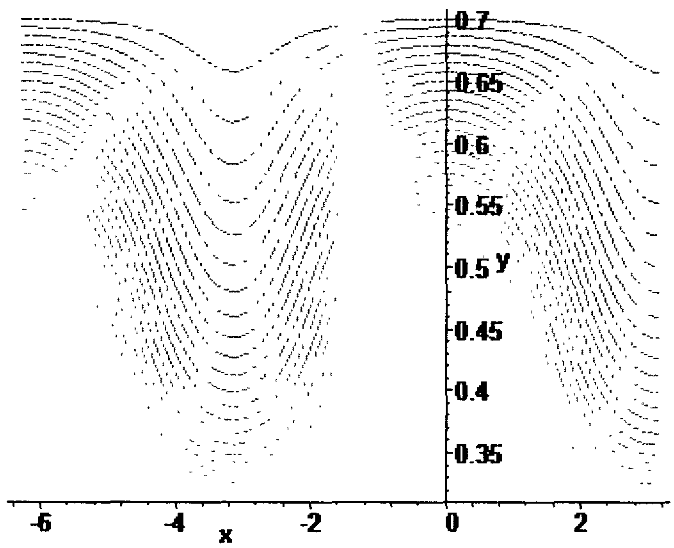

Figure 4.2: $J_0 = 1$ and $J_1 = 1$

When $J_0 = 1$ and $J_1 = 0.79378653779$ ($s = 0.6481239941$, almost equal to s_2), the phase portrait is the picture (4.3);

When $J_0 = 1$ and $J_1 = 0.6$, ($s = 0.489\ldots$), the phase portrait is the picture (4.4);

When $J_0 = 1$ and $J_1 = 0.1$ ($s = 0.081\ldots$), the phase portrait is the picture (4.5);

When $J_0 = 1$ and $J_1 = -0.1$ ($s = -0.081\ldots$), the phase portrait is the picture (4.6);

When $J_0 = 1$ and $J_1 = -0.474897315135$ ($s = -0.3877520341$, almost equal to s_1), the phase portrait is the picture (4.7);

When $J_0 = 1$ and $J_1 = -1$ ($s = -0.816\ldots$), the phase portrait is the picture

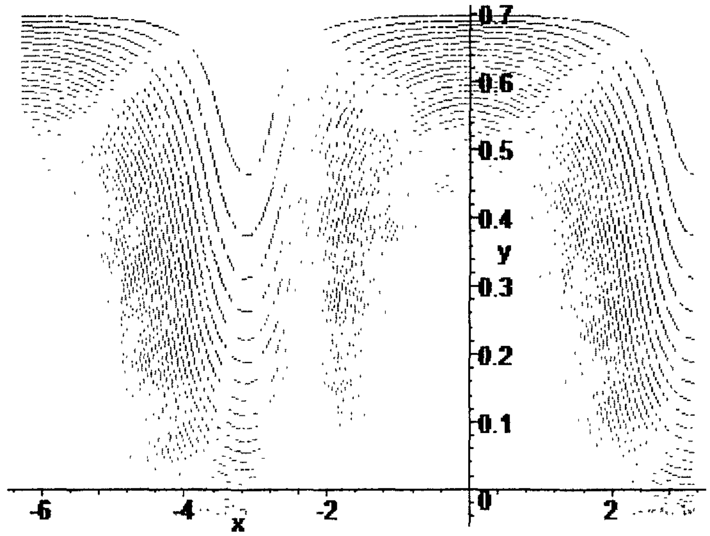

Figure 4.3: $J_0 = 1$ and $J_1 = 0.79378653779$

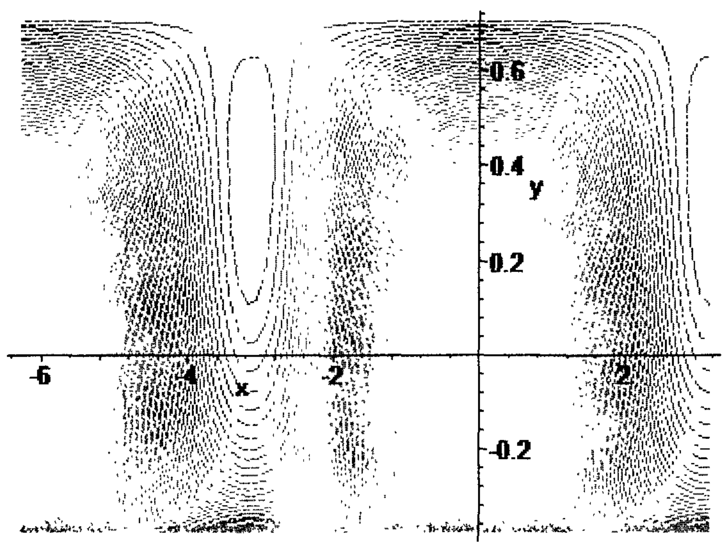

Figure 4.4: $J_0 = 1$ and $J_1 = 0.6$

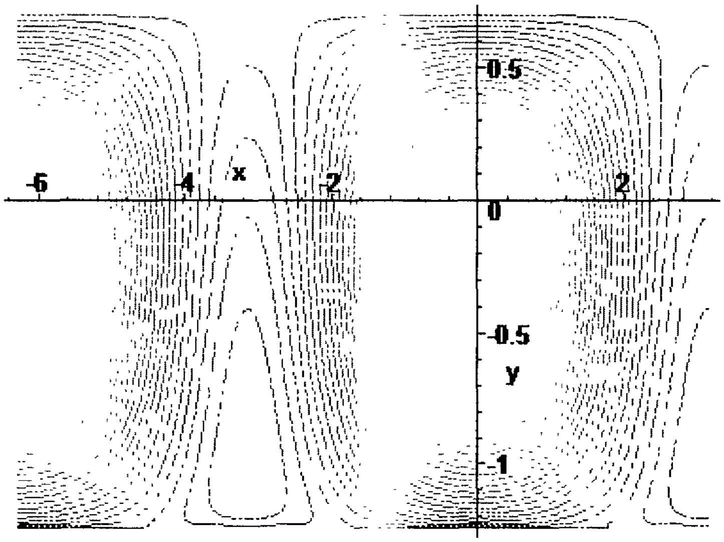

Figure 4.5: $J_0 = 1$ and $J_1 = 0.1$

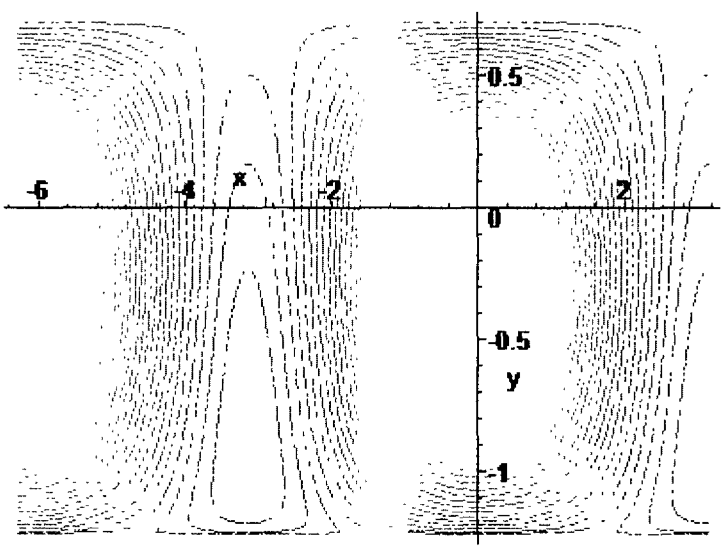

Figure 4.6: $J_0 = 1$ and $J_1 = -0.1$

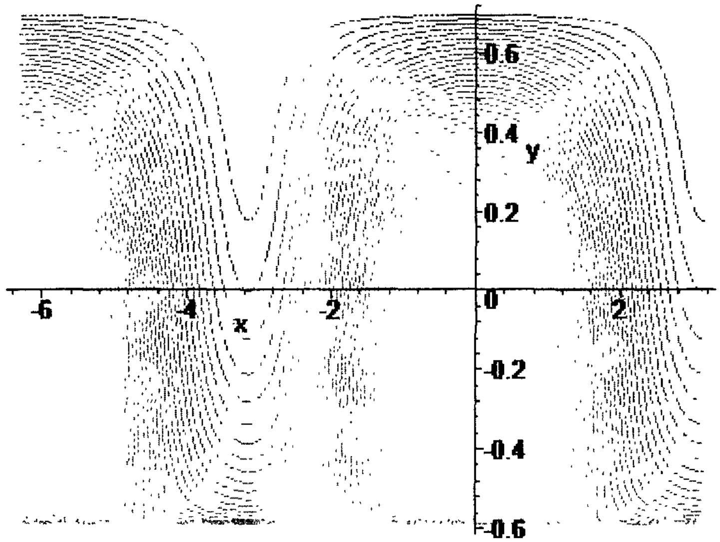

Figure 4.7: $J_0 = 1$ and $J_1 = -0.474897315135$

(4.8).

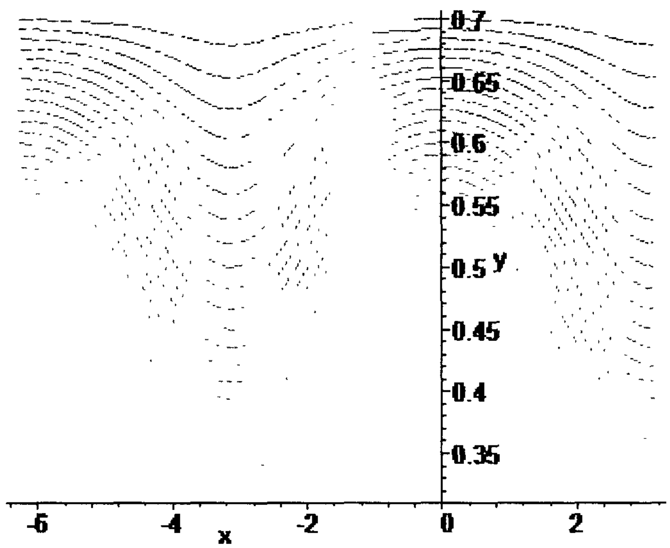

Figure 4.8: $J_0 = 1$ and $J_1 = -1$

Chapter 5

Conclusions

This thesis mainly focus on a Birkhoff normal forms theorem for the partial differential equation

$$\begin{cases} i\psi_t = \frac{1}{2}\psi_{xx} - \frac{x^2}{2}\psi - g|\psi|^2\psi & x \in \mathbb{R}^1 \\ \psi(x,0) = \psi_0(x) & \psi \text{ complex valued,} \end{cases} \tag{5.1}$$

which is known as the Gross-Pitaevskii (GP) equation. The following results are obtained:

1. An estimate of the coupling coefficients for the Hermite functions,

$$C_{mmnn} \approx \frac{2}{\sqrt{m}} \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{1 - s^2 \sin^2 \theta}} = \frac{2}{\sqrt{m}} E(s),$$

where $s = \sqrt{\frac{n}{m}} < 1$. Thus we get a sharp estimate on C_{mmnn}. In the case that one of the indices is much larger than all the others, say $k > N(l + m + n)$, then

$$|C_{klmn}| \lesssim a^{-k}.$$

2. We introduce a family of Hilbert spaces $X^s(\mathbb{R}^m)$, which provide good working spaces for the GP equations with harmonic potential. In particular, many

important spaces such as $S'(\mathbb{R}^n)$, $S(\mathbb{R}^n)$, $L^2(\mathbb{R}^m)$ and virial space can all be embedded in this Hilbert scale. We introduce the operator A^σ, which acts as isometry mapping between spaces $X^s(\mathbb{R}^m)$ with different regularity index, similar to the operator $(I-\Delta)^\sigma$ acting on the Sobolev spaces. In particular, the unitary group associated with the operator e^{iAt} acts on these spaces preserving the norm, and the Fourier transformation is essentially embedded in this group.

This family of Hilbert spaces provide very natural spaces for the N-presentation theory for the rapid decreasing functions and tempered distribution [Si]. They are also closely related with the work is by V. Bargmann, especially the construction of Bargmann's spaces. In fact it is implied from his work [Bar2] that our function spaces $X^s(\mathbb{R}^m)$ are isometric to the Bargmann's spaces F_m^s under the Segal-Bargmann transformation, which originally was defined as a transformation from the space $L^2(\mathbb{R}^m)$ onto the Fock space.

Meanwhile, the function spaces have very close relationship with the Sobolev spaces. When s is a nonnegative integer, say $s = n \geq 0$, then $X^n(\mathbb{R}^m) = H^n(\mathbb{R}^m) \bigcap \mathcal{F}H^n(\mathbb{R}^m)$; and in the case of negative integer we have $X^{-n}(\mathbb{R}^m) = H^{-n}(\mathbb{R}^m) + \mathcal{F}H^{-n}(\mathbb{R}^m)$. In particular, since $S'(\mathbb{R}^m) = \bigcup_{n=1}^{+\infty} X^{-n}(\mathbb{R}^m)$, the last result above implies that for any tempered distribution there is an integer $n \in \mathbb{N}$ such that it can be decomposed into two parts: one is in Sobolev space $H^{-n}(\mathbb{R}^m)$ and the other one is in $\mathcal{F}H^{-n}(\mathbb{R}^m)$, which must be locally integrable. It is believable that in general same result holds for any real number s. Although we do not yet have a proof of this conjecture, the following fact, which can be deduced out by the conjecture, has been proved independently in this thesis: let $\lceil m/2 \rceil$ denote the smallest integer greater than $m/2$; then for any real number $s \geq \lceil m/2 \rceil$ the space $X^s(\mathbb{R}^m)$ is an ideal and also a subalgebra of the space $H^s(\mathbb{R}^m)$, which itself is an algebra. If the conjecture is true, then for $s \geq 0$ any function in $X^{-s}(\mathbb{R}^m)$ can be decomposed into two parts: one is in $H^{-s}(\mathbb{R}^m)$ and the other one is in $\mathcal{F}H^{-s}(\mathbb{R}^m)$.

3. We apply these function spaces in the study of the local and global well-posedness problem for the GP equation. It is known that initial data in the space $X^1(\mathbb{R})$ will lead to a unique global flow in time in the same space, no matter whether it is in the defocusing or focusing case. We generalize this result into the cases of the spaces $X^n(\mathbb{R})$ with any integer $n \geq 2$. In the case of $n = 1$, its X^n norm remains bounded; in the cases of $n \geq 2$, its norm can have a growth at most at an exponential rate. When the initial data is not particularly smooth, say $\psi_0(x) \in X^s(\mathbb{R})$ with $1/2 < s < 1$, our conjecture ($X^s(\mathbb{R}) = H^s(\mathbb{R}) \bigcap \mathcal{F}H^s(\mathbb{R})$) suggests that there should be a unique local flow in the space $X^s(\mathbb{R})$.

We give a proof for a Birkhoff normal forms theorem for the equation (5.1). Specifically, the Hamiltonian function H of the one dimensional GP equation (5.1) can be transformed by symplectic transformations into the form

$$H \circ \Gamma(p) = \sum_{k \geq 0} \omega_k |p_k|^2 + \frac{g}{2} \sum_{k+l=m+n} C_{klmn} \overline{p_k} \overline{p_l} p_m p_n + R(p),$$

where the remainder term $R(p)$ is real analytic and of order 6 near the origin, for which X_R is a real analytic Hamiltonian vector fields in the function space $X^s(\mathbb{R})$. In this way, the original Hamiltonian PDE system is transformed to the problem of a perturbation of the Hamiltonian system with the Hamiltonian function

$$H_{tr}(p) = \sum_{k \geq 0} \omega_k |p_k|^2 + \frac{g}{2} \sum_{k+l=m+n} C_{klmn} \overline{p_k} \overline{p_l} p_m p_n. \tag{5.2}$$

One class of explicit solutions of the system (5.2) are the following one-mode activated solutions

$$\begin{cases} p_{k_0}(t) = p_{k_0}(0)\exp\{iw_{k_0}t + iC_{k_0k_0k_0k_0}\left|p_{k_0}\right|^2 t\}, \\ p_k(t) = 0, \quad k \neq k_0. \end{cases} \tag{5.3}$$

But unlike integrable cases, in general the solutions of the system (5.2) are not able to be written out explicitly. All of the solutions of the system (5.2) have the following properties: They will preserve the $l^2(\mathbb{Z}_+)$ ($L^2(\mathbb{R})$) norm and $l_2^{1/2}(\mathbb{Z}_+)$ ($X^1(\mathbb{R})$) norms in the p coordinates; if we compare the oscillations on every eigenmodes under the focusing and the defocusing case initialed from a common initial data at time zero, say $I_j^{(f)}(t)$ and $I_j^{(de)}(t)$, then they only differs in the time direction in the rotating coordinates, that is, $I_j^{(f)}(t) = I_j^{(de)}(-t)$. This fact is due to the observation that the system is equivalent to the system

$$H_{tr}(\widetilde{p}) = \frac{g}{2}\sum_{k+l=m+n} C_{klmn}\overline{\widetilde{p}_k}\overline{\widetilde{p}_l}\widetilde{p}_m\widetilde{p}_n, \tag{5.4}$$

which is not sensitive to the signature of the constant g. As a by product, all the solutions of the system (5.4) admit the following symmetry property: if $\widetilde{p}(t)$ is a solution then so is $\lambda\widetilde{p}(\lambda^2 t)$ (λ any real number greater than zero).

We have furthermore studied a simplified system modeled as (5.4), namely the $\{0, 1, 2\}$ system

$$h = \sum_{k\geq 0}^{2}\omega_k\left|p_k\right|^2 + \frac{g}{2}\sum_{k+l=m+n,\ 0\leq k,l,m,n\leq 2} C_{klmn}\overline{p_k}\overline{p_l}p_m p_n.$$

This system is completely integrable in certain action angle variables, and succeptable to a phase portrait analysis. We have classified all the phase portraits for this system.

Considering the original system (5.1), one direct corollary of the Birkhoff normal form is that if the initial data of the system (5.1) is small enough, the X^1

norm of the flow in the coordinate $p = \Gamma^{-1}(q)$ will change very slowly. Meanwhile, the property of the solutions of the system (5.4) suggests that there is no big difference between the focusing case and the defocusing case when the perturbation is sufficiently small.

We are planning future research on solutions of the GP equation from this point of view. Generally, perturbation theory of completely resonant systems is harder to study than that of the nearly integrable cases. In the latter case, each eigenmode's action function is an integral for the unperturbed system, and Nekhoroshev style results state that small perturbation will keep action function of each eigenmode not far away from the initial state for a long time. In the completely resonant system case, even without perturbation those action functions in general are no longer constants of motion. We may turn to consider under small perturbation whether the solution will stay close to the orbit of the unperturbed system. Also we are interested in the following question: Does there exist (quasi) periodic oscillations to our original system? In particular, the class of special solutions in (5.3) is a good starting point. Meanwhile, the introduction of our function spaces makes it possible to do a more careful study on the regularities of the solutions of the nonlinear Gross-Pitaevskii equation. For example, one question is whether we have local or global well-posedness for the Cauchy problem with rough initial data? Harmonic analysis on Hermite expansions series will play an important role in the studies in this field, the kernel of which is provided in this dissertation.

Bibliography

[A] V. I. Arnold, *Proof of a theorem by A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian,* Russian Math Surveys 18 (1963), 9-36.

[A-S] M. Abramowitz and I. A. Stegun, *Handbook of mathematical Functions with Formulas, Graphs, and Mathematical Tables,* (Dover, New York, 1965), chapter 22.

[A-W] R. Askey and S. Wainger, *Mean convergence of expansions in Laguerre and Hermite series,* American Journal of Mathematics, 87, (1965) 695–708.

[Bam1] D. Bambusi, *Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations,* Mathematische Zeitschrift 230, (1999), 345-387.

[Bam2] D. Bambusi, *Birkhoff Normal Form for Some Nonlinear PDEs,* Commun. Math. Phys. 234, 253-285, (2003).

[Bar1] V. Bargmann, *On a Hilbert Space of Analytic Functions and an Associated Integral Transform, Part I,* Communication On Pure and Applied mathematics, Vol. 14, (1961), 187-214.

[Bar2] V. Bargmann, *On a Hilbert Space of Analytic Functions and an Associated Integral Transform, Part II. A Family of Related Function Spaces Application to Distribution Theory,* Communication On Pure and Applied mathematics, Vol. 20, (1967), 1-101.

[Bo] R. Bonic, *Some properties of Hilbert Scales,* Proceeding of the American Mathematical Society, Vol. 18, No. 6.(Dec. 1967), p.1000-1003).

[Bu] Ida W. Busbridge, *Some integrals involving Hermite polynomials*, J. London Math. Soc. 23, (1948). 135–141.

[B-D] J. Bonet and P. Domanski. *Real Analytic Curves in Fréchet Spaces and Their Duals,* Mh. Math. 126, 13-36 (1998).

[B-G] D. Bambusi and B. Grébert, *Birkhoff Normal Form for PDEs with Tame Modulus,* Duke Math. J. Volume 135, Number 3 (2006), 507-567.

[B-H-K] G. Benettin, J. Henrard and S.Kuksin, *Hamiltonian Dynamics Theory and Applications,* Lecture Notes in Mathematics 1861.

[B-S-H] C. C. Bradley, C. A. Sackett and R. G. Hulet, *Bose-Einstein condensation of Lithium: observation of limited condensate number*, Phys. Rev. Lett., 78, 985-989, 1997.

[B-S-T-H] C. C. Bradley, C. A. Sackett, J. J. Tollett and R. G. Hulet, *Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions,* Phys. Rev. Lett., 75, 1687-1690, 1995.

[C1] R. Carles, *Remarks on Nonlinear Schrödinger Equations with Harmonic Potential,* Ann. Henri Poincaré 3 (2002), 757-772.

[C2] R. Carles, *Critical Nonlinear Schrödinger Equations with and without Harmonic Potential,* Mathematical Models and Methods in Applied Sciences, Vol 12, (10), 1513-1523, 2002.

[D-S] J.-M. Delort and J. Szeftel, *Long-time existence for small data nonlinear Klein-Gordon equations on tori and spheres,* Int. Math. Res. Not. 2004, no. 37, 1897–1966.

[E] A. Erdélyi, *Über eine erzeugende Funktion von Produkten Hrmitescher Polynome*, Math. Zeitschr. 44 (1938), 201-211.

[E-M-O-T] A. Erdélyi, W. Magnus, F. Oberhettinger and F.G. Tricomi, *Tables of Integral Transforms I, II,* Mcgraw Hill, New York (1954).

[E-K-M-Y] L.H. Eliasson, S.B. Kuksin, S. Marmi and J.-C. Yoccoz, *Dynamical Systems and Small Divisors,* Lecture Notes in Mathematics 1784.

[F1] D. Fujiwara, *A construction of the fundamental solution for the Schrödinger equation*, J. Analyse Math. 35 (1979), 41-96.

[F2] D. Fujiwara, *Remarks convergence of the Feynman path integrals*, Duke Math. J. 47 (1980), 559-600.

[F-H] R.P. Feymann, A.R. Hibbs, *Quantum Mechanics and path integrals,* in International Series in Pure and Applied Physics McGraw-Hill, Maidenhead, 1965, p.365.

[F-N] G. Freud and G. Németh, *On the L_p-norm of orthonormal Hermite functions,* Studia Scientiarum Mathematicarum Hungarica 8 (1973) 399-404.

[Gre] B. Grébert, *Birkhoff Normal Form and Hamiltonian PDEs*, arxiv: math. AP/0604132 v1, 6 Apr 2006.

[Gro] E. P. Gross, *Hydrodynamics of a superfluid condensate*, J. Math. Phys. 4, 195-207 (1963).

[G-V] J. Ginibre and G. Velo, *On a Class of Nonlinear Schrödinger Equations–I The Cauchy Problem, General Cases,* Journal of Functional Analysis 32, 1–32 (1979).

[H] F. Hausdorff, *Ueber das Momentenproblem für ein endliches Interval,* Mathematische Zeitschrift, vol. 16 (1923), pp. 220-248.

[H-P] E Hille, R S Phillips, *Functional Analysis and Semi-Groups.* American Mathematical Society, 1975.

[J] A. Jeffray, *Handbook of mathematical formulas and integrals,* (Amsterdam, Boston, Elsevier Academic Press, 2004).

[Ko] A. N. Kolmogorov, *On the conservation of conditionally periodic motions for a small change in Hamilton's function,* Dokl. Akad. Nauk SSSR 98 (1954), 527-530, (Russian). English translation in: Lectures Notes in Physics 93, Springer, 1979.

[Kw] M. K. Kwong, *Uniqueness of positive solutions of* $\Delta u - u + u^p = 0$ *in* $\mathbb{R}^n$, Arch. Rational Mech. Anal., 105 (3), 243-266, 1989.

[K-M1] A. Kriegl and P. W. Michor, *Aspects of The Theory of Infinite Dimensional Manifolds,* Differential Geometry and its Applications 1 (1991), 159-176.

[K-M2] A. Kriegl and P. W. Michor, *The Convenient Setting For Real Analytic Mappings,* Acta Mathematica 165 (1990), 105-159.

[K-N-S-Q] E. B. Kolomeneisky, T. J. Newman, J. P. Straley, and X. Qi, *Low-dimensional Bose liquids: Beyond the Gross-Pitaevskii approxiamation,* Phys. Rev. Lett., 85 (6), 1146-1149, 2000.

[K-P] S. Kuksin and J. Pöschel, *Invariant Cantor Manifolds of Quasi-Periodic Oscillations for a Nonlinear Schrödinger Equation,* The Annals of Mathematics, 2nd Ser., Vol. 143, No. 1 (1996), 149-179.

[L] P. Lochak, *Canonical Perturbation Theory via Simulataneous Approxima-*

tion, Uspekhi Math. Nauk, 47 (6) 59-140, 1992, English translation: Russ. Math. Surv. 47 57-133 (1992).

[Ma] C. Markett, *Norm Estimates for (C, δ) Means of Hermite Expansions and Bounds for δ_{eff}*, Acta Math. Hung. 43 (3-4) (1984), 187-198.

[Me1] F. Merle, *Determination of blow-up solutions with minimal mass for nonlinear Schrödinger equations with critical power,* Duke Math. J., 69 (2), 427-454, 1993.

[Me2] F. Merle, *Blow-up phenomena for critical nonlinear Schrödinger and Zakharov equations, in Proceedings of the International Congress of Mathematics,* Vol. III (Berlin, 1998), numbers Extra Vol. III, pages 57-66 (electronic), 1998.

[Mi] B.S. Mitjagin, *Approximate dimension and bases in nuclear spaces,* Russian Math. Surveys, London Math. Soc. 18 (1963), 59-127.

[Mo] J. Moser, *On invariant curves of area preserving mappings of an annulus,* Nachr. Akad. Wiss. Gött., Math. Phys. K1, (1962), 1-20.

[M-O-S] W. Magnus, F. Oberhettinger and R. P. Soni, *Formulas and theorems for the special functions of mathematical physics*, (Berlin, New York 1966).

[M-R1]F. Merle and P. Raphael, *Sharp upper bound on the blow-up rate for the critical nonlinear Schrödinger equation*, Geom. Funct. Anal. 13 (2003), no. 3, 591–642.

[M-R2] F. Merle and P. Raphael, *On a Sharp Lower Bound on the Blow-up Rate for the L^2 Critical Nonlinear Schrödinger Equation,* Journal of the American mathematical society, Vol. 19, No. 1, Pages 37-90, 2005.

[N] N. Nekhoroshev, *An exponential estimate of the time of stability of nearly integrable Hamiltonian system,* Russ. Math. Surveys 32 (1977), 1-65.

[O] Y. G. Oh, *Cauchy problem and Ehrenfest's law of nonlinear Schrödinger equations with potentials,* J. Differential Equations 81 (1989), no. 2, 255–274.

[Pi] L. P. Pitaevskii, *Vortex lines in an imparfect Bose gas*, Sov. Phys. JETP 13, 451-454 (1961).

[Po] J. Pöschel, *On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem by Bambusi*, Nonlinearity 12 (1999), 1587-1600.

[Q-X-W] M. Qi, C. Xu and W. Wang, *An introduction to Mordern Partial differential equations (in Chinese)*, Wuhan University publishing company, 1994.

[R-S1] M. Reed and B. Simon. *Methods of Modern Mathematical Physics,* Academic press, New York and London, vol. I, 1972.

[R-S2] M. Reed and B. Simon. *Methods of Modern Mathematical Physics,* Academic press, New York and London, vol. 2, 1975.

[Si] B. Simon, *Distributions and Their Hermite Expansions,* Journal of Mathematical Physics, Vol. 12, No. 1 (1971), p.140-148.

[Sc] L. Schwartz, *Théorie des distributions,* (Hermann, Paris, 1966), rev. ed.

[T] M. E. Taylor, *Partial Differential Equations I,II,III,* Applied mathematical sciences (Springer-Vertag New York Inc.) v.115-117.

[Webpage1] http://www.math.ucla.edu/~tao/Dispersive/, maintained by Jim Colliander, Mark Keel, Gigliola Staffilani, Hideo Takaoka, and Terry Tao.

[Webpage2] http://en.wikipedia.org/wiki/Hermite_polynomials

[Z] J. Zhang, *Stability of attractive Bose-Einstein condensates,* J. Statist. Phys., 101 (3-4), 731-746, 2000.

4659 03