Stochastic Quantization of Abelian Higgs Model

BY Chunqiu Song

Born in July 14th 1994

Email: s6chsong@uni-bonn.de

Master Thesis of Mathematics

Advisor: Prof. Massimiliano Gubinelli

Second Advisor: Dr. Francesco de Vecchi

INSTITUTE FOR APPLIED MATHEMATICS

Mathematisch-Naturwissenschaftliche Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Table of contents

1 Introduction

The purpose of this thesis is to study the problems behind the stochastic quantization of abelian Higgs model, which is the following system of stochastic partial differential equations

$$\begin{aligned}
\partial_t A_1 &= \partial_2^2 A_1 - \partial_1\partial_2 A_2 - \frac{ie}{2}[\bar{\Phi}\partial_1\Phi - \Phi\partial_1\bar{\Phi}] - e^2 A_1 \Phi\bar{\Phi} + \xi_1 \\
\partial_t A_2 &= \partial_1^2 A_2 - \partial_1\partial_2 A_1 - \frac{ie}{2}[\bar{\Phi}\partial_2\Phi - \Phi\partial_2\Phi] - e^2 A_2 \Phi\,\bar{\Phi} + \xi_2 \\
\partial_t \Phi &= \partial_1^2\Phi + \partial_2^2\Phi - ie\,(\partial_1 A_1 + \partial_2 A_2)\Phi - 2ie\,(A_1\partial_1 + A_2\partial_2)\Phi \\
&\quad - e^2(A_1^2 + A_2^2)\Phi + \zeta
\end{aligned}$$

which is studied in the paper [41]. These equations are motivated by the programs of constructive quantum field theory [21], which try to construct probability measures on the space of distributions that satisfy the Osterwalder-Schrader axioms of Euclidean quantum field theory, and then produce a theory in Minkowski time through the reconstruction theorems [44] [35] [36]. The stochastic quantization method proposed by Parisi and Wu [37] is one of such approaches to produce a candidate of probability measure [14]. The equations given by stochastic quantization method are usually SPDEs, which is in general hard to interpret and solve. The recent theory of paracontrolled analysis developed by Gubinelli, Imkeller and Perkowski [23], and theory of regularity structures developed by Hairer [25] which is used in [41], are powerful tools to tackle such problems. We will introduce both the subject of SPDEs within the framework of paracontrolled analysis, and the axiomatic quantum field theory. The structure of this thesis is following.

Chapter 2 contains a short introduction of theory of tempered distributions, invented by Laurent Schwartz. Distributions are needed when we try to describe singular objects. In general, random objects are distributions and stochastic differential equations are equations of distributions. We study how to do operations on tempered distributions, such as transformations, Fourier transformations, convolutions, and differentiations, etc. Theorems about tempered distributions with compact supports are discussed.

Chapter 3 deals with the subject of paracontrolled analysis. The central difficulty in the study of SPDEs is to interpret the nonlinear functions of tempered distributions, in particular the product of tempered distributions. We start by introducing a way to measure the singular behavior of tempered distributions, the so called Besov space is discussed. Then we introduce Bony's paraproduct, from which we can define products of tempered distributions and separate out the singular part. After that we develop the first order paracontrolled calculus.

Chapter 4 is about the white noise and Gaussian analysis. In most stochastic partial differential equations the random force term are usually given by white noise due to the random nature of background. We talks about the Wick product of Gaussian random variables and an important estimate of Gaussian variables, called the Gaussian hypercontractivity. Then we find out the regularity exponent of a white noise.

Chapter 5 is an application of all the machinery we have developed so far to a particular stochastic partial differential equation, called the parabolic Anderson model. We first discuss heuristically how to set up the equations that one can apply use fix point argument. Then we produce the space where we want to find the solution and the Schauder estimates. Finally we sketch ideas in proving the existence and uniqueness of the solution, and discuss about the renormalization.

Chapter 6 is devoted to the subject of axiomatic quantum field theory. We first talk about what is the correct mathematical object to model a quantum field. We argue from both mathematics and physics that one can not define quantum field by assigning an operator to each point of space-time. Quantum field has to be averaged by some good functions over some space-time regions. Then we come to the concept of operator-valued distributors. Then we study the Lorentz group and their unitary representations, which is needed in describing relativistic symmetry in quantum theory. After that we talk about the Wightman axioms of fields, the vacuum correlation functions and their properties. The distributions with these properties are called Wightman distributions and a Wightman field theory can be reconstructed from these correlation functions. We then move on to the analytic continuation of there Wightman distributions to get Schwinger functions, and a system of properties of such functions can be obtained. One can add some assumptions and assuming these properties to get Osterwalder-Schrader axioms, then it can be shown that Wightman axioms can be recovered from these axioms.

Chapter 7 concerns the stochastic quantization of the abelian Higgs model. We first introduce the method of stochastic quantization. In order to find the Euclidean path integral measure of a quantum field theory, one can go to one higher dimension to study a non-equilibrium process described by Langevin dynamics. One hope that the stationary measure can be obtained by solving such equations at infinite time. Then we talk about the abelian Higgs model, including how the equations is calculated, what we have tried to deal with it and problems.

Acknowledgment I would like to express my deepest appreciation to my advisor Prof. Massimiliano Gubinelli for his kindness and patience to have discussions, and for everything he taught me, including many great lectures in probability, stochastic analysis and modern advanced topics. I would like to extend my sincere thanks to Dr. Francesco de Vecchi for his numerous help and discussion. I am also grateful to many of my friends and classmates, for their moral support and help. Finally I would like to thank my family for their constant emotional, moral and economical support.

2 Distributions

This chapter is devoted to the study of tempered distributions. The motivation originates from making a rigorous understanding of Dirac delta function, and many singular functions appeared in the calculations of quantum field theory. The main references for this chapter are [20], [4] and [39].

2.1 The Fourier Transform on Schwartz Space

The space $\mathcal{S}(\mathbb{R}^n)$ of Schwartz functions on $\mathbb{R}^n$ consists of all smooth functions whose derivatives fall off faster than any reciprocal power of polynomials. More precisely, $f\in\mathcal{S}(\mathbb{R}^n)$ if and only if $f\in C^\infty(\mathbb{R}^n)$, and for any $k\in\mathbb{N}$, the following quantities

$$\|f\|_k := \sup_{|\alpha|\leqslant k, |\beta|\leqslant k, x\in\mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty$$

are finite, where $\alpha=(\alpha_1,\ldots,\alpha_n)\in\mathbb{N}^n$, $\beta=(\beta_1,\ldots,\beta_n)\in\mathbb{N}^n$ are multi-indices with length $|\alpha|=\alpha_1+\cdots+\alpha_n$, $|\beta|=\beta_1+\cdots+\beta_n$, and $x^\alpha=x_1^{\alpha_1}\times\cdots\times x_n^{\alpha_n}$, $\partial^\beta=\partial_1^{\beta_1}\ldots\partial_n^{\beta_n}$.

Clearly $\mathcal{S}(\mathbb{R}^n)$ is a vector space, for each $k\in\mathbb{N}$, $\|\cdot\|_k$ defines a norm on $\mathcal{S}(\mathbb{R}^n)$, hence $\mathcal{S}(\mathbb{R}^n)$ is a countably normed space. The topology is then defined by giving the neighborhoods, and it is this topology we are mostly interested in. So there are other equally good families of norms which defines the same topology, for example

$$\|f\|_k' = \sup_{|\alpha|\leqslant k, x\in\mathbb{R}^n} (1+|x|^k)|\partial^\alpha f(x)|$$

gives another choice. Moreover, one can show that $\mathcal{S}(\mathbb{R}^n)$ is indeed a Frechet space, which means this countably normed space is complete.

For any $f\in\mathcal{S}(\mathbb{R}^n)$, the Fourier transform $\mathcal{F}(f)$ of f is defined as

$$\mathcal{F}(f)(\xi)=\hat{f}(\xi):=\int_{\mathbb{R}^n} f(x)e^{-i\langle\xi,x\rangle}\,dx$$

Theorem 2.1. *The map $\mathcal{F}:\mathcal{S}(\mathbb{R}^n)\to\mathcal{S}(\mathbb{R}^n)$ is bijective and continuous. More precisely, for any given $k\in\mathbb{N}$, there exists a constant C and an integer $K\in\mathbb{N}$ such that*

$$\|\hat{f}\|_k\leqslant C\|f\|_K$$

for all $f\in\mathcal{S}(\mathbb{R}^n)$. The inverse of the map is given by

$$\mathcal{F}^{-1}(f)(x):=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(\xi)e^{i\langle\xi,x\rangle}\,d\xi$$

for any $f\in\mathcal{S}(\mathbb{R}^n)$.

Proof. For any multi-indices α and β such that $|\alpha|, |\beta| \leqslant k$, we have

$$\begin{aligned} |\xi^\alpha \partial^\beta \hat{f}(\xi)| &= |\mathcal{F}(\partial^\alpha (x^\beta f))(\xi)| \\ &\leqslant \|\partial^\alpha (x^\beta f)\|_{L^1} \\ &= \int_{\mathbb{R}^n} (1+|x|)^{n+1} \cdot |\partial^\alpha (x^\beta f)(x)| \cdot \frac{1}{(1+|x|)^{n+1}} d^n x \\ &\leqslant c_n \|(1+|x|)^{n+1} \partial^\alpha (x^\beta f)\|_{L^\infty} \\ &\leqslant C \|f\|_{k+n+1} \end{aligned}$$

where the last inequality follows an estimation of the expansion of the term in the L^∞-norm. Thus we have

$$\|\hat{f}\|_k \leqslant C \|f\|_{k+n+1}$$

by definition. □

We define the convolution $f * g$ of two Schwartz functions $f, g \in \mathcal{S}(\mathbb{R}^n)$ to be the function

$$(f * g)(x) := \int_{\mathbb{R}^n} f(x-y) g(y) dy$$

This operation has following simple properties.

Lemma 2.2. *We have*

(1) $f * g = g * f$ and $f * g \in \mathcal{S}(\mathbb{R}^n)$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$;

(2) for fixed $g \in \mathcal{S}(\mathbb{R}^n)$, the map defined by $f \mapsto g * f$, for all $f \in \mathcal{S}(\mathbb{R}^n)$, is a continuous map from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$;

(3) $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$ and $\mathcal{F}(fg) = \mathcal{F}(f) * \mathcal{F}(g)$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$;

(4) $(f * g) * h = f * (g * h)$ for all $f, g, h \in \mathcal{S}(\mathbb{R}^n)$.

Proof.

(1) Since the function $f(x-y)$, as a function of y, is obtained by applying translation $f(y) \to f(y-x)$ and reflection of the function $f(y-x) \to f(x-y)$, thus the resulting function is also a Schwartz function. Then the formula

$$\int_{\mathbb{R}^n} f(x-y) g(y) dy$$

is the L^2-inner product(we use the convention that the inner product is conjugate linear in the first vector, linear in the second) of two Schwartz functions, we can apply the Fourier transform to each function which preserves this L^2-inner product up to a constant, so

$$\begin{aligned} (f * g)(x) &= \int_{\mathbb{R}^n} \widehat{f(x-\cdot)}(\xi) \widehat{g(\cdot)}(\xi) d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left[\overline{\int_{\mathbb{R}^n} f(x-y) e^{-i\langle \xi, y \rangle} dy} \int_{\mathbb{R}^n} g(h) e^{-i\langle \xi, h \rangle} dh \right] d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y) e^{-i\langle \xi, x-y \rangle} dy \int_{\mathbb{R}^n} g(h) e^{-i\langle \xi, h \rangle} dh \right] e^{i\langle \xi, x \rangle} d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{f(\cdot)}(\xi) \widehat{g(\cdot)}(\xi) e^{i\langle \xi, x \rangle} d\xi \end{aligned}$$

which clearly shows that $f*g\in\mathcal{S}(\mathbb{R}^n)$. The equation $f*g=g*f$ is an easy consequence of changing the variable of the integration.

(2) According to our computation above, we have

$$g*f=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\widehat{g(\cdot)}(\xi)\widehat{f(\cdot)}(\xi)e^{i\langle\xi,\cdot\rangle}d\xi=\mathcal{F}^{-1}(\mathcal{F}(g)\mathcal{F}(f))$$

which is clearly continuous about f, since the Fourier transform and its inverse are continuous by theorem 2.1.

(3) This is clear from previous computation.

(4) We have

$$\begin{aligned}(f*g)*h &= \mathcal{F}^{-1}(\mathcal{F}(f*g)\mathcal{F}(h))\\ &= \mathcal{F}^{-1}(\mathcal{F}(f)\mathcal{F}(g)\mathcal{F}(h))\\ &= \mathcal{F}^{-1}(\mathcal{F}(f)\mathcal{F}(g*h))\\ &= f*(g*h)\end{aligned}$$

which concludes the proof. □

2.2 Tempered Distributions

A tempered distribution F on $\mathbb{R}^n$ is a continuous linear functional on the Schwartz space $\mathcal{S}(\mathbb{R}^n)$. The continuity of the linear map $u:\mathcal{S}(\mathbb{R}^n)\to\mathbb{R}$ can be characterized by existence of constant C and an integer k such that the inequality $|u(f)|\leqslant C\|f\|_k$ holds for all $f\in\mathcal{S}(\mathbb{R}^n)$. We also use the notation $u(f)=\langle u,f\rangle$

The set of all tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$. Clearly it is a vector space, and the natural topology on $\mathcal{S}'(\mathbb{R}^n)$ is the weak topology. Equivalently the topology is given by giving the notion of limit, a sequence of distributions $\{u_n\}_{n\in\mathbb{N}}$ is said to converge to the limit $u\in\mathcal{S}'(\mathbb{R}^n)$ iff for any $f\in\mathcal{S}(\mathbb{R}^n)$, we have $\lim_{n\to\infty}u_n(f)=u(f)$.

Example 2.3. The map $f\mapsto f(x)$ for some fixed $x\in\mathbb{R}^n$ defines a tempered distribution on $\mathcal{S}(\mathbb{R}^n)$. More generally, the map $f\mapsto\partial^\alpha f(x)$ for fixed $x\in\mathbb{R}^n$ and $\alpha\in\mathbb{N}^n$ defines a tempered distribution on $\mathcal{S}(\mathbb{R}^n)$.

Example 2.4. The space $L^p(\mathbb{R}^n)$ for all $p\in[1,+\infty]$ is contained in the space of tempered distribution $\mathcal{S}'(\mathbb{R}^n)$, each $u\in L^p(\mathbb{R}^n)$ is identified with the map

$$f\mapsto\int_{\mathbb{R}^n}u(x)f(x)dx$$

A tempered distribution $u\in\mathcal{S}'(\mathbb{R}^n)$ vanishes on an open set $U\subset\mathbb{R}^n$ if $\langle u,f\rangle=0$ for any $f\in\mathcal{S}(\mathbb{R}^n)$ with $\operatorname{supp}(f)\subset U$. The support of a tempered distribution $u\in\mathcal{S}'(\mathbb{R}^n)$ is complement of the union of those open sets where u vanishes, thus the support is always a closed set.

A continuous linear map $A: \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a linear map such that for all $k \in \mathbb{N}$, there exists a constant C and an integer K such that the inequality $\|A(f)\|_k \leqslant C\|f\|_K$ holds for all $f \in \mathcal{S}(\mathbb{R}^n)$. To define operators on the space of tempered distributions out of some given continuous linear map on the Schwartz space, we use duality.

Theorem 2.5. *Suppose $A: \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a continuous linear map, then the map $A^T: \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ defined by*

$$\langle A^T(u), f \rangle := \langle u, A(f) \rangle$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$, is linear and continuous. Here the continuity means for any sequence $\{u_n\}_{n \in \mathbb{N}} \subset \mathcal{S}'(\mathbb{R}^n)$ with limit $u \in \mathcal{S}'(\mathbb{R}^n)$, the sequence $\{A^T(u_n)\}_{n \in \mathbb{N}}$ has limit $A^T(u)$.

Proof. We first show that $A^T(u)$ is a tempered distribution, which means $A^T(u)$ is a continuous linear functional on $\mathcal{S}(\mathbb{R}^n)$. Linearity is clear from the definition. We proof the continuity.

Since $u \in \mathcal{S}'(\mathbb{R}^n)$, there exists a constant C and an integer k, such that

$$|u(g)| \leqslant C\|g\|_k$$

for all $g \in \mathcal{S}(\mathbb{R}^n)$.

Since $A: \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a continuous linear map, then there exists a constant C' and an integer K such that the inequality

$$\|A(f)\|_k \leqslant C'\|f\|_K$$

holds for all $f \in \mathcal{S}(\mathbb{R}^n)$.

Combine there two inequalities, we have

$$|u(A(f))| \leqslant C\|A(f)\|_k \leqslant C \cdot C'\|f\|_K$$

holds for all $f \in \mathcal{S}(\mathbb{R}^n)$.

Hence $A^T(u)$ is indeed a tempered distribution. The second statement follows directly from the definition of the convergence of a sequence of tempered distribution. □

Example 2.6.

1. The differential operator $(-\partial)^\alpha$ with $\alpha \in \mathbb{N}^n$, acting on a Schwartz function by taking derivatives, is clearly a continuous linear map from $\mathcal{S}(\mathbb{R}^n)$ to itself, this defines the differential operator ∂^α on tempered distributions.

2. Given a smooth function $h \in C^\infty(\mathbb{R}^n)$ whose partial derivatives of any order is at most polynomial growth, that is for each $\alpha \in \mathbb{N}^n$ there is a constant C and a polynomial function $P_\alpha(x)$ such that $|\partial^\alpha h(x)| \leqslant C \cdot P_\alpha(x)$ for all $x \in \mathbb{R}^n$. Clearly multiplying Schwartz function by h is a continuous linear map from $\mathcal{S}(\mathbb{R}^n)$ to itself, this defines the multiplication operator h on tempered distributions.

3. Suppose $L \in \mathrm{GL}(\mathbb{R}^n)$, then the map

$$\mathcal{S}(\mathbb{R}^n) \ni f(\cdot) \mapsto \frac{1}{\det(L)} f(L^{-1}(\cdot))$$

is a continuous linear map from $\mathcal{S}(\mathbb{R}^n)$ to itself, this defines coordinate change $L(\cdot)$ of tempered distributions. More generally, one can define coordinate change with respect to other elements in the diffeomorphism group of $\mathbb{R}^n$. This operator also makes it possible to talk about invariant distribution, for example rotation invariant or Lorentz invariant.

4. Given a fixed function $f \in \mathcal{S}(\mathbb{R}^n)$ and denote $f^-(x) = f(-x)$, then the map

$$\mathcal{S}(\mathbb{R}^n) \ni g \mapsto f^- * g$$

is clearly a continuous linear map from $\mathcal{S}(\mathbb{R}^n)$ to itself by lemma 2.2. This gives how to convolute a tempered distribution with a Schwartz function f.

5. We know from Theorem 2.1 that Fourier transform is a continuous linear map from $\mathcal{S}(\mathbb{R}^n)$ to itself, then this defines the Fourier transform $\mathcal{F}$ on the space of tempered distribution. In the same way, we can define the inverse Fourier transform $\mathcal{F}^{-1}$ on the space of tempered distribution, and we will prove later they are indeed inverse of each other.

Just like the case of Fourier transform on the Schwartz space, we have following formulas.

Lemma 2.7. *For any $u \in \mathcal{S}'(\mathbb{R}^n)$ and $f \in \mathcal{S}(\mathbb{R}^n)$, we have*

$$\begin{aligned} (i\partial)^\alpha \mathcal{F}(u) &= \mathcal{F}(x^\alpha u) \\ (i\xi)^\alpha \mathcal{F}(u) &= \mathcal{F}(\partial^\alpha u) \\ \mathcal{F}(f * u) &= \mathcal{F}(f)\,\mathcal{F}(u) \\ \mathcal{F}(fu) &= \mathcal{F}(f) * \mathcal{F}(u) \end{aligned}$$

Proof. These identities are easy to check. □

Fourier transform is a continuous linear isomorphism on the space of tempered distribution.

Theorem 2.8. *The Fourier transform $\mathcal{F}$ from $\mathcal{S}'(\mathbb{R}^n)$ to itself is linear, bijective and continuous with continuous linear inverse $\mathcal{F}^{-1}$.*

Proof. It is an easy consequence of theorem 2.1. □

It can be shown that any tempered distribution is a derivative of some continuous function, the precise statement is given bellow.

Theorem 2.9. *Suppose $u \in \mathcal{S}'(\mathbb{R}^n)$, then there is a polynomially bounded continuous function f, that is*

$$|f(x)| \leqslant C(1 + \|x\|^2)^m$$

for some constant $C \in \mathbb{R}_+$ $m \in \mathbb{N}$, such that

$$u = \partial^\alpha f$$

for some multi index α, where we think of f as a tempered distribution and then take the partial derivatives.

The proof can be found in [39], we omit it here.

2.3 Convolution of Distributions

Taking the convolution of a Schwartz function and a tempered distribution is an important operation, it gives a way to regularize a tempered distribution, which is not regular enough in general.

Theorem 2.10.

(1) For a fixed function $f\in\mathcal{S}(\mathbb{R}^n)$, the map defined by $u\mapsto f*u$ for any $u\in\mathcal{S}'(\mathbb{R}^n)$, is a continuous map.

(2) We have following identities:

$$\partial^\alpha(f*u)=(\partial^\alpha f)*u=f*(\partial^\alpha u)$$

$$(f*g)*u=f*(g*u)$$

$$\mathcal{F}(f*u)=\mathcal{F}(f)\,\mathcal{F}(u)$$

$$\mathcal{F}(fu)=\mathcal{F}(f)*\mathcal{F}(u)$$

for any $f,g\in\mathcal{S}(\mathbb{R}^n)$ and $u\in\mathcal{S}'(\mathbb{R}^n)$.

(3) For any $f\in\mathcal{S}(\mathbb{R}^n)$ and $u\in\mathcal{S}'(\mathbb{R}^n)$, the convolution $f*u$ is a smooth function such that for any multi index α, we have

$$|\partial^\alpha(f*u)(x)|\leqslant C(1+\|x\|^2)^m$$

for some constant $C\in\mathbb{R},m\in\mathbb{N}$.

Proof.

(1) For any sequence $\{u_n\}\subset\mathcal{S}'(\mathbb{R}^n)$ with limit $u\in\mathcal{S}'(\mathbb{R}^n)$, and any $h\in\mathcal{S}(\mathbb{R}^n)$, we have

$$\langle f*u_n,h\rangle=\langle u_n,f^-*h\rangle\to\langle u,f^-*h\rangle=\langle f*u,h\rangle$$

which concludes that the convolution map with fixed Schwartz function f is a continuous map from $\mathcal{S}'(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$.

(2) These identities are easy to check.

(3) By theorem 2.9 we know that here is some continuous function h such that $u=\partial^\alpha h$ for some multi index α, and

$$|h(x)|\leqslant C(1+\|x\|^2)^m$$

for some constant $C\in\mathbb{R}_+$ $m\in\mathbb{N}$. Then we have

$$f*u=f*\partial^\alpha h=(\partial^\alpha f)*h$$

which is a convolution of two functions, and

$$\begin{aligned}(\partial^\alpha f)*h(x) &= \int_{\mathbb{R}^n}\partial^\alpha f(x-y)\cdot h(y)dy\\ &= \int_{\mathbb{R}^n}[(1+\|y\|^2)^m\partial^\alpha f(x-y)]\cdot\frac{h(y)}{(1+\|y\|^2)^m}dy\end{aligned}$$

which shows $f*u$ is a smooth function since the first term is again a Schwartz function and the second term is a bounded continuous function. Moreover

$$\begin{aligned}|f*u(x)| &= \left|\int_{\mathbb{R}^n}[(1+\|y\|^2)^m\partial^\alpha f(x-y)]\cdot\frac{h(y)}{(1+\|y\|^2)^m}dy\right|\\ &\leqslant \int_{\mathbb{R}^n}|(1+\|y\|^2)^m\partial^\alpha f(x-y)|\cdot\frac{|h(y)|}{(1+\|y\|^2)^m}dy\\ &\leqslant C\int_{\mathbb{R}^n}|(1+\|y\|^2)^m\partial^\alpha f(x-y)|dy\\ &= C\int_{\mathbb{R}^n}(1+\|x-y\|^2)^m|\partial^\alpha f(y)|dy\end{aligned}$$

which is clearly polynomially bounded since $\partial^\alpha f$ is a Schwartz function. □

Sometimes we need to approximate a tempered distribution, we introduce here the concept of approximate identity.

Definition 2.11. *Suppose we have a positive smooth function φ, whose support is contained in the centered unit ball $B\subset\mathbb{R}^n$, and*

$$\int_{\mathbb{R}^n}\varphi(x)dx=1$$

Then the sequence of functions $\left\{\varphi_\varepsilon:=\varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right)\right\}$ parametrized by positive real number ε, is called an approximate identity.

Clearly this sequence tends to Dirac delta function as $\varepsilon\to 0$. We have following result.

Theorem 2.12. *Suppose φ_ε is an approximate identity and $u\in\mathcal{S}'(\mathbb{R}^n)$, then we have $\varphi_\varepsilon*u\to u$ in the space $\mathcal{S}'(\mathbb{R}^n)$ as $\varepsilon\to 0$.*

Proof. We need to show that for any Schwartz function f, we have

$$\langle\varphi_\varepsilon*u,f\rangle\to\langle u,f\rangle$$

as $\varepsilon\to 0$. Since

$$\langle\varphi_\varepsilon*u,f\rangle = \langle u,\varphi_\varepsilon^-*f\rangle$$

then we only need to show $\varphi_\varepsilon^-*f\to f$ in the space $\mathcal{S}(\mathbb{R}^n)$. Since Fourier transform is a continuous isomorphism from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$, we only need to show the limit $\mathcal{F}(\varphi_\varepsilon^-)\mathcal{F}(f)\to\mathcal{F}(f)$ in the space $\mathcal{S}(\mathbb{R}^n)$. Since

$$\begin{aligned}\mathcal{F}(\varphi_\varepsilon^-)(\xi) &= \int_{\mathbb{R}^n}\varepsilon^{-n}\varphi\left(\frac{-x}{\varepsilon}\right)e^{-i\langle\xi,x\rangle}dx\\ &= \int_{\mathbb{R}^n}\varphi(x)e^{i\langle\varepsilon\xi,x\rangle}dx\\ &= \int_B\varphi(x)e^{i\langle\varepsilon\xi,x\rangle}dx\end{aligned}$$

then

$$\mathcal{F}(\varphi_\varepsilon^-)\,\mathcal{F}(f)-\mathcal{F}(f) \;=\; \int_B \varphi(x)(e^{i\langle\varepsilon\xi,x\rangle}-1)\,dx\mathcal{F}(f)$$

and

$$\|\mathcal{F}(\varphi_\varepsilon^-)\,\mathcal{F}(f)-\mathcal{F}(f)\|_k \;=\; \sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\left|\xi^\alpha\partial^\beta\Big[\int_B \varphi(x)(e^{i\langle\varepsilon\xi,x\rangle}-1)\,dx\mathcal{F}(f)(\xi)\Big]\right|$$

To show the right hand side tends to 0, it suffice to notice that when we use Leibniz rule to compute the term

$$\partial^\beta\Big[\int_B \varphi(x)(e^{i\langle\varepsilon\xi,x\rangle}-1)\,dx\mathcal{F}(f)(\xi)\Big]$$

one case is there is some partial derivatives act on the first term, for example

$$\begin{aligned}
&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\left|\xi^\alpha\int_B \varphi(x)[\partial^{\beta_1}(e^{i\langle\varepsilon\xi,x\rangle}-1)]\,dx\,(\partial^{\beta-\beta_1}\mathcal{F}(f))(\xi)\right|\\
=\;&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\left|\xi^\alpha\int_B \varphi(x)\xi^{\beta_1}x^{\beta_1}(\varepsilon i)^{|\beta_1|}e^{i\langle\varepsilon\xi,x\rangle}\,dx\,(\partial^{\beta-\beta_1}\mathcal{F}(f))(\xi)\right|\\
=\;&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\varepsilon^{|\beta_1|}\left|\int_B \varphi(x)x^{\beta_1}e^{i\langle\varepsilon\xi,x\rangle}dx\,\xi^\alpha\xi^{\beta_1}(\partial^{\beta-\beta_1}\mathcal{F}(f))(\xi)\right|\\
\leqslant\;&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\varepsilon^{|\beta_1|}\int_B \varphi(x)|x^{\beta_1}|dx\,|\xi^\alpha\xi^{\beta_1}(\partial^{\beta-\beta_1}\mathcal{F}(f))(\xi)|
\end{aligned}$$

which clearly tends to 0 as $\varepsilon\to 0$, since $\xi^\alpha\xi^{\beta_1}\partial^{\beta-\beta_1}\mathcal{F}(f)$ is a Schwartz function, another case is the following

$$\begin{aligned}
&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\left|\int_B \varphi(x)(e^{i\langle\varepsilon\xi,x\rangle}-1)\,dx\xi^\alpha(\partial^\beta\mathcal{F}(f))(\xi)\right|\\
\leqslant\;&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\int_B \varphi(x)|e^{i\langle\varepsilon\xi,x\rangle}-1|dx\,|\xi^\alpha(\partial^\beta\mathcal{F}(f))(\xi)|\\
=\;&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\int_B \varphi(x)\left|\int_0^{\langle\varepsilon\xi,x\rangle} e^{i\lambda}\,d\lambda\right|dx\,|\xi^\alpha(\partial^\beta\mathcal{F}(f))(\xi)|\\
\leqslant\;&\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\int_B \varphi(x)\left|\int_0^{\langle\varepsilon\xi,x\rangle} |e^{i\lambda}|\,d\lambda\right|dx\,|\xi^\alpha(\partial^\beta\mathcal{F}(f))(\xi)|\\
=\;&\varepsilon\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\int_B \varphi(x)|\langle\xi,x\rangle|dx\,|\xi^\alpha(\partial^\beta\mathcal{F}(f))(\xi)|\\
\leqslant\;&\varepsilon\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\int_B \varphi(x)\|x\|dx\,\|\xi\||\xi^\alpha(\partial^\beta\mathcal{F}(f))(\xi)|\\
=\;&\varepsilon\int_B \varphi(x)\|x\|dx\sup_{|\alpha|\leqslant k,|\beta|\leqslant k,\xi\in\mathbb{R}^n}\|\xi\||\xi^\alpha(\partial^\beta\mathcal{F}(f))(\xi)|
\end{aligned}$$

which clearly tends to 0 as $\varepsilon\to 0$. And this concludes the proof of the theorem. □

2.4 Compact Supported Tempered Distribution

A tempered distribution whose Fourier transform is compactly supported behaves good, see book [15].

Theorem 2.13. *Suppose* $u \in \mathcal{S}'(\mathbb{R}^n)$ *and* $\mathcal{F}(u)$ *has compact support, then there exists a smooth function* $\tilde{u}$ *and constants* $C_\alpha \in \mathbb{R}$, $m \in \mathbb{N}$, *which depend on the multi index* α, *with*

$$|\partial^\alpha \tilde{u}(x)| \leqslant C_\alpha (1 + \|x\|^2)^m$$

for all $x \in \mathbb{R}^n$, *such that*

$$\langle u, f \rangle = \int_{\mathbb{R}^n} \tilde{u}(x) f(x) dx$$

Proof. Since $\mathcal{F}(u)$ has compact support, then we can find a compact supported smooth function ρ, whose value equals 1 on the set $\operatorname{supp}(\mathcal{F}(u))$, and clearly we have $\mathcal{F}(u) = \rho \mathcal{F}(u)$. Thus

$$u = \mathcal{F}^{-1}(\rho) * u$$

and notice that ρ is also a Schwartz function, so is $\mathcal{F}^{-1}(\rho)$, we can use theorem 2.10 to conclude that u is given by a smooth function whose derivatives are polynomially bounded. □

3 Besov Spaces and Paracontrolled Calculus

The first difficulty confronted in the study of stochastic partial differential equations, is to find a precise way to understand functions of, or product of the irregular terms appearing in a stochastic partial differential equation. The motivation also arise in quantum field theory, where most people believe that the infinities arise in many calculations are due to multiplying distributions incorrectly. In this chapter we introduce the Besov spaces, which offers a way to measure the regularity of a tempered distribution, and the paracontrolled calculus, which offers a method to manipulate the calculus of these irregular objects.

3.1 Littlewood-Paley Theory

To define the concept of Besov spaces, we need the smooth dyadic partition of unity.

Definition 3.1. *A smooth dyadic partition of unity consists two smooth radial functions $\rho_{-1}, \rho_0 \in C_c^\infty$ that take values in the interval $[0,1]$, where ρ_{-1} is supported in the ball $B = \{x \in \mathbb{R}^n: |x| \leqslant R\}$ and ρ_0 is supported in the annulus $A = \{x \in \mathbb{R}^n: 0 < r_1 \leqslant |x| \leqslant r_2, \text{with}\, r_1 < r_2\}$ for some suitably chosen constants $r_1, r_2, R > 0$, such that:*

1. *for each $j \in \mathbb{N}$, define functions ρ_j by $\rho_j(x) := \rho_0\left(\frac{x}{2^j}\right)$, we have*

$$\sum_{j=-1}^{\infty} \rho_j(x) = 1$$

for all $x \in \mathbb{R}^n$;

2. $\operatorname{supp}(\rho_i) \cap \operatorname{supp}(\rho_j) = \varnothing$*, for all* $|i - j| > 1$.

The existence of dyadic partition of unity can be found in the book [4]. For a given dyadic partition of unity, we have the definition of Littlewood-Paley decomposition of tempered distribution as follows.

Definition 3.2. *For any tempered distribution $u \in \mathcal{S}'(\mathbb{R}^n)$ and integer $j \geqslant -1$, the j-th Littlewood-Paley block of u is defined by*

$$\Delta_j u = \mathcal{F}^{-1}(\rho_j \hat{u})$$

thus $\Delta_j: \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is a continuous operator and we have the Littlewood-Paley decomposition

$$\mathrm{Id} = \sum_{j=-1}^{\infty} \Delta_j$$

For convenience, we assume $\Delta_j = 0$ for $j \leqslant -2$. Since $\mathcal{F}^{-1}(\rho_j \hat{u}) = \mathcal{F}^{-1}(\rho_j) * u$, we denote $K_j = \mathcal{F}^{-1}(\rho_j)$.

Lemma 3.3. *For $j \geqslant 0$, $\|K_j\|_{L^1} = \|K_0\|_{L^1}$.*

Proof. We have for $j \geqslant 0$

$$\begin{aligned} K_j(x) &= \mathcal{F}^{-1}(\rho_j)(x) \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \rho_0\left(\frac{\xi}{2^j}\right) e^{i\langle \xi, x\rangle} \, d\xi \\ &\overset{\xi = 2^j y}{=\!=\!=} \frac{2^{jn}}{(2\pi)^n} \int_{\mathbb{R}^n} \rho_0(y) e^{i\langle y, 2^j x\rangle} \, dy \\ &= 2^{jn} K_0(2^j x) \end{aligned}$$

So

$$\begin{aligned} \int_{\mathbb{R}^n} |K_j(x)| dx &= \int_{\mathbb{R}^n} 2^{jn} |K_0(2^j x)| dx \\ &= \int_{\mathbb{R}^n} |K_0(2^j x)| d 2^j x \\ &= \int_{\mathbb{R}^n} |K_0(x)| dx \end{aligned}$$

which is exactly $\|K_j\|_{L^1} = \|K_0\|_{L^1}$ for $j \geqslant 0$. □

Lemma 3.4. *Suppose $u \in L^p \subset \mathcal{S}'$ with $p \in [1, \infty]$, then*

$$\|\Delta_j u\|_{L^p} \leqslant \|K_0\|_{L^1} \|u\|_{L^p}.$$

Proof. We use Young's inequality for convolution

$$\begin{aligned} \|\Delta_j u\|_{L^p} &= \|K_j * u\|_{L^p} \\ &\leqslant \|K_j\|_{L^1} \|u\|_{L^p} \\ &= \|K_0\|_{L^1} \|u\|_{L^p} \end{aligned}$$

□

Any tempered distribution can be approximated by a sequence of tempered distributions whose Fourier transforms are compactly supported, hence a sequence of compactly supported smooth functions.

Proposition 3.5. *Suppose $u \in \mathcal{S}'(\mathbb{R}^n)$, denote $\Delta_{<j} = \sum_{i \leqslant j-1} \Delta_i$, then*

$$u = \lim_{j \to +\infty} \Delta_{<j} u$$

in the space $\mathcal{S}'(\mathbb{R}^n)$.

Proof. For any Schwartz function $f \in \mathcal{S}(\mathbb{R}^n)$, we need to show

$$\langle u, f \rangle = \lim_{j \to +\infty} \langle \Delta_{<j} u, f \rangle = \lim_{j \to +\infty} \langle u, \Delta_{<j} f \rangle$$

which amounts to prove

$$\lim_{j \to +\infty} \Delta_{<j} f = \lim_{j \to +\infty} \mathcal{F}^{-1}\left(\sum_{i=-1}^{j-1} \rho_i \hat{f} \right) = f$$

convergence in $\mathcal{S}(\mathbb{R}^n)$. Since the Fourier transform $\mathcal{F}$ is a continuous isomorphism from $\mathcal{S}(\mathbb{R}^n)$ to itself with a continuous inverse $\mathcal{F}^{-1}$, we only need to show

$$\lim_{j \to +\infty} \sum_{i=-1}^{j-1} \rho_i \hat{f} = \hat{f}$$

convergence in $\mathcal{S}(\mathbb{R}^n)$, which is equivalent to show that for all $k \in \mathbb{N}$

$$\lim_{j \to +\infty} \left\| \sum_{i=j}^{\infty} \rho_i \hat{f} \right\|_k = 0$$

If $k = 0$, this is clear. For $k > 0$, since $\rho_j(x) := \rho_0\left(\frac{x}{2^j}\right)$, any mixed partial derivative of ρ_j for all $j > 0$ are uniformly bounded, hence

$$\sup_{|\alpha|, |\beta| \leqslant k, x \in \mathbb{R}^n} \left| x^\alpha \partial^\beta \left(\sum_{i=j}^{\infty} \rho_i \hat{f} \right)(x) \right| \leqslant C \sup_{|\alpha|, |\beta| \leqslant k, x \in \operatorname{supp}(\sum_{i=j}^{\infty} \rho_i)} |x^\alpha \partial^\beta \hat{f}(x)|$$

for some constant C. The right hand side clearly tends to 0 when $j \to +\infty$. Thus the result follows. □

3.2 Besov Spaces

Now we introduce a way to characterize the regularity of tempered distribution. Since the smoothness of a function is connected to the decay property of its Fourier transform, we need to control the growth of the Fourier transform of a tempered distribution in each Littlewood-Paley Block to measure its regularity. The main references are [23], [24] and [38], the book [4] contains more details.

For any $u \in \mathcal{S}'(\mathbb{R}^n)$, since the Fourier transform of each tempered distribution $\Delta_j u$ is compactly supported, then it can be identified with a smooth function of at most polynomial growth, so we can consider its L^p-norm.

Definition 3.6. *Suppose $\alpha \in \mathbb{R}$ and $1 \leqslant p, q \leqslant \infty$, the Besov space $B^{\alpha}_{p,q}(\mathbb{R}^n)$ is a subset of $\mathcal{S}'(\mathbb{R}^n)$ which contains all tempered distribution u such that*

$$\|u\|_{B^{\alpha}_{p,q}} := \left(\sum_{j \geqslant -1}^{\infty} (2^{j\alpha} \|\Delta_j u\|_{L^p})^q \right)^{\frac{1}{q}}$$

is finite.

Clearly $\|\cdot\|_{B^{\alpha}_{p,q}}$ is a norm for any $\alpha \in \mathbb{R}$ and $1 \leqslant p, q \leqslant \infty$, we will show the completeness later, hence the Besov space is a Banach space. Note that in this definition, $L^p -$norm is used to measure the amount of frequencies in each blocks and parameter α controls the decay speed of the amount of frequencies in each blocks in the sense that the $l^q -$norm is finite, hence controls the decay of high frequency terms, so α measures the regularity. The norm $\|\cdot\|_{B^{\alpha}_{p,q}}$ depends on the choice of dyadic partition of unity, but the space $B^{\alpha}_{p,q}(\mathbb{R}^n)$ doesn't, we will show this later.

In the application to stochastic partial differential equations, we will be more interested in the special case $B^{\alpha}_{\infty,\infty}(\mathbb{R}^n)$, so we denote it by $\mathcal{C}^{\alpha}(\mathbb{R}^n)$ or $\mathcal{C}^{\alpha}$ for short, and the norm $\|\cdot\|_{B^{\alpha}_{\infty,\infty}}$ by $\|\cdot\|_{\alpha}$ for simplicity.

Lemma 3.7.

1. *If $\alpha \leqslant \beta$, we have $\|u\|_{\alpha} \lesssim \|u\|_{\beta}$ for all $u \in \mathcal{C}^{\beta}$, hence $\mathcal{C}^{\beta} \subset \mathcal{C}^{\alpha}$;*
2. *If $\alpha > 0$, then $\|u\|_{L^{\infty}} \lesssim \|u\|_{\alpha}$ for all $u \in \mathcal{C}^{\alpha}$;*
3. *If $\alpha \leqslant 0$, then $\|u\|_{L^{\infty}} \gtrsim \|u\|_{\alpha}$ for all $u \in L^{\infty}$;*
4. *If $\alpha < 0$, then $\|\Delta_{\leqslant j} u\|_{L^{\infty}} \lesssim 2^{-j\alpha} \|u\|_{\alpha}$ for all $u \in \mathcal{C}^{\alpha}$;*
5. *If $\alpha > 0$, then $\|\Delta_{> j} u\|_{L^{\infty}} \lesssim 2^{-j\alpha} \|u\|_{\alpha}$ for all $u \in \mathcal{C}^{\alpha}$.*

Proof. 1. Since $u \in \mathcal{C}^{\beta}$, we know the norm

$$\|u\|_{\beta} = \sup_{j \geqslant -1} 2^{j\beta} \|\Delta_j u\|_{L^{\infty}}$$

is finite. Since $\beta \geqslant \alpha$, we have

$$\begin{aligned} \sup_{j \geqslant -1} 2^{j\beta} \|\Delta_j u\|_{L^{\infty}} &= 2^{-(\beta-\alpha)} \sup_{j \geqslant -1} 2^{(j+1)(\beta-\alpha)} 2^{j\alpha} \|\Delta_j u\|_{L^{\infty}} \\ &\geqslant 2^{-(\beta-\alpha)} \sup_{j \geqslant -1} 2^{j\alpha} \|\Delta_j u\|_{L^{\infty}} \\ &= 2^{-(\beta-\alpha)} \|u\|_{\alpha} \end{aligned}$$

2. Since $u \in \mathcal{C}^{\alpha}$, we know the norm

$$\|u\|_{\alpha} = \sup_{j \geqslant -1} 2^{j\alpha} \|\Delta_j u\|_{L^{\infty}}$$

is finite, thus each block $\Delta_j u$ is in L^{∞}. Consider a sequence of smooth functions, for integer $N \geqslant 1$

$$u_N = \sum_{i \geqslant -1}^{N} \Delta_i u$$

we claim it is a Cauchy sequence in L^∞. For integers $1 \leqslant N \leqslant M$ we have the estimate

$$\begin{aligned}
\|u_M - u_N\|_{L^\infty} &= \left\| \sum_{i \geqslant N}^{M} \Delta_i u \right\|_{L^\infty} \\
&= \left\| \sum_{i \geqslant N}^{M} 2^{-i\alpha} 2^{i\alpha} \Delta_i u \right\|_{L^\infty} \\
&\leqslant \sum_{i \geqslant N}^{M} 2^{-i\alpha} (2^{i\alpha} \|\Delta_i u\|_{L^\infty}) \\
&\leqslant \sum_{i \geqslant N}^{M} 2^{-i\alpha} \|u\|_\alpha \\
&\lesssim 2^{-N\alpha} \|u\|_\alpha \xrightarrow{N \to +\infty} 0
\end{aligned}$$

Since L^∞ is a Banach space, the sequence u_N converges to some function $\tilde{u} \in L^\infty$. We show that $u = \tilde{u}$ almost everywhere. Since we know that

$$\lim_{N \to +\infty} u_N = \lim_{N \to +\infty} \sum_{i \geqslant -1}^{N} \Delta_i u = u$$

in the space of tempered distribution $\mathcal{S}'$. Thus we have for any Schwartz function $f \in \mathcal{S}$

$$\lim_{N \to +\infty} \int_{\mathbb{R}^n} u_N \cdot f dx = \int_{\mathbb{R}^n} u \cdot f dx$$

however

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} (u_N - \tilde{u}) \cdot f dx \right| &= \left| \int_{\mathbb{R}^n} \left(\sum_{i \geqslant -1}^{N} \Delta_i u - \tilde{u} \right) \cdot f dx \right| \\
&\leqslant \int_{\mathbb{R}^n} \left| \left(\sum_{i \geqslant -1}^{N} \Delta_i u - \tilde{u} \right) \right| \cdot |f| \, dx \\
&\leqslant \int_{\mathbb{R}^n} |f| \, dx \left\| \sum_{i \geqslant -1}^{N} \Delta_i u - \tilde{u} \right\|_{L^\infty}
\end{aligned}$$

after taking $N \to +\infty$, we get

$$\left| \int_{\mathbb{R}^n} (u - \tilde{u}) \cdot f dx \right| \leqslant 0 \Rightarrow \int_{\mathbb{R}^n} (u - \tilde{u}) \cdot f dx = 0$$

which concludes $u = \tilde{u}$ almost everywhere(this also shows u is a continuous function since $\tilde{u}$ is L^∞ limit of a sequence of continuous functions). So we have

$$\begin{aligned}
\|u\|_{L^\infty} &= \|\tilde{u}\|_{L^\infty} \\
&= \left\| \sum_{i \geqslant -1}^{\infty} \Delta_i u \right\|_{L^\infty} \\
&= \left\| \sum_{i \geqslant -1}^{\infty} 2^{-i\alpha} 2^{i\alpha} \Delta_i u \right\|_{L^\infty}
\end{aligned}$$

$$\begin{aligned} &\leqslant \sum_{i\geqslant -1}^{\infty} 2^{-i\alpha}(2^{i\alpha}\|\Delta_i u\|_{L^\infty}) \\ &\leqslant \sum_{i\geqslant -1}^{\infty} 2^{-i\alpha}\|u\|_\alpha \simeq \|u\|_\alpha \end{aligned}$$

3. Since $u\in L^\infty$, and $\alpha\leqslant 0$, we have

$$\begin{aligned} \sup_{j\geqslant -1} 2^{j\alpha}\|\Delta_j u\|_{L^\infty} &= 2^{-\alpha}\sup_{j\geqslant -1} 2^{(j+1)\alpha}\|\Delta_j u\|_{L^\infty} \\ &\leqslant 2^{-\alpha}\sup_{j\geqslant -1}\|\Delta_j u\|_{L^\infty} \\ &\leqslant 2^{-\alpha}(\|K_0\|_{L^1}+\|K_{-1}\|_{L^1})\|u\|_{L^\infty} \end{aligned}$$

4. Since $\alpha<0$ and $u\in\mathcal{C}^\alpha$, we have

$$\begin{aligned} \|\Delta_{\leqslant j}u\|_{L^\infty} &= \left\|\sum_{i=-1}^{j}\Delta_i u\right\|_{L^\infty} \\ &= \left\|\sum_{i=-1}^{j} 2^{-i\alpha}2^{i\alpha}\Delta_i u\right\|_{L^\infty} \\ &\leqslant \sum_{i=-1}^{j} 2^{-i\alpha}(2^{i\alpha}\|\Delta_i u\|_{L^\infty}) \\ &\leqslant \|u\|_\alpha\sum_{i=-1}^{j} 2^{-i\alpha} \\ &= \frac{2^{-j\alpha}-2^{2\alpha}}{1-2^\alpha}\|u\|_\alpha \\ &\lesssim 2^{-j\alpha}\|u\|_\alpha \end{aligned}$$

5. Since $\alpha>0$ and $u\in\mathcal{C}^\alpha$, we have

$$\begin{aligned} \|\Delta_{>j}u\|_{L^\infty} &= \left\|\sum_{i=j+1}^{+\infty}\Delta_i u\right\|_{L^\infty} \\ &= \left\|\sum_{i=j+1}^{+\infty} 2^{-i\alpha}2^{i\alpha}\Delta_i u\right\|_{L^\infty} \\ &\leqslant \sum_{i=j+1}^{+\infty} 2^{-i\alpha}(2^{i\alpha}\|\Delta_i u\|_{L^\infty}) \\ &\leqslant \|u\|_\alpha\sum_{i=j+1}^{+\infty} 2^{-i\alpha} \\ &= \frac{2^{-j\alpha}}{2^\alpha-1}\|u\|_\alpha \\ &\lesssim 2^{-j\alpha}\|u\|_\alpha \end{aligned}$$

□

Now we show that the Besov norm is left continuous with respect to the regularity exponent.

Lemma 3.8. *Suppose $u \in \mathcal{C}^{\alpha}$, then*

$$\lim_{\alpha' \to \alpha^-} \|u\|_{\alpha'} = \|u\|_{\alpha}$$

Proof. We know from lemma 3.7 that $u \in \mathcal{C}^{\alpha'}$, and $\|u\|_{\alpha'}$ is an increasing function of α'. So it is clear that the limit exists and

$$\lim_{\alpha' \to \alpha^-} \|u\|_{\alpha'} \leqslant \|u\|_{\alpha}$$

We only need to show the inequality in the other direction. By definition

$$\|u\|_{\alpha} = \sup_{j \geqslant -1} 2^{j\alpha} \|\Delta_j u\|_{L^{\infty}}$$

Case 1. There is some $j' \geqslant -1$ such that $\|u\|_{\alpha} = 2^{j'\alpha} \|\Delta_{j'} u\|_{L^{\infty}}$, then

$$\begin{aligned} \|u\|_{\alpha} &= 2^{j'\alpha} \|\Delta_{j'} u\|_{L^{\infty}} \\ &\leqslant 2^{j'(\alpha - \alpha')} \|u\|_{\alpha'} \end{aligned}$$

so

$$\begin{aligned} \|u\|_{\alpha} &\leqslant \lim_{\alpha' \to \alpha^-} 2^{j'(\alpha - \alpha')} \|u\|_{\alpha'} \\ &= \lim_{\alpha' \to \alpha^-} \|u\|_{\alpha'} \end{aligned}$$

Case 2. There is an increasing sequence $j_n \to \infty$, such that $2^{j_n \alpha} \|\Delta_{j_n} u\|_{L^{\infty}}$ is increasing with limit $\|u\|_{\alpha}$. Then for any $\varepsilon > 0$, there is some integer N such that for $n \geqslant N$ we have $\|u\|_{\alpha} \leqslant 2^{j_n \alpha} \|\Delta_{j_n} u\|_{L^{\infty}} + \varepsilon$. Then we have

$$\begin{aligned} \|u\|_{\alpha} &\leqslant 2^{j_n \alpha} \|\Delta_{j_n} u\|_{L^{\infty}} + \varepsilon \\ &\leqslant 2^{j_n (\alpha - \alpha')} \|u\|_{\alpha'} + \varepsilon \end{aligned}$$

so

$$\begin{aligned} \|u\|_{\alpha} &\leqslant \lim_{\alpha' \to \alpha^-} 2^{j_n (\alpha - \alpha')} \|u\|_{\alpha'} + \varepsilon \\ &= \lim_{\alpha' \to \alpha^-} \|u\|_{\alpha'} + \varepsilon \end{aligned}$$

Since ε is arbitrary, we get the desired inequality. □

The following lemma is useful when we need to approximate a distribution in a Besov space with slightly smaller regularity exponent.

Lemma 3.9. *Suppose $\alpha > \beta$ and $u \in \mathcal{C}^\alpha$, the the sequence $\Delta_{\leqslant n} u$ converges in $\mathcal{C}^\beta$. Hence $\mathcal{C}^\alpha$ is contained in the closure of Schwartz functions in $\mathcal{C}^\beta$.*

Proof. Since we have $\Delta_j(\Delta_{\leqslant n} u - u) = 0$ for $j < n$, then

$$\begin{aligned}
& \|\Delta_{\leqslant n} u - u\|_\beta \\
= & \sup_{j \geqslant -1} 2^{j\beta} \|\Delta_j(\Delta_{\leqslant n} u - u)\|_{L^\infty} \\
= & \sup_{j \geqslant n} 2^{j(\beta-\alpha)} \cdot 2^{j\alpha} \left\| \Delta_j \sum_{\max\{n, j-1\} \leqslant k \leqslant j+1} \Delta_k u \right\|_{L^\infty} \\
\lesssim & \sup_{j \geqslant n} 2^{j(\beta-\alpha)} \cdot 2^{j\alpha} \|\Delta_j u\|_{L^\infty} \\
\leqslant & 2^{n(\beta-\alpha)} \|u\|_\alpha \xrightarrow{n \to +\infty} 0
\end{aligned}$$

The second statement is true since each $\Delta_{\leqslant n} u$ is contained in $\mathcal{S}$, hence in $\mathcal{C}^\beta$. □

A set B is called a ball if it has the form $\{x \in \mathbb{R}^n : |x| \leqslant R\}$ with $R > 0$, a set A is called an annulus if it has the form $\{x \in \mathbb{R}^n : 0 < r_1 \leqslant |x| \leqslant r_2\}$ with $0 < r_1 < r_2$.

Lemma 3.10. (Bernstein Type Inequalities) *Suppose B is a ball and A is an annulus. For any constants $k \in \mathbb{N}, 1 \leqslant p \leqslant q \leqslant \infty$ and $\lambda > 0$, then*

1. there exists a constant C which depends on k, B, p, q, such that for any function $f \in L^p$ with $\operatorname{supp}(\mathcal{F}(f)) \subset \lambda B$, we have:

$$\max_{|\alpha| = k} \|\partial^\alpha f\|_{L^q} \leqslant C \lambda^{k + n\left(\frac{1}{p} - \frac{1}{q}\right)} \|f\|_{L^p}$$

2. there exists a constant C which depends on k, A, p, such that for any function $f \in L^p$ with $\operatorname{supp}(\mathcal{F}(f)) \subset \lambda A$, we have:

$$\lambda^k \|f\|_{L^p} \leqslant C \cdot \max_{|\alpha| = k} \|\partial^\alpha f\|_{L^p}$$

Proof. 1. Denote r to be the constant satisfies $\frac{1}{p} + \frac{1}{r} = 1 + \frac{1}{q}$. Let $\varphi \in C_c^\infty$, with $\varphi(x) = 1$ when $x \in B$, and denote $\varphi_\lambda(x) = \varphi\left(\frac{x}{\lambda}\right)$. By Young's inequality for convolutions (see B.3), we have

$$\begin{aligned}
\|\partial^\alpha f\|_{L^q} &= \|\partial^\alpha \mathcal{F}^{-1}(\varphi_\lambda \cdot \hat{f})\|_{L^q} \\
&= \|\partial^\alpha(\mathcal{F}^{-1}(\varphi_\lambda)) * f\|_{L^q} \\
&\leqslant \|\partial^\alpha(\mathcal{F}^{-1}(\varphi_\lambda))\|_{L^r} \cdot \|f\|_{L^p}
\end{aligned}$$

Case 1. If $r<\infty$, we estimate the term $\|\partial^\alpha \mathcal{F}^{-1}(\varphi_\lambda)\|_{L^r}$

$$\begin{aligned}\|\partial^\alpha(\mathcal{F}^{-1}(\varphi_\lambda))\|_{L^r} &= \left(\int_{\mathbb{R}^n} |\partial^\alpha(\mathcal{F}^{-1}(\varphi_\lambda))(x)|^r dx\right)^{\frac{1}{r}} \\ &= \left(\int_{\mathbb{R}^n} |\partial^\alpha(\lambda^n(\mathcal{F}^{-1}\varphi)(\lambda x))|^r dx\right)^{\frac{1}{r}} \\ &= \left(\lambda^{nr+|\alpha|r}\int_{\mathbb{R}^n} |(\partial^\alpha(\mathcal{F}^{-1}\varphi))(\lambda x)|^r dx\right)^{\frac{1}{r}} \\ &= \left(\lambda^{nr+|\alpha|r-n}\int_{\mathbb{R}^n} |(\partial^\alpha(\mathcal{F}^{-1}\varphi))(x)|^r dx\right)^{\frac{1}{r}} \\ &= \lambda^{|\alpha|+n\left(1-\frac{1}{r}\right)}\|\partial^\alpha(\mathcal{F}^{-1}\varphi)\|_{L^r}\end{aligned}$$

so we have

$$\begin{aligned}\|\partial^\alpha f\|_{L^q} &\leqslant \lambda^{|\alpha|+n\left(1-\frac{1}{r}\right)}\|\partial^\alpha(\mathcal{F}^{-1}\varphi)\|_{L^r}\cdot\|f\|_{L^p} \\ &\leqslant C\lambda^{k+n\left(\frac{1}{p}-\frac{1}{q}\right)}\|f\|_{L^p}\end{aligned}$$

with $C=\|\partial^\alpha(\mathcal{F}^{-1}\varphi)\|_{L^r}$.
Case 2. If $r=\infty$, then $p=1$, $q=\infty$, we estimate the term $\|\partial^\alpha \mathcal{F}^{-1}(\varphi_\lambda)\|_{L^\infty}$

$$\begin{aligned}\|\partial^\alpha(\mathcal{F}^{-1}(\varphi_\lambda))\|_{L^\infty} &= \sup_{x\in\mathbb{R}^n} |\partial^\alpha(\mathcal{F}^{-1}(\varphi_\lambda))(x)| \\ &= \sup_{x\in\mathbb{R}^n} |\partial^\alpha(\lambda^n(\mathcal{F}^{-1}\varphi)(\lambda x))| \\ &= \lambda^{n+|\alpha|}\sup_{x\in\mathbb{R}^n} |(\partial^\alpha(\mathcal{F}^{-1}\varphi))(\lambda x)| \\ &= \lambda^{n+|\alpha|}\sup_{x\in\mathbb{R}^n} |(\partial^\alpha(\mathcal{F}^{-1}\varphi))(x)| \\ &= \lambda^{n+|\alpha|}\|(\partial^\alpha(\mathcal{F}^{-1}\varphi))(x)\|_{L^\infty}\end{aligned}$$

so we have

$$\begin{aligned}\|\partial^\alpha f\|_{L^\infty} &\leqslant \lambda^{n+|\alpha|}\|(\partial^\alpha(\mathcal{F}^{-1}\varphi))(x)\|_{L^\infty}\cdot\|f\|_{L^p} \\ &\leqslant C\lambda^{k+n\left(\frac{1}{p}-\frac{1}{q}\right)}\|f\|_{L^p}\end{aligned}$$

Then the result follows.

2. Let $\varphi\in C_c^\infty$, $\operatorname{supp}(\varphi)$ is a neighborhood of A which does not intersect a neighborhood of 0 , with $\varphi(x)=1$ when $x\in A$, and denote $\varphi_\lambda(x)=\varphi\left(\frac{x}{\lambda}\right)$. Since we can find a set of integers $(N_\alpha)_\alpha\in\mathbb{N}$ indexed by the multi-index α, such that

$$|x|^{2k}=(x_1^2+\cdots+x_n^2)^k=\sum_{|\alpha|=k} N_\alpha(ix)^\alpha(-ix)^\alpha$$

and we have $\operatorname{supp}(\mathcal{F}(f)) \subset \lambda A$, then we have the decomposition

$$\begin{aligned} f &= \mathcal{F}^{-1}(\mathcal{F}f) \\ &= \mathcal{F}^{-1}(\varphi_\lambda \mathcal{F}f) \\ &= \mathcal{F}^{-1}\left(\varphi_\lambda \cdot \frac{\sum_{|\alpha|=k} N_\alpha (i\xi)^\alpha (-i\xi)^\alpha}{|\xi|^{2k}} \cdot \mathcal{F}f\right) \\ &= \sum_{|\alpha|=k} N_\alpha \mathcal{F}^{-1}\left(\varphi_\lambda \frac{(-i\xi)^\alpha}{|\xi|^{2k}} \cdot (i\xi)^\alpha \mathcal{F}f\right) \\ &= \sum_{|\alpha|=k} N_\alpha \mathcal{F}^{-1}\left(\varphi_\lambda \frac{(-i\xi)^\alpha}{|\xi|^{2k}} \cdot \mathcal{F}(\partial^\alpha f)\right) \\ &= \sum_{|\alpha|=k} N_\alpha \mathcal{F}^{-1}\left(\varphi_\lambda \frac{(-i\xi)^\alpha}{|\xi|^{2k}}\right) * \partial^\alpha f \end{aligned}$$

By Young's inequality for convolutions, we have

$$\begin{aligned} \lambda^k \|f\|_{L^p} &= \lambda^k \left\| \sum_{|\alpha|=k} N_\alpha \mathcal{F}^{-1}\left(\varphi_\lambda \frac{(-i\xi)^\alpha}{|\xi|^{2k}}\right) * \partial^\alpha f \right\|_{L^p} \\ &\leqslant \lambda^k \sum_{|\alpha|=k} N_\alpha \left\| \mathcal{F}^{-1}\left(\varphi_\lambda \frac{(-i\xi)^\alpha}{|\xi|^{2k}}\right) * \partial^\alpha f \right\|_{L^p} \\ &\leqslant \lambda^k \sum_{|\alpha|=k} N_\alpha \left\| \mathcal{F}^{-1}\left(\varphi_\lambda \frac{(-i\xi)^\alpha}{|\xi|^{2k}}\right) \right\|_{L^1} \cdot \|\partial^\alpha f\|_{L^p} \\ &= \lambda^k \sum_{|\alpha|=k} N_\alpha \int_{\mathbb{R}^n} \left| \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \varphi\left(\frac{\xi}{\lambda}\right) \frac{(-i\xi)^\alpha}{|\xi|^{2k}} e^{i\langle \xi, x\rangle}\, d\xi \right| dx \cdot \|\partial^\alpha f\|_{L^p} \\ &= \sum_{|\alpha|=k} N_\alpha \int_{\mathbb{R}^n} \left| \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \varphi\left(\frac{\xi}{\lambda}\right) \frac{(-i\xi/\lambda)^\alpha}{|\xi/\lambda|^{2k}} e^{i\langle \xi/\lambda, \lambda x\rangle}\, d\xi/\lambda \right| d\lambda x \cdot \|\partial^\alpha f\|_{L^p} \\ &= \sum_{|\alpha|=k} N_\alpha \int_{\mathbb{R}^n} \left| \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \varphi(\xi) \frac{(-i\xi)^\alpha}{|\xi|^{2k}} e^{i\langle \xi, x\rangle}\, d\xi \right| dx \cdot \|\partial^\alpha f\|_{L^p} \\ &= \sum_{|\alpha|=k} N_\alpha \left\| \mathcal{F}^{-1}\left(\varphi \cdot \frac{(-i\xi)^\alpha}{|\xi|^{2k}}\right) \right\|_{L^1} \cdot \|\partial^\alpha f\|_{L^p} \\ &\leqslant \left(\sum_{|\alpha|=k} N_\alpha \left\| \mathcal{F}^{-1}\left(\varphi \cdot \frac{(-i\xi)^\alpha}{|\xi|^{2k}}\right) \right\|_{L^1} \right) \cdot \max_{|\alpha|=k} \|\partial^\alpha f\|_{L^p} \end{aligned}$$

which concludes the proof. □

Bernstein type inequalities are useful when we need to estimate functions with compact Fourier transformations.

Corollary 3.11. *Given $u \in \mathcal{C}^\alpha$ where $\alpha \in \mathbb{R}$ and a multi-index $\mu \in \mathbb{N}^n$, we have*

$$\|\partial^\mu u\|_{\alpha - |\mu|} \lesssim \|u\|_\alpha.$$

Proof. By Bernstein type inequality for balls with $\lambda = 2^j$, $j \in \mathbb{N}$, we have

$$\begin{aligned}
\|\Delta_j(\partial^\mu u)\|_{L^\infty} &= \|K_j * (\partial^\mu u)\|_{L^\infty} \\
&= \|\partial^\mu (K_j * u)\|_{L^\infty} \\
&= \|\partial^\mu (\Delta_j u)\|_{L^\infty} \\
&\lesssim 2^{j|\mu|}\|\Delta_j u\|_{L^\infty} \\
&= 2^{j(|\mu|-\alpha)} 2^{j\alpha}\|\Delta_j u\|_{L^\infty} \\
&\leqslant 2^{j(|\mu|-\alpha)}\|u\|_\alpha
\end{aligned}$$

For $j = -1$, use Bernstein type inequality for balls with $\lambda = 1$, we have

$$\begin{aligned}
\|\Delta_{-1}(\partial^\mu u)\|_{L^\infty} &= \|\partial^\mu(\Delta_{-1} u)\|_{L^\infty} \\
&\lesssim \|\Delta_j u\|_{L^\infty} \\
&\leqslant 2^\alpha \|u\|_\alpha
\end{aligned}$$

In summary we get $\|\partial^\mu u\|_{\alpha - |\mu|} \lesssim \|u\|_\alpha$. □

Now we use Bernstein type inequalities to show the equivalence of Besov spaces $B^\alpha_{\infty,\infty}$ and Hölder spaces $C^{0,\alpha}$ when $\alpha \in (0,1)$.

Corollary 3.12. *For $\alpha \in (0,1)$, we have $\mathcal{C}^\alpha = C^{0,\alpha}$.*

Proof.

1. $\mathcal{C}^\alpha \subset C^{0,\alpha}$: Suppose $u \in \mathcal{C}^\alpha$, since $\alpha > 0$, we have $\|u\|_{L^\infty} \lesssim \|u\|_\alpha$. As shown in the proof 3.7 part 2, we have that

$$u = \lim_{n \to \infty} \sum_{j=-1}^{n} \Delta_j u$$

in L^∞ and hence u is continuous and the limit converges at each point in $\mathbb{R}^n$. Suppose $x, y \in \mathbb{R}^n$, since each block $\Delta_j u$ is a smooth function, then there is some point z on the line segment $\overline{xy}$ such that

$$\Delta_j u(x) - \Delta_j u(y) = \nabla(\Delta_j u)(z) \cdot (x - y)$$

hence by Bernstein inequality we have

$$\begin{aligned}
|\Delta_j u(x) - \Delta_j u(y)| &= |\nabla(\Delta_j u)(z) \cdot (x - y)| \\
&\leqslant \max_{i \in \{1,2,\ldots,n\}} \|\partial^i \Delta_j u\|_{L^\infty} \cdot \|x - y\| \\
&\lesssim 2^j \|\Delta_j u\|_{L^\infty} \cdot \|x - y\| \\
&\leqslant 2^{j(1-\alpha)} \|u\|_\alpha \cdot \|x - y\|
\end{aligned}$$

We also have

$$\begin{aligned}
|\Delta_j u(x) - \Delta_j u(y)| &\lesssim \|\Delta_j u\|_{L^\infty} \\
&\leqslant 2^{-j\alpha}\|u\|_\alpha
\end{aligned}$$

If $\|x - y\| \geqslant 1$, we have

$$\begin{aligned}
\frac{|u(x) - u(y)|}{\|x - y\|^\alpha} &\leqslant |u(x) - u(y)| \\
&\lesssim \|u\|_{L^\infty} \lesssim \|u\|_\alpha
\end{aligned}$$

If $\|x-y\|<1$, then there is some $j_0\in\mathbb{N}$ such that $2^{-j_0}\simeq\|x-y\|$ and we have

$$\begin{aligned}
|u(x)-u(y)| &\leqslant \sum_{j=-1}^{\infty}|\Delta_j u(x)-\Delta_j u(y)|\\
&= \sum_{j=-1}^{j_0}|\Delta_j u(x)-\Delta_j u(y)|+\sum_{j=j_0+1}^{\infty}|\Delta_j u(x)-\Delta_j u(y)|\\
&\leqslant \sum_{j=-1}^{j_0}2^{j(1-\alpha)}\|u\|_\alpha\cdot\|x-y\|+\sum_{j=j_0+1}^{\infty}2^{-j\alpha}\|u\|_\alpha\\
&\simeq \|u\|_\alpha(2^{j_0(1-\alpha)}\|x-y\|+2^{-j_0\alpha})\\
&\simeq \|u\|_\alpha\cdot\|x-y\|^\alpha
\end{aligned}$$

In summary we have

$$\|u\|_{C^{0,\alpha}}=\|u\|_{L^\infty}+\sup_{x\neq y}\frac{|u(x)-u(y)|}{\|x-y\|^\alpha}\lesssim\|u\|_\alpha$$

hence $\mathcal{C}^\alpha\subset C^{0,\alpha}$.

2. $\mathcal{C}^\alpha\supset C^{0,\alpha}$: Suppose $f\in C^{0,\alpha}$, then $f\in L^\infty$ and we have

$$\|\Delta_{-1}f\|_{L^\infty}\lesssim\|f\|_{L^\infty}$$

For $j\geqslant 0$, since $\rho_j(0)=0$, then

$$\begin{aligned}
\int_{\mathbb{R}^n}\mathcal{F}^{-1}(\rho_j)(x)dx &= \int_{\mathbb{R}^n}\mathcal{F}^{-1}(\rho_j)(x)e^{-i\langle x,0\rangle}dx\\
&= \mathcal{F}(\mathcal{F}^{-1}(\rho_j))(0)\\
&= \rho_j(0)\\
&= 0
\end{aligned}$$

Since $f\in C^{0,\alpha}$, we have $|f(y)-f(x)|\leqslant\|f\|_{C^{0,\alpha}}\cdot\|x-y\|^\alpha$, then

$$\begin{aligned}
|\Delta_j f(x)| &= |\mathcal{F}^{-1}(\rho_j)*f(x)|\\
&= \left|\int_{\mathbb{R}^n}\mathcal{F}^{-1}(\rho_j)(x-y)f(y)dy\right|\\
&= \left|\int_{\mathbb{R}^n}\mathcal{F}^{-1}(\rho_j)(x-y)(f(y)-f(x))dy\right|\\
&\leqslant \|f\|_{C^{0,\alpha}}\int_{\mathbb{R}^n}|\mathcal{F}^{-1}(\rho_j)(x-y)|\cdot\|y-x\|^\alpha dy\\
&= \|f\|_{C^{0,\alpha}}\cdot 2^{jn}\int_{\mathbb{R}^n}|\mathcal{F}^{-1}(\rho_0)(2^j(x-y))|\cdot\|y-x\|^\alpha dy\\
&= \|f\|_{C^{0,\alpha}}\cdot 2^{-j\alpha}\int_{\mathbb{R}^n}|\mathcal{F}^{-1}(\rho_0)(2^j(x-y))|\cdot\|2^j(y-x)\|^\alpha d\,(2^j y)\\
&= \|f\|_{C^{0,\alpha}}\cdot 2^{-j\alpha}\int_{\mathbb{R}^n}|\mathcal{F}^{-1}(\rho_0)(x-y)|\cdot\|y-x\|^\alpha dy
\end{aligned}$$

which concludes the inequality $\|f\|_\alpha \lesssim \|f\|_{C^{0,\alpha}}$. □

In general, it can be shown that if $\alpha \in (0, +\infty) \backslash \mathbb{N}$, the Besov space $\mathcal{C}^\alpha(\mathbb{R}^n)$ and the Hölder space $C^{\lfloor \alpha \rfloor, \alpha - \lfloor \alpha \rfloor}(\mathbb{R}^n)$ are the same, the Besov norm and Hölder norm are equivalent (see [4], and [1] for a proof), hence the space $\mathcal{C}^\alpha(\mathbb{R}^n)$ and $C^{\lfloor \alpha \rfloor, \alpha - \lfloor \alpha \rfloor}(\mathbb{R}^n)$ are the same. However if $\alpha \in \mathbb{N}$, the Besov space $\mathcal{C}^\alpha(\mathbb{R}^n)$ is strictly larger than $C^\alpha(\mathbb{R}^n)$.

Theorem 3.13. (Besov embedding) *Let* $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$, $1 \leqslant q_1 \leqslant q_2 \leqslant \infty$, *and* $\alpha \in \mathbb{R}$. *Then for any* $u \in B^\alpha_{p_1, q_1}(\mathbb{R}^n)$, *we have*

$$\|u\|_{B_{p_2,q_2}^{\alpha - n\left(\frac{1}{p_1} - \frac{1}{p_2}\right)}} \lesssim \|u\|_{B^\alpha_{p_1,q_1}}$$

hence we have a continuous embedding of $B^\alpha_{p_1,q_1}(\mathbb{R}^n)$ *into* $B_{p_2,q_2}^{\alpha - n\left(\frac{1}{p_1} - \frac{1}{p_2}\right)}(\mathbb{R}^n)$.

Proof. We have

$$\begin{aligned}
\|u\|_{B^\alpha_{p_1,q_1}} &= \left(\sum_{j \geqslant -1}^{\infty} (2^{j\alpha} \|\Delta_j u\|_{L^{p_1}})^{q_1}\right)^{\frac{1}{q_1}} \\
&\gtrsim \left(\sum_{j \geqslant -1}^{\infty} \left(2^{j\alpha} 2^{-jn\left(\frac{1}{p_1} - \frac{1}{p_2}\right)} \|\Delta_j u\|_{L^{p_2}}\right)^{q_1}\right)^{\frac{1}{q_1}} \\
&= \left(\sum_{j \geqslant -1}^{\infty} \left(2^{j\left(\alpha - n\left(\frac{1}{p_1} - \frac{1}{p_2}\right)\right)} \|\Delta_j u\|_{L^{p_2}}\right)^{q_1}\right)^{\frac{1}{q_1}} \\
&\geqslant \left(\sum_{j \geqslant -1}^{\infty} \left(2^{j\left(\alpha - n\left(\frac{1}{p_1} - \frac{1}{p_2}\right)\right)} \|\Delta_j u\|_{L^{p_2}}\right)^{q_2}\right)^{\frac{1}{q_2}} \\
&= \|u\|_{B_{p_2,q_2}^{\alpha - n\left(\frac{1}{p_1} - \frac{1}{p_2}\right)}}
\end{aligned}$$

where we used $\|\cdot\|_{l^{q_2}} \leqslant \|\cdot\|_{l^{q_1}}$.

□

Next, we want to show the Besov space $B^\alpha_{p,q}(\mathbb{R}^n)$ is complete, and is independent of the dyadic partition of unity used in the definition.

Lemma 3.14. *Let* $A \subset \mathbb{R}^n$ *be an annulus and* $\{u_j\}_{j \geqslant -1}$ *be a sequence of smooth functions such that* $\mathcal{F}(u_j)$ *is supported in* $2^j A$, *and* $\|u_j\|_{L^\infty} \lesssim 2^{-j\alpha}$ *for all* j. *Then the limit*

$$\lim_{N \to \infty} \sum_{j=-1}^{N} u_j$$

converges in the space $\mathcal{S}'$.

Proof. For any $f\in\mathcal{S}$, we need to show

$$\left\langle \sum_{j=-1}^{N} u_j, f\right\rangle = \sum_{j=-1}^{N} \langle u_j, f\rangle$$

converges as $N\to\infty$. So we try to estimate a typical term $\langle u_j, f\rangle$. As we did in the proof of Bernstein type inequalities, let $\varphi\in C_c^\infty$, $\mathrm{supp}(\varphi)$ is a neighborhood of A which does not intersect a neighborhood of 0 , with $\varphi(x)=1$ when $x\in A$, and denote $\varphi_\lambda(x)=\varphi\left(\frac{x}{\lambda}\right)$. We have

$$\begin{aligned}
f(x) &= \sum_{|\beta|=k} N_\beta \mathcal{F}^{-1}\left(\varphi_\lambda \frac{(-i\xi)^\beta}{|\xi|^{2k}}\right) * \partial^\beta f(x)\\
&= \sum_{|\beta|=k} \frac{N_\beta}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi\left(\frac{\xi}{\lambda}\right)\frac{(-i\xi)^\beta}{|\xi|^{2k}} e^{i\langle\xi,x-y\rangle}\,\partial^\beta f(y)d\xi dy\\
&= \sum_{|\beta|=k} \frac{\lambda^{n-k}N_\beta}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi\left(\frac{\xi}{\lambda}\right)\frac{\left(-i\frac{\xi}{\lambda}\right)^\beta}{\left|\frac{\xi}{\lambda}\right|^{2k}} e^{i\left\langle\frac{\xi}{\lambda},\lambda(x-y)\right\rangle}\,\partial^\beta f(y)d\frac{\xi}{\lambda}dy\\
&= \sum_{|\beta|=k} \frac{\lambda^{n-k}N_\beta}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(\xi)\frac{(-i\xi)^\beta}{|\xi|^{2k}} e^{i\langle\xi,\lambda(x-y)\rangle}\,\partial^\beta f(y)d\xi dy
\end{aligned}$$

Since the function

$$\int_{\mathbb{R}^n}\varphi(\xi)\frac{(-i\xi)^\beta}{|\xi|^{2k}}e^{i\langle\xi,\lambda y\rangle}d\xi$$

has compact supported Fourier transform, then it is a Schwartz function. Then by using Hölder inequality and Young's inequality for convolution, we have

$$\begin{aligned}
&|\langle u_j, f\rangle|\\
=&\left|\sum_{|\beta|=k}\frac{\lambda^{n-k}N_\beta}{(2\pi)^n}\iiint_{\mathbb{R}^{3n}}\varphi(\xi)\frac{(-i\xi)^\beta}{|\xi|^{2k}}e^{i\langle\xi,\lambda(x-y)\rangle}\,\partial^\beta f(y)u_j(x)d\xi dy dx\right|\\
\leqslant&\sum_{|\beta|=k}\frac{\lambda^{n-k}N_\beta}{(2\pi)^n}\int_{\mathbb{R}^n}\left|\iint_{\mathbb{R}^{2n}}\varphi(\xi)\frac{(-i\xi)^\beta}{|\xi|^{2k}}e^{i\langle\xi,\lambda(x-y)\rangle}\,\partial^\beta f(y)d\xi dy\right||u_j(x)|dx\\
\leqslant&\sum_{|\beta|=k}\frac{\lambda^{n-k}N_\beta}{(2\pi)^n}\|u_j\|_{L^\infty}\left\|\iint_{\mathbb{R}^{2n}}\varphi(\xi)\frac{(-i\xi)^\beta}{|\xi|^{2k}}e^{i\langle\xi,\lambda(x-y)\rangle}\,\partial^\beta f(y)d\xi dy\right\|_{L^1}\\
\leqslant&\sum_{|\beta|=k}\frac{\lambda^{n-k}N_\beta}{(2\pi)^n}\|u_j\|_{L^\infty}\|\partial^\beta f\|_{L^1}\left\|\int_{\mathbb{R}^n}\varphi(\xi)\frac{(-i\xi)^\beta}{|\xi|^{2k}}e^{i\langle\xi,\lambda\cdot\rangle}d\xi\right\|_{L^1}\\
=&\sum_{|\beta|=k}\frac{\lambda^{-k}N_\beta}{(2\pi)^n}\|u_j\|_{L^\infty}\|\partial^\beta f\|_{L^1}\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}^n}\varphi(\xi)\frac{(-i\xi)^\beta}{|\xi|^{2k}}e^{i\langle\xi,\lambda x\rangle}d\xi\right|d\lambda x
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{|\beta|=k} \frac{\lambda^{-k} N_\beta}{(2\pi)^n} \|u_j\|_{L^\infty} \|\partial^\beta f\|_{L^1} \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \varphi(\xi) \frac{(-i\xi)^\beta}{|\xi|^{2k}} e^{i\langle \xi, x\rangle} d\xi \right| dx \\
&= \sum_{|\beta|=k} \frac{\lambda^{-k} N_\beta}{(2\pi)^n} \|u_j\|_{L^\infty} \|\partial^\beta f\|_{L^1} \left\| \int_{\mathbb{R}^n} \varphi(\xi) \frac{(-i\xi)^\beta}{|\xi|^{2k}} e^{i\langle \xi, \cdot\rangle} d\xi \right\|_{L^1} \\
&\lesssim 2^{-j\alpha} \sum_{|\beta|=k} \frac{\lambda^{-k} N_\beta}{(2\pi)^n} \|\partial^\beta f\|_{L^1} \left\| \int_{\mathbb{R}^n} \varphi(\xi) \frac{(-i\xi)^\beta}{|\xi|^{2k}} e^{i\langle \xi, \cdot\rangle} d\xi \right\|_{L^1}
\end{aligned}
$$

We then choose $\lambda = 2^j$ and k to be some integer such that $k + \alpha > 0$. Then

$$
|\langle u_j, f\rangle| \lesssim 2^{-j(\alpha+k)} \sum_{|\beta|=k} \frac{N_\beta}{(2\pi)^n} \|\partial^\beta f\|_{L^1} \left\| \int_{\mathbb{R}^n} \varphi(\xi) \frac{(-i\xi)^\beta}{|\xi|^{2k}} e^{i\langle \xi, \cdot\rangle} d\xi \right\|_{L^1}
$$

hence the series

$$
\sum_{j=-1}^{\infty} \langle u_j, f\rangle
$$

converges absolutely. □

Lemma 3.15.

1. Let $A \subset \mathbb{R}^n$ be an annulus, $\alpha \in \mathbb{R}$, and $\{u_j\}_{j \geqslant -1}$ be a sequence of smooth functions such that $\mathcal{F}(u_j)$ is supported in $2^j A$, and $\|u_j\|_{L^\infty} \lesssim 2^{-j\alpha}$ for all j. Then the limit

$$
\lim_{N\to\infty} \sum_{j=-1}^{N} u_j
$$

converges in the space $\mathcal{S}'$, denote the limit by u, then $u \in \mathcal{C}^\alpha$, and

$$
\|u\|_\alpha \lesssim_\alpha \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\}.
$$

2. Let $B \subset \mathbb{R}^n$ be a ball, $\alpha > 0$, and $\{u_j\}_{j \geqslant -1}$ be a sequence of smooth functions such that $\mathcal{F}(u_j)$ is supported in $2^j B$, and $\|u_j\|_{L^\infty} \lesssim 2^{-j\alpha}$ for all j. Then the limit

$$
\lim_{N\to\infty} \sum_{j=-1}^{N} u_j
$$

converges in the space $\mathcal{S}'$, denote the limit by u, then $u \in \mathcal{C}^\alpha$, and

$$
\|u\|_\alpha \lesssim_\alpha \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\}.
$$

Proof.

1. We have shown in previous lemma that the infinite sum converges in the space $\mathcal{S}'$. Denote the limit by u, we show $u \in \mathcal{C}^\alpha$. Since $\mathcal{F}(u_j)$ is supported in in $2^j A$, then $\Delta_i u_j \neq 0$ only for $i \sim j$. So

$$\begin{aligned} \|\Delta_i u\|_{L^\infty} &= \left\| \Delta_i \sum_{j=-1}^{\infty} u_j \right\|_{L^\infty} \\ &\leqslant \sum_{j \sim i} \|\Delta_i u_j\|_{L^\infty} \\ &\lesssim \sum_{j \sim i} 2^{j\alpha} \|u_j\|_{L^\infty} 2^{-j\alpha} \\ &\leqslant \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\} \sum_{j \sim i} 2^{-j\alpha} \\ &\simeq \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\} 2^{-i\alpha} \end{aligned}$$

thus

$$\|u\|_\alpha = \sup_{i \geqslant -1} 2^{i\alpha} \|\Delta_i u\|_{L^\infty} \lesssim \sup_{j \geqslant -1} \{2^{j\alpha} \|u_j\|_{L^\infty}\}$$

which concludes the proof.

2. For any $f \in \mathcal{S}$, we need to show

$$\left\langle \sum_{j=-1}^{N} u_j, f \right\rangle = \sum_{j=-1}^{N} \langle u_j, f \rangle$$

converges as $N \to \infty$. For a typical term $\langle u_j, f \rangle$, since $\alpha > 0$ and

$$\begin{aligned} |\langle u_j, f \rangle| &= \left| \int_{\mathbb{R}^n} u_j(x) f(x) dx \right| \\ &\leqslant \int_{\mathbb{R}^n} |u_j(x) f(x)| dx \\ &\leqslant \|u_j\|_{L^\infty} \int_{\mathbb{R}^n} |f(x)| dx \\ &= \|u_j\|_{L^\infty} \|f\|_{L^1} \\ &\lesssim 2^{-j\alpha} \|f\|_{L^1} \end{aligned}$$

we then have the limit converges in the space of tempered distributions. Denote the limit by u, we show $u \in \mathcal{C}^\alpha$. Since $\mathcal{F}(u_j)$ is supported in in $2^j B$, then $\Delta_i u_j \neq 0$ only for $i \lesssim j$. So

$$\begin{aligned} \|\Delta_i u\|_{L^\infty} &= \left\| \Delta_i \sum_{j=-1}^{\infty} u_j \right\|_{L^\infty} \\ &\leqslant \sum_{j \gtrsim i} \|\Delta_i u_j\|_{L^\infty} \\ &\lesssim \sum_{j \gtrsim i} 2^{j\alpha} \|u_j\|_{L^\infty} 2^{-j\alpha} \end{aligned}$$

$$\begin{aligned} &\leqslant \sup_{j\geqslant -1}\{2^{j\alpha}\|u_j\|_{L^\infty}\}\sum_{j\gtrsim i} 2^{-j\alpha} \\ &\simeq \sup_{j\geqslant -1}\{2^{j\alpha}\|u_j\|_{L^\infty}\}2^{-i\alpha} \end{aligned}$$

where we used $\alpha > 0$. We then conclude that

$$\|u\|_\alpha = \sup_{i\geqslant -1} 2^{i\alpha}\|\Delta_i u\|_{L^\infty} \lesssim \sup_{j\geqslant -1}\{2^{j\alpha}\|u_j\|_{L^\infty}\}. \qquad \square$$

One can prove a general version of this lemma for Besov spaces $B^\alpha_{p,q}$, which can be found in the book [4]. We can get the following corollary easily from this lemma.

Corollary 3.16. *If $(\tilde{\rho}_j)_{j\geqslant -1}$ is another dyadic partition of unity, and denote $\tilde{\Delta}_j$ to be the corresponding operators in Littlewood-Paley theory, $\tilde{B}^\alpha_{p,q}$ to be the corresponding Besov spaces. Then $\tilde{B}^\alpha_{p,q} = B^\alpha_{p,q}$ as a set, two norms $\|\cdot\|_{\tilde{B}^\alpha_{p,q}}$ and $\|\cdot\|_{B^\alpha_{p,q}}$ are equivalent.*

Corollary 3.17. *The Besov space $B^\alpha_{p,q}(\mathbb{R}^n)$ is complete, for all $\alpha \in \mathbb{R}$ and $1 \leqslant p$, $q \leqslant \infty$.*

Proof. Suppose $\{u_k\}_{k\geqslant 0}$ is a Cauchy sequence in $B^\alpha_{p,q}(\mathbb{R}^n)$, hence for any $\varepsilon > 0$, there is some positive integer M, such that for all $n, m \geqslant M$, we have

$$\|u_m - u_n\|_{B^\alpha_{p,q}} = \left(\sum_{j\geqslant -1}^{\infty} (2^{j\alpha}\|\Delta_j u_m - \Delta_j u_n\|_{L^p})^q\right)^{\frac{1}{q}} < \varepsilon$$

thus for all j

$$\|\Delta_j u_m - \Delta_j u_n\|_{L^p} < 2^{-j\alpha}\varepsilon$$

we then get for each j, the sequence of smooth functions $\{\Delta_j u_n\}_{n\geqslant 0}$ is a Cauchy sequence in L^p space, thus has a limit, denote it by v_j. Let $n \to \infty$ in previous inequalities we get

$$\|\Delta_j u_m - v_j\|_{L^p} < 2^{-j\alpha}\varepsilon$$

for all $m \geqslant M$, and

$$\left(\sum_{j=-1}^{K} (2^{j\alpha}\|\Delta_j u_m - v_j\|_{L^p})^q\right)^{\frac{1}{q}} < \varepsilon$$

for all $K \in \mathbb{N}$, then let $K \to \infty$ to get

$$\left(\sum_{j=-1}^{\infty} (2^{j\alpha}\|\Delta_j u_m - v_j\|_{L^p})^q\right)^{\frac{1}{q}} < \varepsilon$$

Then we have an inequality

$$\begin{aligned}
&\left(\sum_{j\geqslant -1}^{\infty}(2^{j\alpha}\|v_j\|_{L^p})^q\right)^{\frac{1}{q}}\\
\leqslant\; &\left(\sum_{j\geqslant -1}^{\infty}(2^{j\alpha}\|\Delta_j u_m - v_j\|_{L^p})^q\right)^{\frac{1}{q}} + \left(\sum_{j\geqslant -1}^{\infty}(2^{j\alpha}\|\Delta_j u_m\|_{L^p})^q\right)^{\frac{1}{q}}\\
\leqslant\; &\varepsilon + C
\end{aligned}$$

for some constant C since the Cauchy is bounded. Since $\varepsilon > 0$ is arbitrary, we get

$$\left(\sum_{j\geqslant -1}^{\infty}(2^{j\alpha}\|v_j\|_{L^p})^q\right)^{\frac{1}{q}} < C$$

and hence $\|v_j\|_{L^p} < C2^{-j\alpha}$. By Bernstein inequality, since $\mathcal{F}(\Delta_j u_m)$ is supported in some ball $2^j\tilde{B}$ where $\tilde{B}$ is related to the domain used in the Littlewood-Paley decomposition, then

$$\|\Delta_j u_m\|_{L^\infty} \leqslant C' \cdot 2^{\frac{jn}{p}}\|\Delta_j u_m\|_{L^p}$$

also

$$\|\Delta_j u_m - \Delta_j u_k\|_{L^\infty} \leqslant C' \cdot 2^{\frac{jn}{p}}\|\Delta_j u_m - \Delta_j u_k\|_{L^p}$$

thus the sequence $\{\Delta_j u_n\}_{n\geqslant 0}$ is a Cauchy sequence in L^∞ space, hence has a limit denoted by v_j' which is clearly continuous. Then $v_j = v_j'$ almost everywhere, this follows from for any compactly supported test function g, by using Hölder inequality

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n}(v_j - v_j')(x)g(x)dx\right|\\
\leqslant\; &\int_{\mathbb{R}^n}|(v_j - \Delta_j u_k)(x)g(x)|dx + \int_{\mathbb{R}^n}|(v_j' - \Delta_j u_k)(x)g(x)|dx\\
\leqslant\; &\|v_j - \Delta_j u_k\|_{L^p}\cdot\|g\|_{L^{\frac{p}{p-1}}} + \|v_j' - \Delta_j u_k\|_{L^\infty}\cdot\|g\|_{L^1} \xrightarrow{k\to\infty} 0
\end{aligned}$$

Thus taking limit as $m \to \infty$ in our inequality we get

$$\|v_j\|_{L^\infty} \leqslant C' \cdot 2^{\frac{jn}{p}}\|v_j\|_{L^p} < CC' \cdot 2^{-j\left(\alpha - \frac{n}{p}\right)}$$

For $j \geqslant 0$, we know that the Fourier transform of v_j is supported in the domain 2^jA, where A is the annulus used in the Littlewood-Paley decomposition, since we know $\mathcal{F}(\Delta_j u_m)$ is supported in the domain 2^jA, then for any compactly supported test function g supported outside 2^jA, we have

$$0 = \int_{\mathbb{R}^n}\mathcal{F}(\Delta_j u_m)(x)\cdot g(x)dx = \int_{\mathbb{R}^n}\Delta_j u_m(x)\cdot\mathcal{F}^{-1}(g)(x)dx$$

hence

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n} v_j(x)\cdot\mathcal{F}^{-1}(g)\,(x)dx\right|\\
=&\left|\int_{\mathbb{R}^n}(v_j(x)-\Delta_j u_m(x))\cdot\mathcal{F}^{-1}(g)(x)dx\right|\\
\leqslant&\int_{\mathbb{R}^n}|v_j(x)-\Delta_j u_m(x)|\cdot|\mathcal{F}^{-1}(g)(x)|dx\\
\leqslant&\|v_j-\Delta_j u_m\|_{L^\infty}\cdot\|\mathcal{F}^{-1}(g)\,\|_{L^1}\xrightarrow{m\to\infty}0
\end{aligned}$$

so

$$0=\int_{\mathbb{R}^n} v_j(x)\cdot\mathcal{F}^{-1}(g)\,(x)dx=\langle v_j,\mathcal{F}^{-1}(g)\rangle=\langle\mathcal{F}(v_j),g\rangle$$

this shows if we think of v_j as a tempered distribution, then it is supported in 2^jA, which also shows v_j is smooth. Now we can use the lemma 3.14 to get the infinite sum

$$\sum_{j\geqslant -1} v_j$$

converges to some limit $u\in\mathcal{S}'$. We need to show that $u\in B^\alpha_{p,q}$ and $u_k\to u$ in $B^\alpha_{p,q}$.

We first show that $u\in B^\alpha_{p,q}$. Since $\operatorname{supp}(v_j)\subset 2^jA$, we know that $\Delta_i v_j\neq 0$ if and only if $|i-j|<2$. Moreover, we have the estimate

$$\|\Delta_i v_j\|_{L^p}\leqslant\|K_0\|_{L^1}\cdot\|v_j\|_{L^p}$$

for all j, thus

$$\begin{aligned}
\|u\|_{B^\alpha_{p,q}} &= \left(\sum_{i=-1}^{\infty}(2^{i\alpha}\|\Delta_i u\|_{L^p})^q\right)^{\frac{1}{q}}\\
&= \left(\sum_{i=-1}^{\infty}\left(2^{i\alpha}\left\|\sum_{|j-i|<2}\Delta_i v_j\right\|_{L^p}\right)^q\right)^{\frac{1}{q}}\\
&\leqslant \left(\sum_{i=-1}^{\infty}\left(2^{i\alpha}\sum_{|j-i|<2}\|K_0\|_{L^1}\cdot\|v_j\|_{L^p}\right)^q\right)^{\frac{1}{q}}\\
&\leqslant \|K_0\|_{L^1}\sum_{k=-1,0,1}\left(\sum_{i=-1}^{\infty}(2^{i\alpha}\|v_{j+k}\|_{L^p})^q\right)^{\frac{1}{q}}\\
&< +\infty
\end{aligned}$$

where we assume $v_{-2}=v_{-3}=0$ in the sum. This shows $u\in B^\alpha_{p,q}$.

Next we show that $u_k\to u$ in $B^\alpha_{p,q}$. For any $\varepsilon>0$, take integer $M>0$ such that when $k>M$, we have $\|\Delta_j u_k-v_j\|_{L^p}<2^{-j\alpha}\varepsilon$ and

$$\left(\sum_{j\geqslant -1}^{\infty}(2^{j\alpha}\|\Delta_j u_m-v_j\|_{L^p})^q\right)^{\frac{1}{q}}<\varepsilon$$

for all j. Then

$$
\begin{aligned}
& \|u_k - u\|_{B^\alpha_{p,q}} \\
= & \left(\sum_{i=-1}^{\infty} (2^{i\alpha} \|\Delta_i (u_k - u)\|_{L^p})^q \right)^{\frac{1}{q}} \\
= & \left(\sum_{i=-1}^{\infty} \left(2^{i\alpha} \left\| \Delta_i u_k - \sum_{|j-i|<2} \Delta_i v_j \right\|_{L^p} \right)^q \right)^{\frac{1}{q}} \\
= & \left(\sum_{i=-1}^{\infty} \left(2^{i\alpha} \left\| (\Delta_i u_k - v_i) - \left(\sum_{|j-i|<2} \Delta_i v_j - v_i \right) \right\|_{L^p} \right)^q \right)^{\frac{1}{q}} \\
= & \left(\sum_{i=-1}^{\infty} \left(2^{i\alpha} \left\| (\Delta_i u_k - v_i) - \sum_{|j-i|<2} ((\Delta_i v_j - \Delta_i \Delta_j u_k) - (\Delta_j v_i - \Delta_j \Delta_i u_k)) \right\|_{L^p} \right)^q \right)^{\frac{1}{q}} \\
\leqslant & \left(\sum_{i=-1}^{\infty} (2^{i\alpha} \|(\Delta_i u_k - v_i)\|_{L^p})^q \right)^{\frac{1}{q}} + \sum_{l=-1,0,1} \left(\sum_{i=-1}^{\infty} (2^{i\alpha} \|\Delta_i (v_{i+l} - \Delta_{i+l} u_k)\|_{L^p})^q \right)^{\frac{1}{q}} \\
& + \sum_{l=-1,0,1} \left(\sum_{i=-1}^{\infty} (2^{i\alpha} \|\Delta_{i+l} (v_i - \Delta_i u_k)\|_{L^p})^q \right)^{\frac{1}{q}} \\
\lesssim & \ \varepsilon + \left(\sum_{i=-1}^{\infty} (2^{i\alpha} \|v_i - \Delta_i u_k\|_{L^p})^q \right)^{\frac{1}{q}} + \left(\sum_{i=-1}^{\infty} (2^{i\alpha} \|v_i - \Delta_i u_k\|_{L^p})^q \right)^{\frac{1}{q}} \\
\lesssim & \ \varepsilon
\end{aligned}
$$

which gives the result. □

3.3 First Order Paracontrolled Calculus

Now we know how to measure the regularity of distributions. We next consider the problem of defining the multiplication of tempered distributions. Let us begin with some heuristic discussion, we want to define the multiplication uv of $u, v \in \mathcal{S}'$. When u, v are smooth functions, we want it to be agree with the usual product of functions. So we do the following formal computation

$$uv = \sum_{i \geqslant -1} \Delta_i u \sum_{j \geqslant -1} \Delta_j v$$

we can see the problem by following lemma.

Lemma 3.18. *There exists an annulus $\tilde{A}$ such that for any $j \geqslant 1$ and all $i < j - 1$, we have*

$$\operatorname{supp}(\mathcal{F}(\Delta_i u \Delta_j v)) \subset 2^j \tilde{A}$$

for all $u, v \in \mathcal{S}'$*. And There exists a ball* $\tilde{B}$ *such that for any* $i, j \geqslant -1$ *and all* $|i - j| \leqslant 1$*, we have*

$$\operatorname{supp}(\mathcal{F}(\Delta_i u \Delta_j v)) \subset 2^j \tilde{B}$$

for all $u, v \in \mathcal{S}'$*.*

Proof. Consider $j \geqslant 1$ and $i < j - 1$, if $i \geqslant 0$, we have

$$\begin{aligned}
\operatorname{supp}(\mathcal{F}(\Delta_i u \Delta_j v)) &= \operatorname{supp}(\mathcal{F}(\Delta_i u) * \mathcal{F}(\Delta_j v)) \\
&\subset \overline{\operatorname{supp}(\mathcal{F}(\Delta_i u)) + \operatorname{supp}(\mathcal{F}(\Delta_j v))} \\
&\subset 2^i A + 2^j A \\
&= 2^j (2^{i-j} A + A)
\end{aligned}$$

by the construction of dyadic partition of unity we know that $2^{i-j} A \cap A = \varnothing$, then there is an annulus $\tilde{A}$ that $2^{i-j} A + A \subset \tilde{A}$. For the case of $i = 0$, since

$$\begin{aligned}
\operatorname{supp}(\mathcal{F}(\Delta_i u \Delta_j v)) &\subset \overline{\operatorname{supp}(\mathcal{F}(\Delta_i u)) + \operatorname{supp}(\mathcal{F}(\Delta_j v))} \\
&\subset 2^i B + 2^j A \\
&= 2^j (2^{i-j} B + A)
\end{aligned}$$

by the construction of dyadic partition of unity we know that $2^{i-j} B \cap A = \varnothing$, then we may enlarge annulus $\tilde{A}$ so that $2^{i-j} B + A \subset \tilde{A}$.

Consider $i \geqslant j \geqslant -1$ and $|i - j| \leqslant 1$, then

$$\begin{aligned}
\operatorname{supp}(\mathcal{F}(\Delta_i u \Delta_j v)) &\subset \overline{\operatorname{supp}(\mathcal{F}(\Delta_i u)) + \operatorname{supp}(\mathcal{F}(\Delta_j v))} \\
&\subset 2^i B + 2^j B \ \text{ or } \ 2^i A + 2^j B \ \text{ or } \ 2^i A + 2^j A \\
&= 2^j (2^{i-j} B + B) \ \text{ or } \ 2^j (2^{i-j} A + B) \ \text{ or } \ 2^j (2^{i-j} A + A)
\end{aligned}$$

in each case, $2^{i-j} B \cap B \neq \varnothing$ or $2^{i-j} A \cap B \neq \varnothing$ or $2^{i-j} A \cap A \neq \varnothing$. Which means we can only find a ball $\tilde{B}$ such that

$$\operatorname{supp}(\mathcal{F}(\Delta_i u \Delta_j v)) \subset 2^j \tilde{B}$$

which concludes the proof. □

From this lemma we can see that if we decompose the sum as

$$uv = \sum_{i \geqslant -1} \Delta_i u \sum_{j \geqslant -1} \Delta_j v = \sum_{i \geqslant 1} \sum_{i - j > 1} \Delta_i u \Delta_j v + \sum_{j \geqslant 1} \sum_{j - i > 1} \Delta_i u \Delta_j v + \sum_{j \geqslant -1} \sum_{i : |i - j| \leqslant 1} \Delta_i u \Delta_j v$$

the third term has a problem since we may add too much amplitudes to the frequencies around 0, so we have to control the decay of product of blocks in order to have a well-defined product distribution. For the first two terms there are no such problems, since

$$\sum_{i \geqslant 1} \sum_{i - j > 1} \Delta_i u \Delta_j v = \sum_{i \geqslant 1} \Delta_i u \Delta_{< i - 1} v \qquad \sum_{j \geqslant 1} \sum_{j - i > 1} \Delta_i u \Delta_j v = \sum_{j \geqslant 1} \Delta_{< j - 1} u \Delta_j v$$

there is no infinite sum involved in any annulus $2^j\tilde{A}$ of Fourier spaces. We use the following notations to denote the three terms appeared in the formal decomposition

$$u\prec v=v\succ u=\sum_{j\geqslant 1}\Delta_{<j-1}u\Delta_j v\quad u\circ v=\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\Delta_i u\Delta_j v$$

and we call $u\prec v$ and $v\prec u$ paraproducts, and $u\circ v$ the resonant product.

Theorem 3.19. (Paraproduct estimate)

1. $\|u\prec v\|_\beta\lesssim_\beta\|u\|_{L^\infty}\|v\|_\beta$ *for all* $\beta\in\mathbb{R}$, $u\in L^\infty$ *and* $v\in\mathcal{C}^\beta$*;*

2. $\|u\prec v\|_{\alpha+\beta}\lesssim_{\alpha,\beta}\|u\|_\alpha\|v\|_\beta$ *for all* $\beta\in\mathbb{R}$, $\alpha<0$, $u\in\mathcal{C}^\alpha$ *and* $v\in\mathcal{C}^\beta$*;*

3. $\|u\circ v\|_{\alpha+\beta}\lesssim_{\alpha,\beta}\|u\|_\alpha\|v\|_\beta$ *for all* $\alpha,\beta\in\mathbb{R}$ *such that* $\alpha+\beta>0$, $u\in\mathcal{C}^\alpha$ *and* $v\in\mathcal{C}^\beta$*.*

Proof. 1. We know that there is an annulus $\tilde{A}$ such that $\mathrm{supp}(\mathcal{F}(\Delta_{<j-1}u\Delta_j v))\subset 2^j\tilde{A}$, we need to estimate $\|\Delta_{<j-1}u\Delta_j v\|_{L^\infty}$. Since for all $x\in\mathbb{R}^n$ we have

$$\begin{aligned}1 &= \rho_{-1}(x)+\sum_{i=0}^{+\infty}\rho_i(x)\\ &= \rho_{-1}(x)+\sum_{i=0}^{+\infty}\rho_0\Big(\frac{x}{2^i}\Big)\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}1 &= \rho_{-1}\Big(\frac{x}{2^{j-1}}\Big)+\sum_{i=0}^{+\infty}\rho_0\Big(\frac{x}{2^{i+j-1}}\Big)\\ &= \rho_{-1}\Big(\frac{x}{2^{j-1}}\Big)+\sum_{i=j-1}^{+\infty}\rho_0\Big(\frac{x}{2^i}\Big)\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}\rho_{-1}\Big(\frac{x}{2^{j-1}}\Big) &= \rho_{-1}(x)+\sum_{i=0}^{j-2}\rho_0\Big(\frac{x}{2^i}\Big)\\ &= \rho_{-1}(x)+\sum_{i=0}^{j-2}\rho_i(x)\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}\Delta_{<j-1}u &= \mathcal{F}^{-1}((\rho_{-1}+\rho_0+\cdots+\rho_{j-2})\mathcal{F}u)\\ &= \mathcal{F}^{-1}\Big(\rho_{-1}\Big(\frac{1}{2^{j-1}}\cdot\Big)\mathcal{F}u\Big)\\ &= 2^{(j-1)n}\mathcal{F}^{-1}(\rho_{-1})(2^{j-1}\cdot)*u\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}\|\Delta_{<j-1}u\|_{L^\infty} &= \|2^{(j-1)n}\mathcal{F}^{-1}(\rho_{-1})(2^{j-1}\cdot)*u\|_{L^\infty}\\ &= \Big\|\int_{\mathbb{R}^n}2^{(j-1)n}\mathcal{F}^{-1}(\rho_{-1})(2^{j-1}y)u(x-y)dy\Big\|_{L^\infty}\\ &\leqslant \|u\|_{L^\infty}\|\mathcal{F}^{-1}(\rho_{-1})\|_{L^1}\end{aligned}$$

$\Rightarrow$

$$\begin{aligned}\|\Delta_{<j-1}u\Delta_j v\|_{L^\infty} &\leqslant \|\Delta_{<j-1}u\|_{L^\infty}\|\Delta_j v\|_{L^\infty}\\ &\leqslant \|u\|_{L^\infty}\|\mathcal{F}^{-1}(\rho_{-1})\|_{L^1}2^{-j\beta}\|v\|_\beta\end{aligned}$$

Then by lemma 3.15 we have

$$u \prec v = \sum_{j\geqslant 1} \Delta_{<j-1}u\Delta_j v \in \mathcal{C}^\beta$$

and moreover

$$\|u\prec v\|_\beta \lesssim_\beta \sup_{j\geqslant -1} \{2^{j\beta}\|\Delta_{<j-1}u\Delta_j v\|_{L^\infty}\} \lesssim_\beta \|u\|_{L^\infty}\|v\|_\beta$$

which concludes the proof.

2. Again we know that there is an annulus $\tilde{A}$ such that $\operatorname{supp}(\mathcal{F}(\Delta_{<j-1}u\Delta_j v)) \subset 2^j\tilde{A}$, we need to estimate $\|\Delta_{<j-1}u\Delta_j v\|_{L^\infty}$. Since $\alpha < 0$, we have

$$\begin{aligned}\|\Delta_{<j-1}u\Delta_j v\|_{L^\infty} &\leqslant \sum_{i<j-1} \|\Delta_i u\|_{L^\infty}\|\Delta_j v\|_{L^\infty}\\ &\leqslant \|u\|_\alpha\|v\|_\beta \sum_{i<j-1} 2^{-i\alpha-j\beta}\\ &\lesssim_\alpha 2^{-j(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta\end{aligned}$$

$\Rightarrow u\prec v \in \mathcal{C}^{\alpha+\beta}$ and $\|u\prec v\|_{\alpha+\beta} \lesssim_{\alpha+\beta,\alpha} \|u\|_\alpha\|v\|_\beta$ which is exactly the inequality

$$\|u\prec v\|_{\alpha+\beta} \lesssim_{\alpha,\beta} \|u\|_\alpha\|v\|_\beta$$

.

3. We know that there is a ball $\tilde{B}$ such that

$$\operatorname{supp}\left(\mathcal{F}\left(\sum_{i:|i-j|\leqslant 1} \Delta_i u\Delta_j v\right)\right) \subset 2^j\tilde{B},$$

and

$$\begin{aligned}\left\|\sum_{i:|i-j|\leqslant 1} \Delta_i u\Delta_j v\right\|_{L^\infty} &\leqslant \sum_{i:|i-j|\leqslant 1} \|\Delta_i u\|_{L^\infty}\|\Delta_j v\|_{L^\infty}\\ &\leqslant \sum_{i:|i-j|\leqslant 1} 2^{-i\alpha-j\beta}\|u\|_\alpha\|v\|_\beta\\ &\simeq_\alpha 2^{-j(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta\end{aligned}$$

By lemma 3.15 and $\alpha+\beta>0$ we have $u\circ v\in\mathcal{C}^{\alpha+\beta}$ and $\|u\circ v\|_{\alpha+\beta} \lesssim_{\alpha,\beta} \|u\|_\alpha\|v\|_\beta$.

□

From this theorem, we can then define the product of two distributions when certain regularity conditions are satisfied.

Corollary 3.20. *Suppose* $\alpha+\beta>0$*, the the map from* $\mathcal{C}^{\alpha}\times\mathcal{C}^{\beta}\to\mathcal{C}^{\alpha\wedge\beta}$ *defined by* $(u,v)\mapsto u\prec v+u\succ v+u\circ v$ *for any* $u\in\mathcal{C}^{\alpha}, v\in\mathcal{C}^{\beta}$ *is a bounded bilinear map. The terms* $u\prec v, u\succ v, u\circ v$ *depend on the specific dyadic partition of unity, the sum does not.*

Proof.

Case 1. $\beta\geqslant\alpha>0$. Then

$$\|u\circ v\|_{\alpha}\leqslant\|u\circ v\|_{\alpha+\beta}\lesssim_{\alpha,\beta}\|u\|_{\alpha}\|v\|_{\beta}$$

$$\|u\prec v\|_{\alpha}\leqslant\|u\prec v\|_{\beta}\lesssim_{\beta}\|u\|_{L^{\infty}}\|v\|_{\beta}\lesssim\|u\|_{\alpha}\|v\|_{\beta}$$

$$\|u\succ v\|_{\alpha}\lesssim_{\alpha}\|u\|_{\alpha}\|v\|_{L^{\infty}}\lesssim\|u\|_{\alpha}\|v\|_{\beta}$$

which shows the sum $u\prec v+u\succ v+u\circ v$ is well-defined in $\mathcal{C}^{\alpha}$.

Case 2. $\beta>0>\alpha$. Then

$$\|u\circ v\|_{\alpha}\leqslant\|u\circ v\|_{\alpha+\beta}\lesssim_{\alpha,\beta}\|u\|_{\alpha}\|v\|_{\beta}$$

$$\|u\prec v\|_{\alpha}\leqslant\|u\prec v\|_{\alpha+\beta}\lesssim_{\alpha,\beta}\|u\|_{\alpha}\|v\|_{\beta}$$

$$\|u\succ v\|_{\alpha}\lesssim_{\alpha}\|u\|_{\alpha}\|v\|_{L^{\infty}}\lesssim\|u\|_{\alpha}\|v\|_{\beta}$$

which shows the sum $u\prec v+u\succ v+u\circ v$ is well-defined in $\mathcal{C}^{\alpha}$.

Case 3. $\beta>\alpha=0$. Then

$$\|u\circ v\|_{\alpha}\leqslant\|u\circ v\|_{\alpha+\beta}\lesssim_{\alpha,\beta}\|u\|_{\alpha}\|v\|_{\beta}$$

$$\|u\prec v\|_{\alpha}\leqslant\|u\prec v\|_{\beta-\frac{\beta}{2}}\lesssim_{\beta}\|u\|_{-\frac{\beta}{2}}\|v\|_{\beta}\leqslant\|u\|_{\alpha}\|v\|_{\beta}$$

$$\|u\succ v\|_{\alpha}\lesssim_{\alpha}\|u\|_{\alpha}\|v\|_{L^{\infty}}\lesssim\|u\|_{\alpha}\|v\|_{\beta}$$

which shows the sum $u\prec v+u\succ v+u\circ v$ is well-defined in $\mathcal{C}^{\alpha}$.

Now we show that the sum is independent of the specific dyadic partition of unity, we denote the corresponding paraproduct and resonant product by $u\tilde{\prec}v$, $u\tilde{\succ}v$ and $u\tilde{\circ}v$. Clearly when u and v are smooth, we have

$$u\prec v+u\succ v+u\circ v=uv=u\tilde{\prec}v+u\tilde{\succ}v+u\tilde{\circ}v$$

and for general $u\in\mathcal{C}^{\alpha}$ and $v\in\mathcal{C}^{\beta}$, we take slight smaller $\alpha'<\alpha$, $\beta'<\beta$ such that $\alpha'+\beta'>0$, and we know that $u\in\mathcal{C}^{\alpha'}$ and $v\in\mathcal{C}^{\beta'}$. Then the maps

$$(u,v)\mapsto u\prec v+u\succ v+u\circ v$$

$$(u,v)\mapsto u\tilde{\prec}v+u\tilde{\succ}v+u\tilde{\circ}v$$

are both continuous bilinear maps from $\mathcal{C}^{\alpha'}\times\mathcal{C}^{\beta'}$ to $\mathcal{C}^{\alpha'\wedge\beta'}$, which agree with each other on smooth elements. So the continuous extension of these two maps have same value on the closure of $(C^{\infty}\cap\mathcal{C}^{\alpha'})\times(C^{\infty}\cap\mathcal{C}^{\beta'})$, which contains $\mathcal{C}^{\alpha}\times\mathcal{C}^{\beta}$ as a subspace, hence two maps agree on space $\mathcal{C}^{\alpha}\times\mathcal{C}^{\beta}$ and the uniqueness is proved. □

In mathematical analysis, we approximate a differentiable function by a linear function, with a remainder of higher order infinitesimal. Now we show some similar result, which constitutes the so called first order paracontrolled calculus. The main reference for the theorems is [23].

Lemma 3.21. *Suppose* $0<\alpha<1$ *and* $\beta\in\mathbb{R}$*, then for any* $u\in\mathcal{C}^\alpha$, $g\in\mathcal{C}^\beta$*and* $j\geqslant -1$*, we have*

$$\|\Delta_j(u\prec v)-u\Delta_j v\|_{L^\infty}\lesssim 2^{-j(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta$$

Proof. According to lemma 3.18, we know that there is an annulus $\tilde{A}$, such that the Fourier transform $\mathcal{F}(\Delta_{<i-1}u\Delta_i v)$ is supported on $2^i\tilde{A}$. Then

$$\begin{aligned}
&\Delta_j(u\prec v)-u\Delta_j v\\
=\;&\Delta_j\left(\sum_{i\geqslant 1}\Delta_{<i-1}u\Delta_i v\right)-u\Delta_j\left(\sum_{i\geqslant -1}\Delta_i v\right)\\
=\;&\sum_{i:i\sim j}[\Delta_j(\Delta_{<i-1}u\Delta_i v)-u\Delta_j\Delta_i v]\\
=\;&\sum_{i:i\sim j}[\Delta_j(\Delta_{<i-1}u\Delta_i v)-\Delta_{<i-1}u\Delta_j\Delta_i v]-\sum_{i:i\sim j}\Delta_{\geqslant i-1}u\Delta_j\Delta_i v
\end{aligned}$$

For the second term, since $\alpha>0$, we have

$$\begin{aligned}
\left\|\sum_{i:i\sim j}\Delta_{\geqslant i-1}u\Delta_j\Delta_i v\right\|_{L^\infty} &\lesssim \sum_{i:i\sim j}\|\Delta_{\geqslant i-1}u\|_{L^\infty}\|\Delta_j\Delta_i v\|_{L^\infty}\\
&\lesssim \sum_{i:i\sim j}2^{-i\alpha}\|u\|_\alpha 2^{-i\beta}\|v\|_\beta\\
&\simeq 2^{-j(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta
\end{aligned}$$

For the first term, we have

$$\begin{aligned}
&|(\Delta_j(\Delta_{<i-1}u\Delta_i v)-\Delta_{<i-1}u\Delta_j\Delta_i v)(x)|\\
=\;&\left|\int_{\mathbb{R}^n}K_j(x-y)(\Delta_{<i-1}u(y)-\Delta_{<i-1}u(x))\Delta_i v(y)dy\right|\\
\leqslant\;&\int_{\mathbb{R}^n}|K_j(x-y)|\cdot|\Delta_{<i-1}u(y)-\Delta_{<i-1}u(x)|\cdot|\Delta_i v(y)|dy\\
\lesssim\;&\int_{\mathbb{R}^n}|K_j(x-y)|\cdot\max_{|\mu|=1}\|\partial^\mu\Delta_{<i-1}u\|_{L^\infty}\cdot|x-y|\cdot\|\Delta_i v\|_{L^\infty}dy\\
\leqslant\;&2^{-i\beta}\|v\|_\beta\cdot\max_{|\mu|=1}\|\mathcal{F}^{-1}[(i\xi)^\mu\rho_{<i-1}\mathcal{F}(u)]\|_{L^\infty}\int_{\mathbb{R}^n}|K_j(x-y)\cdot(x-y)|dy\\
=\;&2^{-i\beta}\|v\|_\beta\cdot\max_{|\mu|=1}\|\mathcal{F}^{-1}[\rho_{<i-1}\mathcal{F}(\partial^\mu u)]\|_{L^\infty}\int_{\mathbb{R}^n}|K_j(x-y)\cdot(x-y)|dy\\
=\;&2^{-i\beta}\|v\|_\beta\cdot\max_{|\mu|=1}\|\Delta_{<i-1}\partial^\mu u\|_{L^\infty}\int_{\mathbb{R}^n}|K_j(x-y)\cdot(x-y)|dy\\
\lesssim\;&2^{-i\beta}\|v\|_\beta\cdot\max_{|\mu|=1}2^{-i(\alpha-1)}\|\partial^\mu u\|_{\alpha-1}\int_{\mathbb{R}^n}|K_j(y)\cdot y|dy\\
\lesssim\;&2^{i(1-\alpha-\beta)}\|u\|_\alpha\|v\|_\beta\int_{\mathbb{R}^n}|K_j(y)\cdot y|dy
\end{aligned}$$

If $j=-1$, we have

$$\begin{aligned} & \left\| \sum_{i:i\sim -1} [\Delta_{-1}(\Delta_{<i-1}u\Delta_i v) - \Delta_{<i-1}u\Delta_{-1}\Delta_i v] \right\|_{L^\infty} \\ \lesssim & \sum_{i:i\sim -1} 2^{i(1-\alpha-\beta)}\|u\|_\alpha \|v\|_\beta \int_{\mathbb{R}^n} |K_{-1}(y)\cdot y|dy \\ \simeq & \ 2^{-(-1)(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta \end{aligned}$$

If $j\geqslant 0$, we have

$$\begin{aligned} & \int_{\mathbb{R}^n} |K_j(y)\cdot y|dy \\ = & \int_{\mathbb{R}^n} |2^{jn}K_0(2^j y)\cdot y|dy \\ = & \ 2^{-j}\int_{\mathbb{R}^n} |K_0(y)\cdot y|dy \end{aligned}$$

thus

$$\begin{aligned} & \left\| \sum_{i:i\sim j} [\Delta_j(\Delta_{<i-1}u\Delta_i v) - \Delta_{<i-1}u\Delta_j\Delta_i v] \right\|_{L^\infty} \\ \lesssim & \sum_{i:i\sim j} 2^{i(1-\alpha-\beta)}\|u\|_\alpha\|v\|_\beta 2^{-j}\int_{\mathbb{R}^n} |K_0(y)\cdot y|dy \\ = & \sum_{i:i\sim j} 2^{(i-j)}2^{-i(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta\int_{\mathbb{R}^n} |K_0(y)\cdot y|dy \\ \simeq & \ 2^{-j(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta \end{aligned}$$

which concludes the proof. □

Lemma 3.22. *Suppose $\alpha\in(0,1)$ and $\beta,\gamma\in\mathbb{R}$, such that $\alpha+\beta+\gamma>0$ and $\beta+\gamma<0$. Then there exists a bounded trilinear operator $C:\mathcal{C}^\alpha\times\mathcal{C}^\beta\times\mathcal{C}^\gamma\to\mathcal{C}^{\alpha+\beta+\gamma}$ such that*

$$C(u,v,w) = ((u\prec v)\circ w) - u(v\circ w)$$

whenever $u,v,w\in\mathcal{S}$.

Proof. Suppose $u,v,w\in\mathcal{S}$, then

$$\begin{aligned} & C(u,v,w) \\ = & ((u\prec v)\circ w) - u(v\circ w) \\ = & \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1} (\Delta_i(u\prec v) - u\Delta_i v)\Delta_j w \\ = & \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\sum_{k\geqslant -1} (\Delta_i(\Delta_k u\prec v) - \Delta_k u\Delta_i v)\Delta_j w \end{aligned}$$

By definition, we have

$$\begin{aligned}\Delta_k u \prec v &= \sum_{l\geqslant 1} \Delta_{<l-1}\Delta_k u \Delta_l v\\ &= \sum_{l\geqslant \max\{k+1,1\}} \Delta_{<l-1}\Delta_k u \Delta_l v\end{aligned}$$

then according to lemma 3.18 there exists a ball $\tilde{B}$ such that the support of Fourier transform $\mathcal{F}(\Delta_k u \prec v)$ is contained in the $\mathbb{R}^n \backslash 2^k \tilde{B}$, then

$$\Delta_i(\Delta_k u \prec v) \neq 0 \Leftrightarrow k \lesssim i$$

Hence the commutator can be written as

$$\begin{aligned}&C(u,v,w)\\ =& \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1} (\Delta_i(\Delta_{\lesssim i} u \prec v) - \Delta_{\lesssim i} u \Delta_i v)\Delta_j w - \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1} \Delta_{\gtrsim i} u \Delta_i v \Delta_j w\end{aligned}$$

First, let us look at the second term

$$\begin{aligned}&\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1} \Delta_{\gtrsim i} u \Delta_i v \Delta_j w\\ =& \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\sum_{k\gtrsim i} \Delta_k u \Delta_i v \Delta_j w\\ =& \sum_{j\geqslant -1}\sum_{k\geqslant -1}\sum_{i:|i-j|\leqslant 1, i\lesssim k} \Delta_k u \Delta_i v \Delta_j w\\ =& \sum_{k\geqslant -1}\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1, i\lesssim k} \Delta_k u \Delta_i v \Delta_j w\end{aligned}$$

The commutativity of all the summation symbols follow from the absolute convergence of the sum, which is from the estimation

$$\begin{aligned}&\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\sum_{k\gtrsim i} \|\Delta_k u \Delta_i v \Delta_j w\|_{L^\infty}\\ \leqslant& \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\sum_{k\gtrsim i} \|\Delta_k u\|_{L^\infty}\|\Delta_i v\|_{L^\infty}\|\Delta_j w\|_{L^\infty}\\ \leqslant& \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\sum_{k\gtrsim i} 2^{-k\alpha}\|u\|_\alpha 2^{-i\beta}\|v\|_\beta 2^{-j\gamma}\|w\|_\gamma\\ \simeq& \sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1} 2^{-i(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta 2^{-j\gamma}\|w\|_\gamma\\ \simeq& \sum_{j\geqslant -1} 2^{-j(\alpha+\beta+\gamma)}\|u\|_\alpha\|v\|_\beta\|w\|_\gamma\end{aligned}$$

where we used $\alpha > 0, \alpha + \beta + \gamma > 0$.

For any $k \geqslant -1$, by lemma 3.18 there exists a ball $\tilde{B}'$ such that the Fourier transform of the sum

$$\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1, i\lesssim k} \Delta_k u \Delta_i v \Delta_j w = \sum_{j\geqslant -1, j\lesssim k}\sum_{i:|i-j|\leqslant 1} \Delta_k u \Delta_i v \Delta_j w$$

is supported in $2^k\tilde{B}'$. Moreover, we have

$$\begin{aligned}
&\left\|\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1, i\lesssim k}\Delta_k u\Delta_i v\Delta_j w\right\|_{L^\infty}\\
\leqslant&\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1, i\lesssim k}\|\Delta_k u\|_{L^\infty}\|\Delta_i v\|_{L^\infty}\|\Delta_j w\|_{L^\infty}\\
\leqslant&\|u\|_\alpha\|v\|_\beta\|w\|_\gamma\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1, i\lesssim k}2^{-k\alpha}2^{-i\beta}2^{-j\gamma}\\
=&\|u\|_\alpha\|v\|_\beta\|w\|_\gamma\sum_{j\geqslant -1, j\lesssim k}\sum_{i:|i-j|\leqslant 1}2^{-k\alpha}2^{-i\beta}2^{-j\gamma}\\
\simeq&\|u\|_\alpha\|v\|_\beta\|w\|_\gamma\sum_{j\geqslant -1, j\lesssim k}2^{-k\alpha}2^{-j(\beta+\gamma)}\\
\simeq&2^{-k(\alpha+\beta+\gamma)}\|u\|_\alpha\|v\|_\beta\|w\|_\gamma
\end{aligned}$$

where we used $\beta+\gamma<0$ in the last line. Since $\alpha+\beta+\gamma>0$, we use the lemma 3.15 to get

$$\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\Delta_{\gtrsim i}u\Delta_i v\Delta_j w\in\mathcal{C}^{\alpha+\beta+\gamma}$$

and

$$\begin{aligned}
&\left\|\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}\Delta_{\gtrsim i}u\Delta_i v\Delta_j w\right\|_{\alpha+\beta+\gamma}\\
\lesssim&\sup_{l\geqslant -1}\left\{2^{l(\alpha+\beta+\gamma)}\left\|\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1, i\lesssim l}\Delta_l u\Delta_i v\Delta_j w\right\|_{L^\infty}\right\}\\
\lesssim&\sup_{l\geqslant -1}\{2^{l(\alpha+\beta+\gamma)}2^{-l(\alpha+\beta+\gamma)}\|u\|_\alpha\|v\|_\beta\|w\|_\gamma\}\\
=&\|u\|_\alpha\|v\|_\beta\|w\|_\gamma
\end{aligned}$$

Next, let us look at the first term. Clearly, for any $j\geqslant -1$, the Fourier transform of the term

$$\sum_{i:|i-j|\leqslant 1}(\Delta_i(\Delta_{\lesssim i}u\prec v)-\Delta_{\lesssim i}u\Delta_i v)\Delta_j w$$

is supported in some ball $2^j\tilde{B}''$. By using out previous lemma, we have the estimation

$$\begin{aligned}
&\left\|\sum_{i:|i-j|\leqslant 1}(\Delta_i(\Delta_{\lesssim i}u\prec v)-\Delta_{\lesssim i}u\Delta_i v)\Delta_j w\right\|_{L^\infty}\\
\leqslant&\sum_{i:|i-j|\leqslant 1}\|\Delta_i(\Delta_{\lesssim i}u\prec v)-\Delta_{\lesssim i}u\Delta_i v\|_{L^\infty}\|\Delta_j w\|_{L^\infty}\\
\lesssim&\sum_{i:|i-j|\leqslant 1}2^{-i(\alpha+\beta)}\|\Delta_{\lesssim i}u\|_\alpha\|v\|_\beta 2^{-j\gamma}\|w\|_\gamma\\
\lesssim&\sum_{i:|i-j|\leqslant 1}2^{-i(\alpha+\beta)}\|u\|_\alpha\|v\|_\beta 2^{-j\gamma}\|w\|_\gamma\\
\simeq&2^{-j(\alpha+\beta+\gamma)}\|u\|_\alpha\|v\|_\beta\|w\|_\gamma
\end{aligned}$$

where we used

$$\begin{aligned}\|\Delta_{\lesssim i}u\|_{\alpha} &= \sup_{l\geqslant -1} 2^{l\alpha}\|\Delta_l\Delta_{\lesssim i}u\|_{L^\infty}\\ &= \sup_{l\geqslant -1} 2^{l\alpha}\left\|\sum_{k\lesssim i,k\sim l}\Delta_l\Delta_k u\right\|_{L^\infty}\\ &\leqslant \sup_{l\geqslant -1} 2^{l\alpha}\sum_{k\lesssim i,k\sim l}\|\Delta_k\Delta_l u\|_{L^\infty}\\ &\lesssim \sup_{l\geqslant -1} 2^{l\alpha}\sum_{k\lesssim i,k\sim l}\|\Delta_l u\|_{L^\infty}\\ &\simeq \|u\|_{\alpha}\end{aligned}$$

Again, since $\alpha+\beta+\gamma>0$, we use the lemma 3.15 to get

$$\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}(\Delta_i(\Delta_{\lesssim i}u\prec v)-\Delta_{\lesssim i}u\Delta_i v)\Delta_j w\in\mathcal{C}^{\alpha+\beta+\gamma}$$

and

$$\begin{aligned}&\left\|\sum_{j\geqslant -1}\sum_{i:|i-j|\leqslant 1}(\Delta_i(\Delta_{\lesssim i}u\prec v)-\Delta_{\lesssim i}u\Delta_i v)\Delta_j w\right\|_{\alpha+\beta+\gamma}\\ \lesssim\ & \sup_{l\geqslant -1}\left\{2^{l(\alpha+\beta+\gamma)}\left\|\sum_{i:|i-j|\leqslant 1}(\Delta_i(\Delta_{\lesssim i}u\prec v)-\Delta_{\lesssim i}u\Delta_i v)\Delta_l w\right\|_{L^\infty}\right\}\\ \lesssim\ & \sup_{l\geqslant -1}\{2^{l(\alpha+\beta+\gamma)}2^{-l(\alpha+\beta+\gamma)}\|u\|_{\alpha}\|v\|_{\beta}\|w\|_{\gamma}\}\\ =\ & \|u\|_{\alpha}\|v\|_{\beta}\|w\|_{\gamma}\end{aligned}$$

So in summary, we have for any $u,v,w\in\mathcal{S}$, the following commutator estimate

$$\|C(u,v,w)\|_{\alpha+\beta+\gamma}\lesssim\|u\|_{\alpha}\|v\|_{\beta}\|w\|_{\gamma}$$

is true.

Now we need to extend this bounded map to the whole space by continuity. To do this, we first choose slightly smaller regularity exponents $\alpha'\in(0,1)$, $\beta',\gamma'\in\mathbb{R}$ such that $\alpha'<\alpha$, $\beta'<\beta$, $\gamma'<\gamma$ and $\alpha'+\beta'+\gamma'>0$, $\beta'+\gamma'<0$. Since we know from lemma 3.9 the space $\mathcal{C}^{\alpha}\times\mathcal{C}^{\beta}\times\mathcal{C}^{\gamma}$ is contained in the closure of the Schwartz functions in space $\mathcal{C}^{\alpha'}\times\mathcal{C}^{\beta'}\times\mathcal{C}^{\gamma'}$, and the space $\mathcal{C}^{\alpha+\beta+\gamma}$ is contained in the closure of the Schwartz functions in space $\mathcal{C}^{\alpha'+\beta'+\gamma'}$. So we first obtain the estimate for Schwartz functions in $\mathcal{C}^{\alpha'}\times\mathcal{C}^{\beta'}\times\mathcal{C}^{\gamma'}$ and then extend the commutator continuously to get a definition of the commutator for any three elements in $\mathcal{C}^{\alpha}\times\mathcal{C}^{\beta}\times\mathcal{C}^{\gamma}$, the extension is independent of α',β',γ' since $\mathcal{C}^{\alpha}\times\mathcal{C}^{\beta}\times\mathcal{C}^{\gamma}$ is continuously embedded in the space $\mathcal{C}^{\alpha'}\times\mathcal{C}^{\beta'}\times\mathcal{C}^{\gamma'}$. And the desired boundedness follows from

$$\begin{aligned}\|C(u,v,w)\|_{\alpha+\beta+\gamma} &= \lim_{\alpha'\to\alpha^-,\beta'\to\beta^-,\gamma'\to\gamma^-}\|C(u,v,w)\|_{\alpha'+\beta'+\gamma'}\\ &\lesssim \lim_{\alpha'\to\alpha^-,\beta'\to\beta^-,\gamma'\to\gamma^-}\|u\|_{\alpha'}\|v\|_{\beta'}\|w\|_{\gamma'}\\ &= \|u\|_{\alpha}\|v\|_{\beta}\|w\|_{\gamma}\end{aligned}$$

where the second inequality may have a constant that depends on α', β', γ' but the limit exists and is finite, which can be seen easily from previous lemmas. $\square$

Now we prove the paralinearization theorem, which is useful when dealing with nonlinearity.

Theorem 3.23. (Paralinearization) *Suppose $\alpha \in (0,1), \beta \in (0,\alpha]$, and let $F \in C^{1,\beta/\alpha}$. The map $R_F\colon \mathcal{C}^\alpha \to \mathcal{C}^{\alpha+\beta}$ defined by the equation*

$$F(u) = F'(u) \prec u + R_F(u)$$

for any $u \in \mathcal{C}^\alpha$, is locally bounded. that is

$$\|R_F(u)\|_{\alpha+\beta} \lesssim \|F\|_{C^{1,\beta/\alpha}}(1 + \|u\|_\alpha^{1+\beta/\alpha})$$

If $F \in C^{2,\beta/\alpha}$, then R_F is locally Lipschitz continuous, that is

$$\|R_F(u) - R_F(v)\|_{\alpha+\beta} \lesssim \|F\|_{C^{2,\beta/\alpha}}(1 + \|u\|_\alpha + \|v\|_\alpha)^{1+\beta/\alpha}\|u - v\|_\alpha$$

for all $u, v \in \mathcal{C}^\alpha$.

Proof. By definition, we have

$$\begin{aligned} R_F(u) &= F(u) - F'(u) \prec u \\ &= \sum_{i \geqslant -1} (\Delta_i F(u) - \Delta_{<i-1} F'(u) \Delta_i u) \end{aligned}$$

Note that there is a ball $\tilde{B}$, such that for any $i \geqslant -1$, the Fourier transform of $\Delta_i F(u) - \Delta_{<i-1} F'(u) \Delta_i u$ is supported in ball $2^i \tilde{B}$. Also notice that

$$\int_{\mathbb{R}^n} K_{-1}(x) dx = \int_{\mathbb{R}^n} \mathcal{F}^{-1}(\rho_{-1})(x) e^{-i\langle 0,x\rangle} dx = \rho_{-1}(0) = 1$$

and for $i \geqslant 0$, we have

$$\int_{\mathbb{R}^n} K_i(x) dx = \int_{\mathbb{R}^n} \mathcal{F}^{-1}(\rho_i)(x) e^{-i\langle 0,x\rangle} dx = \rho_i(0) = 0$$

We use the fact that $\mathcal{C}^\alpha \cong C^{0,\alpha}$ to identify u as a $\alpha-$Hölder continuous function.

If $i < 1$, we have

$$\begin{aligned} & \|\Delta_i F(u) - \Delta_{<i-1} F'(u) \Delta_i u\|_{L^\infty} \\ = & \|\Delta_i F(u)\|_{L^\infty} \\ = & \left\| \int_{\mathbb{R}^n} K_i(x-y) F(u(y)) dy \right\|_{L^\infty} \\ = & \left\| \int_{\mathbb{R}^n} K_i(x-y)(F(u(x)) - F(u(y))) dy - F(u(x)) \int_{\mathbb{R}^n} K_i(x-y) dy \right\|_{L^\infty} \\ \leqslant & \|F\|_{L^\infty} + \int_{\mathbb{R}^n} |K_i(x-y)| \cdot \|F\|_{C^1} \|u\|_\alpha |x-y|^\alpha dy \\ \lesssim & \|F\|_{C^1}(1 + \|u\|_\alpha) \end{aligned}$$

If $i \geqslant 1$, we have

$$\begin{aligned}
& \Delta_i F(u)(x) - \Delta_{<i-1} F'(u) \Delta_i u(x) \\
= & \int_{\mathbb{R}^n} K_i(x-y) F(u(y)) dy - \int_{\mathbb{R}^n} K_i(x-y) u(y) dy \int_{\mathbb{R}^n} K_{<i-1}(x-z) F'(u(z)) dz \\
= & \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_i(x-y) K_{<i-1}(x-z)(F(u(y)) - F'(u(z))u(y)) dy dz \\
= & \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_i(x-y) K_{<i-1}(x-z)(F(u(y)) - F(u(z)) - F'(u(z))(u(y)-u(z))) dy dz
\end{aligned}$$

By mean value theorem, there exists $\xi \in (0,1)$, such that

$$\begin{aligned}
& |F(u(y)) - F(u(z)) - F'(u(z))(u(y)-u(z))| \\
= & |(F'(\xi u(y) + (1-\xi)u(z)) - F'(u(z)))(u(y)-u(z))| \\
\leqslant & \|F\|_{C^{1,\beta/\alpha}} |\xi u(y) + (1-\xi)u(z) - u(z)|^{\beta/\alpha} |u(y)-u(z)| \\
= & \|F\|_{C^{1,\beta/\alpha}} |\xi|^{\beta/\alpha} |u(y)-u(z)|^{1+\beta/\alpha} \\
\leqslant & \|F\|_{C^{1,\beta/\alpha}} \|u\|_\alpha^{1+\beta/\alpha} |y-z|^{\alpha(1+\beta/\alpha)}
\end{aligned}$$

Thus

$$\begin{aligned}
& |\Delta_i F(u)(x) - \Delta_{<i-1} F'(u) \Delta_i u(x)| \\
\leqslant & \|F\|_{C^{1,\beta/\alpha}} \|u\|_\alpha^{1+\beta/\alpha} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |K_i(x-y) K_{<i-1}(x-z)| \cdot |y-z|^{\alpha+\beta} dy dz \\
= & \|F\|_{C^{1,\beta/\alpha}} \|u\|_\alpha^{1+\beta/\alpha} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |K_i(y) K_{<i-1}(z)| \cdot |y-z|^{\alpha+\beta} dy dz
\end{aligned}$$

Now we try to estimate the last integral

$$\begin{aligned}
& \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |K_i(y) K_{<i-1}(z)| \cdot |y-z|^{\alpha+\beta} dy dz \\
= & \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |2^{in} K_0(2^i y) \cdot 2^{(i-1)n} K_{-1}(2^{i-1} z)| \cdot |y-z|^{\alpha+\beta} dy dz \\
= & 2^{-n} 2^{-i(\alpha+\beta)} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |K_0(y) K_{-1}(2^{-1} z)| \cdot |y-z|^{\alpha+\beta} dy dz \\
\simeq & 2^{-i(\alpha+\beta)}
\end{aligned}$$

In summary, we get

$$|\Delta_i F(u)(x) - \Delta_{<i-1} F'(u) \Delta_i u(x)| \lesssim 2^{-i(\alpha+\beta)} \|F\|_{C^{1,\beta/\alpha}} \|u\|_\alpha^{1+\beta/\alpha}$$

By lemma 3.15, we know that $R_F(u) \in \mathcal{C}^{\alpha+\beta}$ and

$$\begin{aligned}
\|R_F(u)\|_{\alpha+\beta} & \lesssim \sup_{j \geqslant -1} \{2^{j(\alpha+\beta)} \|\Delta_i F(u) - \Delta_{<i-1} F'(u) \Delta_i u\|_{L^\infty}\} \\
& \lesssim \|F\|_{C^{1,\beta/\alpha}} (1 + \|u\|_\alpha^{1+\beta/\alpha})
\end{aligned}$$

which concludes the proof of the first statement.

The second statement can be proved in a similar way. By definition, we have

$$\begin{aligned}
R_F(u) - R_F(v) & = (F(u) - F(v)) - (F'(u) \prec u - F'(v) \prec v) \\
& = \sum_{i \geqslant -1} (\Delta_i (F(u) - F(v)) - (\Delta_{<i-1} F'(u) \Delta_i u - \Delta_{<i-1} F'(v) \Delta_i v))
\end{aligned}$$

Note that there is a ball $\tilde{B}'$, such that for any $i \geqslant -1$, the Fourier transform of $\Delta_i(F(u) - F(v)) - (\Delta_{<i-1}F'(u)\Delta_i u - \Delta_{<i-1}F'(v)\Delta_i v)$ is supported in ball $2^i\tilde{B}'$.

If $i < 1$, we have

$$\begin{aligned}
& \|\Delta_i(F(u) - F(v)) - (\Delta_{<i-1}F'(u)\Delta_i u - \Delta_{<i-1}F'(v)\Delta_i v)\|_{L^\infty} \\
= & \|\Delta_i(F(u) - F(v))\|_{L^\infty} \\
= & \left\| \int_{\mathbb{R}^n} K_i(x-y)[F(u(y)) - F(v(y))]dy \right\|_{L^\infty} \\
= & \left\| \int_{\mathbb{R}^n} K_i(x-y)[F'(\xi u(y) + (1-\xi)v(y))(u(x) - v(x)) - (F(u(y)) - F(v(y)))]dy - \right. \\
& \left. (u(x) - v(x)) \int_{\mathbb{R}^n} K_i(x-y)F'(\xi u(y) + (1-\xi)v(y))dy \right\|_{L^\infty} \\
\lesssim & \|F\|_{C^1}\|u - v\|_\alpha + \int_{\mathbb{R}^n} |K_i(x-y)| \cdot \|F\|_{C^1}\|u - v\|_\alpha |x-y|^\alpha dy \\
\simeq & \|F\|_{C^1}\|u - v\|_\alpha
\end{aligned}$$

where ξ depends on y such that

$$F'(\xi u(y) + (1-\xi)v(y))(u(y) - v(y)) = F(u(y)) - F(v(y))$$

whose existence is given by the mean value theorem.

If $i \geqslant 1$, we have

$$\begin{aligned}
& \Delta_i(F(u) - F(v))(x) - (\Delta_{<i-1}F'(u)\Delta_i u - \Delta_{<i-1}F'(v)\Delta_i v)(x) \\
= & \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_i(x-y)K_{<i-1}(x-z)[(F(u(y)) - F(v(y))) - \\
& (F'(u(z))u(y) - F'(v(z))v(y))]\, dydz \\
= & \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} K_i(x-y)K_{<i-1}(x-z)[(F(u(y)) - F(v(y))) - (F(u(z)) - F(v(z))) \\
& -(F'(u(z))(u(y) - u(z)) - F'(v(z))(v(y) - v(z)))]\, dydz
\end{aligned}$$

□

Corollary 3.24. *Suppose* $\alpha \in (0,1), \beta \in (0, \alpha]$ and $\gamma < 0$ *which satisfies* $\alpha + \beta + \gamma > 0$ *and* $\alpha + \gamma < 0$, *let* $F \in C^{1,\beta/\alpha}$. *Then there exists a locally bounded map* $\Pi_F: \mathcal{C}^\alpha \times \mathcal{C}^\gamma \to \mathcal{C}^{\alpha+\beta+\gamma}$ *such that*

$$F(u) \circ v = F'(u)(u \circ v) + \Pi_F(u, v)$$

for any $u \in \mathcal{C}^\alpha$ *and smooth* $v \in \mathcal{C}^\gamma$. *Here the locally boundedness means the following inequality*

$$\|\Pi_F(u, v)\|_{\alpha+\beta+\gamma} \lesssim \|F\|_{C^{1,\beta/\alpha}}(1 + \|u\|_\alpha^{1+\beta/\alpha})\|v\|_\gamma$$

If $F \in C^{2,\beta/\alpha}$, *then is* Π_F *locally Lipschitz continuous*

$$\begin{aligned}
& \|\Pi_F(u_1, v_1) - \Pi_F(u_2, v_2)\|_{\alpha+\beta+\gamma} \\
\lesssim & \|F\|_{C^{2,\beta/\alpha}}(1 + (\|u_1\|_\alpha + \|u_2\|_\alpha)^{1+\beta/\alpha} + \|v_2\|_\gamma)(\|u_1 - u_2\|_\alpha + \|v_1 - v_2\|_\gamma)
\end{aligned}$$

for any $u_1, u_2 \in \mathcal{C}^\alpha$ *and* $v_1, v_2 \in \mathcal{C}^\gamma$.

Proof. For any $u \in \mathcal{C}^\alpha$ and smooth $v \in \mathcal{C}^\gamma$, we have

$$\begin{aligned}\Pi_F(u,v) &= F(u)\circ v - F'(u)(u\circ v)\\ &= (F'(u)\prec u)\circ v - F'(u)(u\circ v) + R_F(u)\circ v\\ &= C(F'(u),u,v) + R_F(u)\circ v\end{aligned}$$

then we use paralinearization and commutator estimate to get

$$\begin{aligned}\|\Pi_F(u,v)\|_{\alpha+\beta+\gamma} &= \|C(F'(u),u,v) + R_F(u)\circ v\|_{\alpha+\beta+\gamma}\\ &\leqslant \|C(F'(u),u,v)\|_{\alpha+\beta+\gamma} + \|R_F(u)\circ v\|_{\alpha+\beta+\gamma}\\ &\lesssim \|F'(u)\|_\beta \|u\|_\alpha \|v\|_\gamma + \|R_F(u)\|_{\alpha+\beta}\|v\|_\gamma\\ &\lesssim \|F'\|_{\beta/\alpha}\|u\|_\alpha^{1+\beta/\alpha}\|v\|_\gamma + \|F\|_{C^{1,\beta/\alpha}}(1+\|u\|_\alpha^{1+\beta/\alpha})\|v\|_\gamma\\ &\lesssim \|F\|_{C^{1,\beta/\alpha}}(1+\|u\|_\alpha^{1+\beta/\alpha})\|v\|_\gamma\end{aligned}$$

The last inequality follows from a similar estimate. ☐

Converse to the paralinearization theorem, the collection of distributions that look like some reference distribution in the sense of regularity, or intuitively a function of the reference distribution, should be important, we have following concept.

Definition 3.25. (Paracontrolled distribution) *Suppose $\alpha \in (0,1)$, $\beta \in (0,\alpha]$ and $Z \in \mathcal{C}^\alpha$. A distribution $u \in \mathcal{C}^\alpha$ is called paracontrolled by Z, if there exists $u' \in \mathcal{C}^\beta$ such that*

$$u^\# := u - u' \prec Z \in \mathcal{C}^{\alpha+\beta}$$

the collection of distributions paracontrolled by Z is denoted by $\mathcal{D}^\beta(Z)$, and to emphasise this structure, we write $(u,u',u^\#) \in \mathcal{D}^\beta(Z)$. The norm

$$\|(u,u',u^\#)\|_{\mathcal{D}^\beta(Z)} := \|u\|_\alpha + \|u'\|_\beta + \|u^\#\|_{\alpha+\beta}$$

is then well-defined.

It can be shown that $\mathcal{D}^\beta(Z)$ is a Banach space.

3.4 Higher Order Paracontrolled Calculus

In the paper [5], they developed a higher order paracontrolled calculus, which generalizes the paralinearization theorem.

Theorem 3.26. *Suppose $\alpha \in (0,1), k \in \mathbb{N}^+$, and let $F \in C^{k+1}$ whose 4-th derivative is bounded. There is a remainder map $R_F: \mathcal{C}^\alpha \to \mathcal{C}^{(k+1)\alpha}$ such that*

$$F(u) = \sum_{n=1}^{k} \frac{1}{n!}\sum_{j=0}^{n-1}(-1)^j\binom{n}{j}(u^j F^{(n)}(u)) \prec (u^{n-j}) + R_F(u)$$

for any $u \in \mathcal{C}^\alpha$.

For more informations on the estimate of remainder map, and various commutator estimates, see [5].

4 Gaussian Analysis

In this chapter we will introduce the subject of Gaussian analysis, which is about the theory of white noise.

4.1 White Noise

We begin by a heuristic discussion of what is white noise. Before introduce the white noise, we introduce the concept of generalized random processes, see [31] and the references there in.

A generalized function or a tempered distribution $u\in\mathcal{S}'(\mathbb{R}^n)$ is used to model a physical field, for example the temperature field, charge density field, electromagnetic field, etc. A measurement by a physical instrument is modeled by a test function $f\in\mathcal{S}(\mathbb{R}^n)$, since usually we can only measure certain average of the field instead of measure its accurate value at a point in the space. The result of the measurement is given by the value $\langle u, f\rangle$.

If the field is random, then we expect the result of the measurement by $f\in\mathcal{S}(\mathbb{R}^n)$ gives a random variable instead of just one deterministic number, that is we have a probability measure on the space of tempered distribution which models the random field and the measurement $\langle u, f\rangle$ is then a random variable. Moreover we expect that the resulting random measurement result look similar if we use similar instrument, this requires some kind of continuity.

Definition 4.1. *A generalized random process or a random field on $\mathbb{R}^n$ is a map Φ from the space of Schwartz functions $\mathcal{S}(\mathbb{R}^n)$ to the space of random variables on a fixed probability space $(\Omega,\mathcal{F},\mathbb{P})$, such that:*

(1) Linearity: $\Phi(c_1f_1+c_2f_2)=c_1\Phi(f_1)+c_2\Phi(f_2)$ almost surely, for any $c_1,c_2\in\mathbb{R}$ and $f_1,f_2\in\mathcal{S}(\mathbb{R}^n)$;

(2) Continuity: If we have m sequences of test functions $\{f_n^{(j)}\}_{1\leqslant j\leqslant m, n\in\mathbb{N}}$ and $j-\mathrm{th}$ sequence converges to $f^{(j)}$ in $\mathcal{S}(\mathbb{R}^n)$ for each $1\leqslant j\leqslant m$, then the random vector $(\Phi(f_n^{(1)}),\ldots,\Phi(f_n^{(m)}))$ converges in distribution to the random vector $(\Phi(f^{(1)}),\ldots,\Phi(f^{(m)}))$ as $n\to\infty$.

The central limit theorem tells us the combined effect of an infinite number of weakly correlated random variables is given by a Gaussian random variable. So it is reasonable to expect that any measurement of a white noise ξ, which is the combined result due to an infinite number of random background signal, is given by a Gaussian random variable, that is all the random variables $\xi(f)$ are Gaussian and the image of this random field is a Gaussian linear space (linear subspace of all random variables such that any finite elements are distributed as centered joint Gaussian). It is also natural to assume that the white noise at different spacial regions are independent from each other and the noise is the same under translation or rotations, that is we can consider the Dirac normalization condition

$$\mathbb{E}[\xi(\delta(\cdot-x))\xi(\delta(\cdot-y))]=\delta(x-y)$$

Formally, for any $f_1, f_2 \in \mathcal{S}(\mathbb{R}^n)$, we have

$$\begin{aligned}
\mathbb{E}[\xi(f_1)\xi(f_2)] &= \mathbb{E}\Big[\xi\Big(\int f_1(x)\delta(x-\cdot)dx\Big)\xi\Big(\int f_2(y)\delta(y-\cdot)dy\Big)\Big]\\
&= \int f_1(x)f_2(y)\mathbb{E}[\xi(\delta(x-\cdot))\xi(\delta(y-\cdot))]dxdy\\
&= \iint f_1(x)f_2(y)\delta(x-y)dxdy\\
&= \int f_1(x)f_2(x)dx
\end{aligned}$$

which is the L^2-inner product on the space of Schwartz functions. One can complete the space of Schwartz functions to $L^2(\mathbb{R}^n)$, since the Schwartz functions form a dense subspace of $L^2(\mathbb{R}^n)$. And convergence in distribution is equivalent to convergence in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ for a Gaussian Hilbert space, then we can complete this generalized random process to have the following formal definition.

Definition 4.2. *A white noise ξ on $\mathbb{R}^n$ is an isometry from $L^2(\mathbb{R}^n)$ to a Gaussian Hilbert space (a closed subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ which consists only centered Gaussian random variables), that is*

$$\mathbb{E}[\xi(f_1)\xi(f_2)] = \langle f_1, f_2\rangle_{L^2(\mathbb{R}^n)}$$

for any $f_1, f_2 \in L^2(\mathbb{R}^n)$.

Details about the Gaussian Hilbert spaces can be founded in the book [30].

4.2 Wick product

Now we study the Wick product of random variables. Let's first consider the case of a single random variable.

Definition 4.3. *Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable X, such that $\mathbb{E}[|X|^n] < \infty$ for any $n \in \mathbb{N}$. Define a sequence of random variables $\{:X^n:\}_{n\in\mathbb{N}}$, which are polynomial functions of X, recursively by*

(1) $:X^0: = 1$*;*

(2) $\frac{\partial}{\partial X}:X^n: = n:X^{n-1}:$ *for* $n = 1, 2, 3, \ldots$*;*

(3) $\mathbb{E}[:X^n:] = 0$ *for* $n = 1, 2, 3, \ldots$

in the second equation, we use the formal derivative of formal power series, and the equality is in the sense of the corresponding coefficients of formal power series are the same. We call $:X^n:$ the n-th Wick power of X.

Note that this definition depends on the probability space and the given random variable on it. The motivation for the third equation comes from quantum field theory, where one need to get rid of the infinities coming from the vacuum expectation. The first few Wick powers are given in the following example.

Example 4.4. With the same assumptions in the definition, we have

$$\begin{aligned} :X^1: &= X - \mathbb{E}[X] \\ :X^2: &= X^2 - 2\mathbb{E}[X]X - \mathbb{E}[X^2] + 2\mathbb{E}[X]^2 \\ :X^3: &= X^3 - 3\mathbb{E}[X]X^2 - 3\mathbb{E}[X^2]X + 6\mathbb{E}[X]^2X - \mathbb{E}[X]^3 + 6\mathbb{E}[X]\mathbb{E}[X^2] - 6\mathbb{E}[X^3] \end{aligned}$$

We study the Wick powers $\{:X^n:\}_{n\in\mathbb{N}}$ when X is a Gaussian random variable with mean 0. The Hermite polynomials are defined by the power series of the following analytical function

$$e^{xt-\frac{1}{2}t^2} = \sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!}.$$

First few Hermite polynomials are

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= x \\ H_2(x) &= x^2 - 1 \\ H_3(x) &= x^3 - 3x \end{aligned}$$

Lemma 4.5. *The Hermite polynomials has following properties:*

(1) $H_n(x)$ is a polynomial of degree n with coefficient of x^n equal to 1;

(2) $H_n'(x) = nH_{n-1}(x)$ for $n \geqslant 1$;

(3) $H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}$.

Proof. Property (1) is clear from multiplying the power series of e^{xt} and $e^{-\frac{1}{2}t^2}$. The property (2) follows from the equation

$$\partial_x e^{xt-\frac{1}{2}t^2} = t\, e^{xt-\frac{1}{2}t^2}$$

by comparing the coefficients of power series on both sides. For the property (3), we observe that

$$\frac{d^n}{dx^n} e^{-\frac{x^2}{2}} = (-1)^n \frac{d^n}{dt^n} e^{-\frac{1}{2}(x-t)^2}|_{t=0}$$

and

$$\begin{aligned} e^{xt-\frac{1}{2}t^2} &= e^{xt-\frac{1}{2}t^2-\frac{1}{2}x^2} e^{\frac{x^2}{2}} \\ &= e^{\frac{x^2}{2}} e^{-\frac{1}{2}(x-t)^2} \\ &= e^{\frac{x^2}{2}} \sum_{n=0}^{\infty} \frac{d^n}{dt^n} e^{-\frac{1}{2}(x-t)^2}|_{t=0} \frac{t^n}{n!} \end{aligned}$$

thus the result follows by comparing the coefficients of power series on both sides. □

Lemma 4.6. *Suppose $X\sim\mathcal{N}(0,\sigma^2)$ where $\sigma>0$, then*

$$:X^n:=\sigma^n H_n\left(\frac{X}{\sigma}\right).$$

Proof. We check that the formula on the right hand side satisfy the definition of Wick power. Clearly $\sigma^0 H_0(\frac{X}{\sigma})=1$. For the second condition, we have

$$\begin{aligned}\frac{\partial}{\partial X}\sigma^{n+1}H_{n+1}\left(\frac{X}{\sigma}\right) &= \sigma^n\frac{\partial}{\partial\frac{X}{\sigma}}H_{n+1}\left(\frac{X}{\sigma}\right)\\ &= (n+1)\sigma^n H_n\left(\frac{X}{\sigma}\right).\end{aligned}$$

For the third condition, since $X\sim\mathcal{N}(0,\sigma^2)$, then $\frac{X}{\sigma}\sim\mathcal{N}(0,1)$, so we have

$$\begin{aligned}1 &= \mathbb{E}\left[e^{it\frac{X}{\sigma}+\frac{1}{2}t^2}\right]\\ &= \mathbb{E}\left[\sum_{n=0}^{\infty}H_n\left(\frac{X}{\sigma}\right)\frac{(it)^n}{n!}\right]\\ &= \sum_{n=0}^{\infty}\mathbb{E}\left[H_n\left(\frac{X}{\sigma}\right)\right]\frac{(it)^n}{n!}\end{aligned}$$

and then compare the coefficients on both sides. □

Remark 4.7. Notice that this lemma also works when $\sigma=0$, that is $X=0$ also surely. Since $H_n(x)$ is a polynomial of degree n, the right hand side of the equation is a polynomial function of two variables X and σ, thus $\sigma=0$ does not make any trouble.

Then under the assumption $X\sim\mathcal{N}(0,\sigma^2)$ for $\sigma>0$, we have the identity

$$e^{\frac{X}{\sigma}t-\frac{1}{2}t^2}=\sum_{n=0}^{\infty}\left(\frac{t}{\sigma}\right)^n\frac{:X^n:}{n!}$$

and if we replace $\frac{t}{\sigma}$ by t, equivalently we get

$$e^{Xt-\frac{1}{2}\sigma^2t^2}=\sum_{n=0}^{\infty}t^n\frac{:X^n:}{n!}.$$

Note that this formula also works when $\sigma=0$. A simple consequence of this formula is following lemma.

Lemma 4.8. *Suppose $(X,Y)\sim\mathcal{N}(0,\Sigma)$ is a Gaussian vector with covariance matrix Σ, then we have*

$$\mathbb{E}[:X^n::Y^m:]=\delta_{nm}n!\mathbb{E}[XY]^n$$

for any $n, m \in \mathbb{N}$.

Proof. Since we know that

$$e^{Xt-\frac{1}{2}\mathbb{E}[X^2]t^2} = \sum_{n=0}^{\infty} t^n \frac{:X^n:}{n!}, e^{Yr-\frac{1}{2}\mathbb{E}[Y^2]r^2} = \sum_{m=0}^{\infty} r^m \frac{:Y^m:}{m!}$$

and since $Xt + Yr$ is again a centered Gaussian random variable, we have

$$e^{Xt-\frac{1}{2}\mathbb{E}[X^2]t^2} e^{Yr-\frac{1}{2}\mathbb{E}[Y^2]r^2} = e^{(Xt+Yr)-\frac{1}{2}\mathbb{E}[(Xt+Yr)^2]} e^{\mathbb{E}[XY]tr}$$

then

$$\mathbb{E}\Big[e^{Xt-\frac{1}{2}\mathbb{E}[X^2]t^2} e^{Yr-\frac{1}{2}\mathbb{E}[Y^2]r^2}\Big] = \mathbb{E}\Big[e^{(Xt+Yr)-\frac{1}{2}\mathbb{E}[(Xt+Yr)^2]} e^{\mathbb{E}[XY]tr}\Big] = e^{\mathbb{E}[XY]tr}$$

thus

$$\begin{aligned}
\mathbb{E}\Big[e^{Xt-\frac{1}{2}\mathbb{E}[X^2]t^2} e^{Yr-\frac{1}{2}\mathbb{E}[Y^2]r^2}\Big] &= \mathbb{E}\left[\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} t^n \frac{:X^n:}{n!} r^m \frac{:Y^m:}{m!}\right] \\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{t^n r^m}{n!m!} \mathbb{E}[:X^n::Y^m:] \\
&= \sum_{n=0}^{\infty} \frac{(tr)^n}{n!} \mathbb{E}[XY]^n
\end{aligned}$$

hence the result follows by comparing the coefficients on both sides of the equation. □

Taking a special case of this lemma, if $X = Y \sim \mathcal{N}(0,1)$, we have

$$\mathbb{E}[:X^n::X^m:] = \delta_{nm} n!$$

for any $n, m \in \mathbb{N}$. In other words, we have

$$\mathbb{E}[H_n(X) H_m(X)] = \delta_{nm} n!$$

which is equivalent to say that the Hermite polynomials are orthogonal with respect to the weighted Lebesgue measure

$$\int_{\mathbb{R}} H_n(x) H_m(x) e^{-\frac{1}{2}x^2} dx = \delta_{nm} n!.$$

Wick product can be generalized to several random variables, and many similar properties can be proved. We only give a definition here without go into the properties, since we will not need that much.

Definition 4.9. *Fix a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and random variables* $X_1, \dots X_m$, *such that* $\mathbb{E}[|X_l|^n] < \infty$ *for any* $n \in \mathbb{N}$ *and* $1 \leqslant l \leqslant m$. *Define a sequence of random variables* $\{:X_1^{n_1} \dots X_m^{n_m}:\}_{n_1, \dots, n_m \in \mathbb{N}}$, *which are polynomial functions of* $X_1, \dots X_m$, *recursively by*

(1) $:X_1^0 \dots X_m^0: = 1$;

(2) $\frac{\partial}{\partial X_l}: X_1^{n_1} \ldots X_l^{n_l} \ldots X_m^{n_m}: = n_l : X_1^{n_1} \ldots X_l^{n_l - 1} \ldots X_m^{n_m}:$ *for all* $1 \leqslant l \leqslant m$ *and* $n_l > 0$*;*
(3) $\mathbb{E}[: X_1^{n_1} \ldots X_m^{n_m}:] = 0$ *for* $n_1 + \cdots + + n_m > 0$

in the second equation, we use the formal derivative of formal power series, and the equality is in the sense of the corresponding coefficients of formal power series are the same.

4.3 Gaussian Hypercontractivity and Regularity of White Noise

For Gaussian random variables, we have following estimates.

Theorem 4.10. (Gaussian Hypercontractivity) *Assume* X *is a Gaussian random variable on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and* $\mathbb{E}[X] = 0$*. Then for any* $p \geqslant 2$*, there is a consant* $c(p)$*, such that*

$$\mathbb{E}[|X|^p] \leqslant c(p) \mathbb{E}[X^2]^{p/2}$$

Proof. We have

$$\begin{aligned}
\mathbb{E}[|X|^p] =& \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{\mathbb{R}} |x|^p e^{-\frac{x^2}{2\sigma^2}} dx \\
=& 2 \frac{1}{(2\pi\sigma^2)^{1/2}} \int_0^\infty |\sigma\sqrt{2y}|^p e^{-\frac{(\sigma\sqrt{2y})^2}{2\sigma^2}} d\sigma\sqrt{2y} \\
=& \frac{2^{\frac{p+1}{2}} \sigma^p}{(2\pi)^{1/2}} \int_0^\infty y^{\frac{p+1}{2} - 1} e^{-y} dy \\
=& \frac{2^{\frac{p+1}{2}} \sigma^p}{(2\pi)^{1/2}} \Gamma\left(\frac{p+1}{2}\right) \\
=& \frac{2^{\frac{p+1}{2}}}{(2\pi)^{1/2}} \Gamma\left(\frac{p+1}{2}\right) \mathbb{E}[X^2]^{p/2}
\end{aligned}$$

where Γ is the Gamma function. Thus the result follows. □

There are generalizations to Wick product of Gaussian variables in a Gaussian Hilbert spaces, see theorem 3.50 in [30]. As an application, we compute the regularity of white noise.

Corollary 4.11. *Suppose* ξ *is a white noise on* d *dimensional torus* $\mathbb{T}^d$*, that is for any* $x, y \in \mathbb{R}^d$*, we have* $\mathbb{E}[\xi(x)\xi(y)] = \delta(x - y)$*, then* $\xi \in \mathcal{C}^{-\frac{d}{2} - \varepsilon}$ *for any* $\varepsilon > 0$*.*

Proof. Denote $\alpha = -\frac{d}{2}$, for any $p \geqslant 2$, by Besov embedding we have

$$\|\xi\|_{B_{\infty,\infty}^{\alpha - \frac{d}{p}}} \lesssim \|\xi\|_{B_{p,p}^\alpha}$$

thus

$$
\begin{aligned}
& \mathbb{E}\bigg[\|\xi\|^p_{\mathcal{C}^{\alpha-\frac{d}{p}}}\bigg] \\
\lesssim\ & \mathbb{E}[\|\xi\|^p_{B^\alpha_{p,p}}] \\
=\ & \mathbb{E}\Bigg[\sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}|\Delta_j\xi|^p_{L^p}\Bigg] \\
=\ & \sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}\int_{\mathbb{T}^d}\mathbb{E}[|\Delta_j\xi|^p]dx \\
\lesssim\ & \sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}\int_{\mathbb{T}^d}\mathbb{E}[(\Delta_j\xi)^2]^{p/2}dx \\
=\ & \sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}\int_{\mathbb{T}^d}\mathbb{E}\bigg[\int K_j(x-y_1)\xi(y_1)dy_1\int K_j(x-y_2)\xi(y_2)dy_2\bigg]^{p/2}dx \\
=\ & \sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}\int_{\mathbb{T}^d}\bigg[\int K_j(x-y_1)K_j(x-y_2)\mathbb{E}[\xi(y_1)\xi(y_2)]dy_1dy_2\bigg]^{p/2}dx \\
=\ & \sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}\int_{\mathbb{T}^d}\bigg[\int K_j(x-y_1)K_j(x-y_2)\delta(y_1-y_2)dy_1dy_2\bigg]^{p/2}dx \\
=\ & \sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}\|K_j\|^p_{L^2}\int_{\mathbb{T}^d}dx \\
=\ & \sum_{j\geqslant -1}^{\infty} 2^{j\alpha p}\cdot 2^{-jdp/2}\|K_0\|^p_{L^2}\int_{\mathbb{T}^d}dx \\
<\ & +\infty
\end{aligned}
$$

where we used the Gaussian hypercontractivity for Gaussian variables $\Delta_j\xi$. □

The case of white noise on $\mathbb{R}^d$ can be shown with a similar method, the only difference is we need some kind of modified Besov space.

5 Parabolic Anderson Model

In this chapter we will present a complete solution of parabolic Anderson model as an application of the tools we have developed so far. The main references are [23], [24] and [38].

5.1 The Model and Paracontrolled Distributions

The parabolic Anderson model is a stochastic partial differential equation on $[0,T)\times\mathbb{T}^2$

$$\partial_t u = \Delta u + \xi u$$

where ξ is spatial white noise, that is $\mathbb{E}[\xi(x)\xi(x')] = \delta(x-x')$. This equation described the diffusion in a random potential.

The first thing to do is to find the regularity of each term. We know the regularity of the spatial white noise $\xi \in \mathcal{C}^{-1-}$, which can be seen by noting that it can be written as tensor product of two one dimensional white noise, and one dimensional white noise can be think of derivative of Brownian motion which has regularity exponent $\frac{1}{2}-$. In general the white noise in d dimension has regularity $-\frac{d}{2}-$.

Then we can only expect the solution u to be in $\mathcal{C}^{1-}$, since the heat operator may improve the regularity of last term by 2. Observe that $(1-)+(-1-)<0$, so the term ξu is not well defined. Thus we identified where the problem is, next we discuss heuristically how to deal with the equation, in order to build up a solution theory by fixed point argument.

Suppose we have a solution $u\in\mathcal{C}^\alpha$, where $\alpha\in(2/3,1)$ (2/3 comes from the need for commutator estimate, we will see it later). Then the regularity of ξ is $\alpha-2$. Define Z to be the solution of $\partial_t Z = \Delta Z + \xi$, which is given by

$$Z(t) = \int_0^t P_{t-s}\xi ds$$

where P_t is the heat kernel. Then Z has regularity α according to Schauder estimate. We use the paraproduct decomposition to get

$$(\partial_t - \Delta)u = \xi \succ u + \xi \circ u + \xi \prec u$$

if we denote $L = \partial_t - \Delta$, we have

$$u = L^{-1}(\xi \succ u + \xi\circ u + \xi \prec u) = u \prec Z + [L^{-1}, u\prec]\xi + L^{-1}(\xi\circ u + \xi\prec u)$$

according to the paraproduct estimate, we know that $(\xi\prec u)\in\mathcal{C}^{2\alpha-2}$, suppose we can show that $(\xi\circ u)\in\mathcal{C}^{2\alpha-2}$, then the third term is in $\mathcal{C}^{2\alpha}$ by Schauder estimate, one can also show that the form of the second term is also in $\mathcal{C}^{2\alpha}$ (see lemma 18 of [24]). Thus we postulate the paracontrolled ansatz

$$u = u\prec Z + u^{\#}$$

with $u^{\#}\in\mathcal{C}^{2\alpha}$, which means we also consider the solution with some extra structure of being paracontrolled by Z. Having this at hand, we would like to ask what is the equation for $u^{\#}$? Since

$$Lu = L(u\prec Z + u^{\#}) = u\prec LZ + [L, u\prec]Z + Lu^{\#}$$

So

$$\begin{aligned} Lu^{\#} &= \xi u - u\prec LZ - [L, u\prec]Z \\ &= \xi\succ u + \xi\circ u + \xi\prec u - u\prec\xi - [L, u\prec]Z \\ &= \xi\circ(u\prec Z + u^{\#}) + \xi\prec u - [L, u\prec]Z \\ &= u(Z\circ\xi) + C(u, Z, \xi) + u^{\#}\circ\xi + \xi\prec u - [L, u\prec]Z \end{aligned}$$

and

$$u^{\#} = \int_0^t P_{t-s}(u(Z\circ\xi) + \xi\prec u - [L, u\prec]Z + C(u, Z, \xi) + u^{\#}\circ\xi)ds$$

Everything is well-defined (since $\alpha>\frac{2}{3}$), except the term $Z\circ\xi$, but this term does not involve unknowns, so that we can assume it is given at first hand. Construction of this term uses probability theory.

Then we have an equivalent system of equations

$$\begin{aligned} u &= u\prec Z + u^{\#} \\ u^{\#} &= \int_0^t P_{t-s}(u(Z\circ\xi) + \xi\prec u - [L, u\prec]Z + C(u, Z, \xi) + u^{\#}\circ\xi)ds \end{aligned}$$

with input $(u(0), \xi, Z, Z\circ\xi)\in\mathcal{C}^{\alpha}\times\mathcal{C}^{\alpha-2}\times\mathcal{C}^{\alpha}\times\mathcal{C}^{2\alpha-2}$, and we know that $u\in\mathcal{C}^{\alpha}$, $u^{\#}\in\mathcal{C}^{2\alpha}$. Now we can apply Picard iteration and Banach fixed point theorem to this system of equation, on the space of paracontrolled disrtribution $\mathcal{D}^{\alpha}$ for some suitable time interval $[0, T]$, to get a unique solution that continuously depends on the initial data $(u(0), \xi, Z, Z\circ\xi)$. We will do this in the following sections.

5.2 Schauder Estimates

Before the discussion we introduce the space needed for the discussion. We denote $C_T\mathcal{C}^{\alpha}$ by the continuous functions from $[0, T)$ to $\mathcal{C}^{\alpha}$, and $C_T^{\alpha'}\mathcal{C}^{\alpha}$ be the $\mathcal{C}^{\alpha}$ valued α'-Hölder continuous functions. We define $\mathscr{L}_T^{\alpha} = C_T\mathcal{C}^{\alpha}\cap C_T^{\alpha/2}L^{\infty}$ for $\alpha\in(0, 2)$, and we equip $\mathscr{L}_T^{\alpha}$ with the norm

$$\|\cdot\|_{\mathscr{L}_T^{\alpha}} = \max\{\|\cdot\|_{C_T\mathscr{C}^{\alpha}}, \|\cdot\|_{C_T^{\alpha/2}L^{\infty}}\}$$

We refine the definition of paracontrolled distribution for our convenience.

Definition 5.1. (Paracontrolled distribution) *Suppose $\alpha\in(0,1)$ and $Z\in\mathscr{L}_T^{\alpha}$. A distribution $u\in\mathscr{L}_T^{\alpha}$ is called paracontrolled by Z, if there exists $u'\in\mathscr{L}_T^{\alpha}$ such that*

$$u^{\#} := u - u'\prec Z\in\mathscr{L}_T^{2\alpha}$$

the collection of distributions paracontrolled by Z *is denoted by* $\mathcal{D}_T^\alpha(Z)$, *and to emphasize this structure, we write* $(u, u', u^\#) \in \mathcal{D}_T^\alpha(Z)$. *The norm*

$$\|(u, u', u^\#)\|_{\mathcal{D}_T^\alpha(Z)} := \|u\|_{\mathscr{L}_T^\alpha} + \|u'\|_{\mathscr{L}_T^\alpha} + \|u^\#\|_{\mathscr{L}_T^{2\alpha}}$$

is well-defined. $\mathcal{D}_T^\alpha(Z)$ *is a Banach space.*

Before talking about the Schauder estimate for the paracontrolled distributions, we first introduce the standard Schauder estimate.

Theorem 5.2. (Schauder Estimate) *Let* $\alpha \in (0,2)$, *let* $(P_t)_{t \geq 0}$ *be the semigroup generated by the periodic Laplacian on* $\mathbb{T}^d$. *For* $f \in C_T \mathcal{C}^{\alpha-2}$, *then the solution of*

$$L u = f, u(0) = 0,$$

is given by

$$u(t) = L^{-1} f(t) = \int_0^t P_{t-s} f(s)\, d s$$

for any $T > 0$, *moreover, we have the following estimates*

$$\|L^{-1} f\|_{\mathscr{L}_T^\alpha} \lesssim (1 + T) \|f\|_{C_T \mathcal{C}^{\alpha-2}}$$

and for any $u \in \mathcal{C}^\alpha$, *we have*

$$\|t \mapsto P_t u\|_{\mathscr{L}_T^\alpha} \lesssim \|u\|_{\mathscr{C}^\alpha}$$

We will not prove this theorem, see lemma 11 of [24] and references therein.

Theorem 5.3. (Schauder Estimate for the Paracontrolled Distribution) *Let* $\alpha \in (0,1)$, $\xi \in C_T \mathcal{C}^{\alpha-2}$ *and* $L Z = \xi$ *with* $Z(0) = 0$. *Let* $u \in \mathscr{L}_T^\alpha$, $f^\sharp \in C_T \mathcal{C}^{2\alpha-2}$, *and* $u_0 \in \mathcal{C}^{2\alpha}$. *Then* $(g, u, g - u \prec Z) \in \mathcal{D}_T^\alpha(Z)$, *where* g *solves*

$$L g = u \prec \xi + f^\sharp, \quad g(0) = u_0,$$

and we have

$$\|g\|_{\mathcal{D}_T^\alpha(Z)} \lesssim \|u_0\|_{2\alpha} + (1 + T)\, (\|u\|_{\mathscr{L}_T^\alpha} (1 + \|\xi\|_{C_T \mathcal{C}^{\alpha-2}}) + \|f^\sharp\|_{C_T \mathcal{C}^{2\alpha-2}})$$

for all $T > 0$.

If furthermore $\tilde{\xi}, \tilde{Z}, \tilde{u}, \tilde{f^\sharp}, \tilde{u}_0, \tilde{g}$ *satisfy the same assumptions as* $\xi, Z, u, f^\sharp, u_0, g$ *respectively, and if* $M = \max\{\|u\|_{\mathscr{L}_T^\alpha}, \|\tilde{\xi}\|_{C_T \mathcal{C}^{\alpha-2}}, 1\}$, *then*

$$d_{\mathcal{D}_T^\alpha(Z)}(g, \tilde{g}) \lesssim \|u_0 - \tilde{u}_0\|_{2\alpha} +$$

$$(1 + T)\, M\, (\|u - \tilde{u}\|_{\mathscr{L}_T^\alpha} + \|\xi - \tilde{\xi}\|_{C_T \mathcal{C}^{\alpha-2}} + \|f^\sharp - \tilde{f}^\sharp\|_{C_T \mathcal{C}^{2\alpha-2}}).$$

We will not prove this theorem, see theorem 7 of [24] and references therein.

5.3 Existence of Solutions and Renormalization

We want to solve the system of equations

$$\begin{aligned} u &= u \prec Z + u^{\#} \\ u^{\#} &= \int_0^t P_{t-s} (u(Z \circ \xi) + \xi \prec u - [L, u \prec] Z + C(u, Z, \xi) + u^{\#} \circ \xi) ds \end{aligned}$$

with input $(u(0), \xi, Z, Z \circ \xi) \in \mathcal{C}^{\alpha} \times \mathcal{C}^{\alpha-2} \times \mathcal{C}^{\alpha} \times \mathcal{C}^{2\alpha-2}$, and we expect solutions satisfy $u \in \mathcal{C}^{\alpha}, u^{\#} \in \mathcal{C}^{2\alpha}$.

We define the map $\Phi_T: \mathcal{D}_T^{\alpha}(Z) \to \mathcal{D}_T^{\alpha}(Z)$ by

$$\Phi_T \begin{pmatrix} u \\ u^{\#} \end{pmatrix} = \begin{pmatrix} u \prec Z + u^{\#} \\ \int_0^t P_{t-s} (u(Z \circ \xi) + \xi \prec u - [L, u \prec] Z + C(u, Z, \xi) + u^{\#} \circ \xi) ds \end{pmatrix}$$

where we only write down the derivative part and remainder part of the paracontrolled distribution, then we need to solve the fixed point problem $\Phi_T \begin{pmatrix} u \\ u^{\#} \end{pmatrix} = \begin{pmatrix} u \\ u^{\#} \end{pmatrix}$. Start from any element $(u, u^{\#}) \in \mathcal{D}_T^{\alpha}(Z)$, we have

$$L\Phi_T(u) = L(u \prec Z) + u(Z \circ \xi) + \xi \prec u - [L, u \prec] Z + C(u, Z, \xi) + u^{\#} \circ \xi = u \prec \xi + f^{\#}$$

where

$$f^{\#} = u(Z \circ \xi) + \xi \prec u + C(u, Z, \xi) + u^{\#} \circ \xi \in C_T \mathcal{C}^{2\alpha-2}$$

together with $\Phi_T(u)(0) = u_0$, then the Schauder estimate for the paracontrolled distributions tells us that $\Phi_T(u)$ is an element of $\mathcal{D}_T^{\alpha}(Z)$ with derivative u, and

$$\|\Phi_T(u)\|_{\mathcal{D}_T^{\alpha}(Z)} \lesssim \|u_0\|_{2\alpha} + (1 + T) \left(\|u\|_{\mathscr{L}_T^{\alpha}} (1 + \|\xi\|_{C_T \mathcal{C}^{\alpha-2}}) + \|f^{\sharp}\|_{C_T \mathcal{C}^{2\alpha-2}} \right)$$

Observe that since ξ is a spacial white noise, then $\|\xi\|_{C_T \mathcal{C}^{\alpha-2}} = \|\xi\|_{\mathcal{C}^{\alpha-2}}$ is a constant. The estimate of the term $\|f^{\sharp}\|_{C_T \mathcal{C}^{2\alpha-2}}$ is too technical to produce here, see the discussion below lemma 5.3 of [23], section 5.4 and theorem 8 of [24] and references therein. The crucial fact is that one can take time T small enough, such that the following inequality holds

$$\|\Phi_T(u)\|_{\mathcal{D}_T^{\alpha}(Z)} \leqslant C \|u\|_{\mathcal{D}_T^{\alpha}(Z)}$$

for some constant $C < 1$ which depends only on the data set $(u(0), \xi, Z, Z \circ \xi)$, hence one can run the Banach fixed point theorem to get the solution of the system of equations. Moreover, one can prove that this solution depends on the data set $(u(0), \xi, Z, Z \circ \xi)$ in a locally Lipschitz continuous way, and this implies the uniqueness of the solution.

Theorem 5.4. *Suppose $\alpha \in (2/3, 1)$, then for any given set of data $(u_0, \xi, Z, Z \circ \xi) \in \mathcal{C}^{\alpha} \times \mathcal{C}^{\alpha-2} \times \mathcal{C}^{\alpha} \times \mathcal{C}^{2\alpha-2}$, there is a finite time $T > 0$, such that a unique solution $(u \prec Z + u^{\#}, u, u^{\#}) \in \mathcal{D}_T^{\alpha}(Z)$ of the system of equations*

$$\begin{aligned} u &= u \prec Z + u^{\#} \\ u^{\#} &= \int_0^t P_{t-s} (u(Z \circ \xi) + \xi \prec u - [L, u \prec] Z + C(u, Z, \xi) + u^{\#} \circ \xi) ds \end{aligned}$$

with initial condition $u(0) = u^{\#}(0) = u_0$ *exists. Moreover, the solution depends on* $(u_0, \xi, Z, Z \circ \xi) \in \mathcal{C}^{\alpha} \times \mathcal{C}^{\alpha-2} \times \mathcal{C}^{\alpha} \times \mathcal{C}^{2\alpha-2}$ *in a locally Lipschitz continuous way.*

Finally we say a few words on renormalization. We should notice that here $Z \circ \xi$ is just a notation, not a legal computation, and we thought it as an independent element from Z and ξ. To construct this term, one may consider a sequence of smooth approximations ξ^{ε} and Z^{ε} (convolution with some approximation of identity). Instead of the convergence of $Z^{\varepsilon} \circ \xi^{\varepsilon}$, one find there is a sequence of numbers $c_{\varepsilon}(t)$ which tends to infinity as $\varepsilon \to 0$, such that

$$Z^{\varepsilon}(t) \circ \xi^{\varepsilon} - c^{\varepsilon}(t)$$

converges in a suitable norm (expectation of some Besov norm). For each ε we can solve previous system of equation, then take limit $\varepsilon \to 0$ and use the continuity with respect to the input data. So we actually solved the renormalized PAM

$$\partial_t u = \Delta u + \xi u - \infty u$$

note that for each sample point in the probability space, the time interval for the solution may be different, so the renormalized PAM exists within a random time interval.

6 Axiomatic Quantum Field Theory

In this chapter, we introduce the subject called axiomatic quantum field theory, which try to extract basic properties that a relativistic quantum field theory should satisfy. This is given by the Wightman axioms, and we show that one can use correlation functions to reconstruct Wightman quantum fields. Then we introduce how this is linked to the Schwinger functions in Euclidean quantum field theory, and the axioms given by Osterwalder-Schrader.

6.1 Quantum Fields as Operator-Valued Distributions

In classical physics, a field is a function of space-time, which means an assignment of quantities of particular type (for example scalar, vector, tensor or spinor, etc.) to each point of space time. Such a concept allow us to know the observational value of fields at each point with an infinite accuracy.

In quantum theory, an observable is a self-adjoint operator on some separable Hilbert space, here we adopt von Neumann's separable Hilbert space formulation of quantum mechanics. It seems natural to think of a quantum field as a operator-valued function of space-time, that is to each point of space-time, we assign a self-adjoint operator on some fixed separable Hilbert space, which represents an observable of some field components of a particular type. But this is not correct.

To see this from a mathematical viewpoint, let us see the case of free scalar Boson field, the corresponding Hilbert space is the Fock space, namely

$$\mathbb{C}\oplus L^2(\mathbb{R}^3)\oplus(L^2(\mathbb{R}^3)\otimes L^2(\mathbb{R}^3))\oplus\ldots=\mathcal{H}_0\oplus\mathcal{H}_1\oplus\mathcal{H}_2\oplus\cdots$$

where $\mathcal{H}_n=\otimes^n L^2(\mathbb{R}^3)$ and a typical vector has form $(\Psi_0,\Psi_1,\Psi_2,\cdots)$ with $\Psi_0\in\mathbb{C}$, $\Psi_1\in L^2(\mathbb{R}^3)$, $\Psi_2\in L^2(\mathbb{R}^3)\otimes L^2(\mathbb{R}^3),\ldots$ represents vacuum wave function, one particle wave function, two particle wave function,etc. The scalar product is given by

$$\langle\Phi,\Psi\rangle=\bar{\Phi}_0\,\Psi_0+\int_{\mathbb{R}^3}\bar{\Phi}_1(x)\,\Psi_1(x)\,d^3x+\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\bar{\Phi}_2(x_1,x_2)\,\Psi_2(x_1,x_2)d^3x_1d^3x_2+\cdots$$

if we define annihilation operator $\psi(x)$ at each point of space by

$$(\psi(x)\,\Psi)_n(x_1,\cdots,x_n)=\sqrt{n+1}\,\Psi_{n+1}(x,x_1,\cdots,x_n)$$

then define and compute the creation operator $\psi^*(x)$ at the same point as the adjoint operator of annihilation operator, we will find that the domain of the adjoint only contains 0 vector, that is the field is too singular such that the operator at any point can not act on any nonzero vector in the Hilbert space (see the arguments in [48], or chapter 12 of the book [9]). Moreover, Wightman showed that in general, one can not define the quantum fields at any space time point [47].

Theorem 6.1. *Suppose we have a quantum theory with a separable Hilbert space $\mathcal{H}$, and a strong continuous unitary representation of space-time translation $\mathbb{R}^{1,3} \ni a \mapsto U(a) \in \mathcal{U}(\mathcal{H})$ such that the spectrum of the energy momentum operator is contained in the closed forward light cone. Suppose the quantum theory has a unique vacuum vector Ψ_0, which is invariant under the action of space-time translation, that is*

$$U(a)\Psi_0 = \Psi_0 \,\text{for all}\, a \in \mathbb{R}^{1,3}$$

Then a map B from a bounded open set $\mathcal{O} \subset \mathbb{R}^{1,3}$ to Von Neumann algebra of bounded operators on $\mathcal{H}$, with the following properties

$$U(a)B(x)U(-a) = B(x+a)$$

$$[B(x), B(y)^{(*)}] = 0$$

where a is small enough and $(x-y)$ is a space-like vector. Then B has constant value equal to a constant multiple of identity.

See theorem 3.1 in [47], where the possibility of generalizing to unbounded operators is also discussed.

From physical viewpoint, this is a consequence of uncertainty principle for fields as given by Bohr and Rosenfeld [11]. From Bohr and Rosenfeld's analysis on measurability of electromagnetic fields in the theory of quantum electrodynamics (see [11] and [12], the English translation can be found in [46]), only the quantities formally corresponds to the average of its classical analog over finite space-time regions are measurable, and hence are observables, namely things like

$$\frac{1}{|O|}\int_O F_{\mu\nu}(x)\, d^4 x$$

where O is an open set in $\mathbb{R}^{1,3}$, $|O|$ represents its volume, and $F_{\mu\nu}$ is the electromagnetic tensor.

It was Heisenberg who first used smeared fields as fundamental object, since by using smeared fields, he could avoid the infinite fluctuation in the computation of Einstein's fluctuation formula of blackbody radiation (see [26] and discussions in [48]). He also argued that in general, to measure the field in a sharply defined region, which is a mathematical idealization, one has to use an infinite amount of energy, thus one can only measure the smeared field. The case of electromagnetic field is special in the analysis of Bohr and Rosenfeld [11]. But there is still a question that what differentiability and regularity conditions one should assume in order to define the smeared field. Heisenberg used second differentiable function to define the smeared field, but the free electromagnetic field need not be smeared due to the analysis of Bohr and Rosenfeld.

Inspired by the Laurent Schwartz's theory of distribution, Wightman and Gårding [49] first try to use the Schwartz function space as the test function to define the smeared field, which is the following definition. In this chapter, we use the complex-valued Schwartz function space, namely the real and imaginary parts are both Schwartz function, this space is also denoted by $\mathcal{S}$.

Definition 6.2. (Operator-valued distribution) *Suppose $\mathcal{H}$ is a Hilbert space, an operator-valued distribution is a complex linear map φ from complex-valued Schwartz function space $\mathcal{S}$ to the set of operators (bounded or unbounded) on Hilbert space $\mathcal{H}$, such that all the operators $\varphi(f), \forall f \in \mathcal{S}$ have a common dense domain D, and the map*

$$\mathcal{S} \to \mathbb{C}, f \mapsto \langle \Phi, \varphi(f)\, \Psi \rangle$$

is continuous, where $\Phi \in \mathcal{H}, \Psi \in D$ are fixed vectors.

We should remark that there are different choice of test function space, not all quantum field theories can be described by just using Schwartz function space as the test function (see the discussion on page 804 [43], and [29]), more restrictive test functions are needed for more singular behavior of vacuum expectations, and there is no single choice for all the quantum field theory models. We also remark that the requirement of the map φ to be weak continuous and together with choosing Schwartz function as test functions in the Wightman axioms will lead to only renormalizable models in quantum field theory, see chapter 15 of [9].

6.2 Wightman Axioms of Relativistic Quantum Fields

Since we need to discuss the quantum theory of fields in Minkowski space-time, we first set up the notations and tools in special relativity. For convenience, we assume the speed of light c is 1. Denote $\mathbb{R}^{1,d}$ the $1+d-$dimensional Minkowski space, with the scalar product of two typical vectors $x_i = (x_i^\mu) = (x_i^0, x_i^1, \ldots, x_i^d) \in \mathbb{R}^{1,d}$, $i = 1, 2$ and $\mu = 0, 1, \ldots, d$, given by

$$\langle x_1, x_2 \rangle = x_1^0 x_2^0 - x_1^1 x_2^1 - \cdots - x_1^d x_2^d = g_{\mu\nu} x_1^\mu x_2^\nu$$

where $g_{00} = 1, g_{11} = \ldots = g_{dd} = -1$ and $g_{\mu\nu} = 0$ if $\mu \neq \nu$. A vector $x \in \mathbb{R}^{1,d}$ is called

$$\begin{cases} \text{time-like}, & \text{if} \quad \langle x, x \rangle > 0 \\ \text{space-like}, & \text{if} \quad \langle x, x \rangle < 0 \\ \text{light-like}, & \text{if} \quad \langle x, x \rangle = 0 \end{cases}$$

The set of all time-like vectors whose first component is positive is an open set, which is called the forward light cone, denoted by V_+, this is the causal future of the origin. Clearly the closure $\overline{V_+}$ is the set of all time-like and light-like vectors whose first component is non-negative.

Consider $d = 3$, which is the case of our physical world. The set of all linear transformations Λ which preserves the Minkowski metric

$$\langle \Lambda\, x, \Lambda\, y \rangle = \langle x, y \rangle \Leftrightarrow g_{\alpha\beta} \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu = g_{\mu\nu}$$

forms a Lie group, called the extended Lorentz group, denoted by $O(1,3)$. Two identities

$$\det(\Lambda)^2 = 1, (\Lambda^0{}_0)^2 - (\Lambda^1{}_0)^2 - (\Lambda^2{}_0)^2 - (\Lambda^3{}_0)^2 = 1$$

can be obtained directly from the definition. Thus, the Lorentz group has four connected components, classified by four different range of conditions $\det(\Lambda) = \pm 1$ and $|\Lambda^0{}_0| \geqslant 1$ (see page 10 of [44]). The connected component of the identity is a subgroup, called the restricted Lorentz group, denoted by $SO^+(1,3)$, which preserves the orientation of the whole Minkowski space ($\det(\Lambda) = 1$) and direction of time ($\Lambda^0{}_0 \geqslant 1$). The other three components can be obtained by multiplying

$$T = \begin{bmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, P = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}, PT = \begin{bmatrix} -1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}$$

which are time reversal operator, space inversion operator and their product.

The restricted Lorentz group is a six dimensional group, six independent symmetries are given by the rotation $\{R_x^\theta, R_y^\theta, R_z^\theta\}$ about three spacial axis

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta & -\sin\theta \\ & & \sin\theta & \cos\theta \end{bmatrix}, \begin{bmatrix} 1 & & & \\ & \cos\theta & & \sin\theta \\ & & 1 & \\ & -\sin\theta & & \cos\theta \end{bmatrix}, \begin{bmatrix} 1 & & & \\ & \cos\theta & -\sin\theta & \\ & \sin\theta & \cos\theta & \\ & & & 1 \end{bmatrix}$$

and the boosts $\{M_x^\phi, M_y^\phi, M_z^\phi\}$ about three axis

$$\begin{bmatrix} \cosh\phi & -\sinh\phi & & \\ -\sinh\phi & \cosh\phi & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \begin{bmatrix} \cosh\phi & & -\sinh\phi & \\ & 1 & & \\ -\sinh\phi & & \cosh\phi & \\ & & & 1 \end{bmatrix}, \begin{bmatrix} \cosh\phi & & & -\sinh\phi \\ & 1 & & \\ & & 1 & \\ -\sinh\phi & & & \cosh\phi \end{bmatrix}.$$

Together with the translation symmetry given by vectors in $\mathbb{R}^{1,3}$, we can define the Poincaré group P to be the set

$$\{(a, \Lambda) |\, \Lambda \in O(1,3), a \in \mathbb{R}^{1,3}\}$$

with group law

$$(a_1, \Lambda_1)(a_2, \Lambda_2) = (a_1 + \Lambda_1 a_2, \Lambda_1 \Lambda_2).$$

Clearly an element (a, Λ) of Poincaré group P should act on any $x \in \mathbb{R}^{1,3}$ by

$$(a, \Lambda)v = \Lambda x + a.$$

The Poincaré group also has four connected components, we call the connected component of the identity is called the restricted Poincaré group, denoted by $\mathbb{R}^{1,3} \rtimes SO^+(1,3)$ according to its construction.

The Lie algebra of the Poincaré group is called the Poincaré algebra, which is generated by generators of translations in four space-time directions $\{p^0, p^1, p^2, p^3\}$ and generators of the rotation $\{L_1, L_2, L_3\}$ about three spacial axis

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and the boosts $\{M_1, M_2, M_3\}$ about three axis

$$\begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

The restricted Lorentz group $SO^+(1,3)$ has fundamental group $\mathbb{Z}_2$, its covering group is $SL(2,\mathbb{C})$, which is called the inhomogeneous Lorentz group. This is given by following construction, there is an isomorphism of vector space between $\mathbb{R}^{1,3}$ and the set of 2-by-2 Hermitian matrix, given by

$$(x^0, x^1, x^2, x^3) \leftrightarrow \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{bmatrix}$$

then clearly for any $A \in SL(2,\mathbb{C})$, the matrix

$$A \begin{bmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{bmatrix} A^*$$

is still a Hermitian matrix, which induce a linear representation of $SL(2,\mathbb{C})$ on $\mathbb{R}^{1,3}$. Moreover, we see that

$$\begin{aligned} \det\left(A \begin{bmatrix} t+x & y-iz \\ x+iy & t-x \end{bmatrix} A^* \right) =& \det \begin{bmatrix} t+x & y-iz \\ y+iz & t-x \end{bmatrix} \\ =& t^2 - x^2 - y^2 - z^2 \end{aligned}$$

which means A acts as a Lorentz transformation. Since the group $SL(2,\mathbb{C})$ is simply connected, this induce a group homomorphism from $SL(2,\mathbb{C})$ to $SO^+(1,3)$ (for surjectivity, see page 134 of [9]). Two elements $A, B \in SL(2,\mathbb{C})$ has the same image if and only if $B = -A$.

The double cover of the restricted Poincaré group P is then given by the

$$\{(a, A) |\, A \in SL(2,\mathbb{C}), a \in \mathbb{R}^{1,3}\}$$

with group law

$$(a_1, A_1)(a_2, A_2) = (a_1 + A_1 a_2, A_1 A_2)$$

where the action on the Minkowski space vector is given by the representation described above, this group is denoted by $\mathbb{R}^{1,3} \rtimes SL(2,\mathbb{C})$ according to its construction.

In the following discussion of quantum mechanics, we use the language of Hilbert space in the spirit of von Neumann, where states are represented by unit rays $\{\Psi\}$ in a Hilbert space $\mathcal{H}$, that is Ψ and $c\Psi$ represent the same state, for all $cc^* = 1$. It turns out that we can not always construct the superposition of two states in the quantum theory, for example the charge is conserved in nature, we don't see a state of a system which has nonzero probabilities to be observed with different charge number. Such a property of non-existence of certain superpositions is called a superselection rule, the maximal subspace of Hilbert space where superposition principle holds is called a superselection sector. See chapter 1 of [44] for more discussion on this issue. In the following discussion, we restrict ourself in a superselection sector.

A symmetry is a transformation of viewpoint which does not produce any observational physical effect, such transformations are divided into two classes, ones which change the mathematical labelings and descriptions are called passive, and ones which do change the status of the experimental apparatus are called active. Lorentz transformations are both active and passive, gauge symmetries are only passive, not active. In quantum mechanics, the only observational quantity is the probability of a prepared normalized state Ψ observed in given normalized state Φ, which is given by Born's rule $|\langle\Psi,\Phi\rangle|^2$. Thus a symmetry U any normalized state Ψ into a new state $U\Psi$ in the same Hilbert space, such that

$$|\langle U\Psi, U\Phi\rangle|^2 = |\langle\Psi,\Phi\rangle|^2$$

clearly this condition is independent of the representatives chosen in each unit ray.

Wigner showed such symmetries are either unitary or anti-unitary operators, see [51] and [8]. A transformation U is called anti-unitary if the following condition holds:

$$U(a\,\Psi + b\,\Phi) = a^*\,U(\Psi) + b^*\,U(\Phi) \text{ for } \forall a, b \in \mathbb{C}, \Psi, \Phi \in \mathcal{H};$$

$$\langle U\,\Psi, U\,\Phi\rangle = \langle\Phi,\Psi\rangle \text{ for } \forall\Psi,\Phi \in \mathcal{H}.$$

Clearly the product of two anti-unitary operators is a unitary operator.

In special relativity, we require the physical laws are invariant under Poincaré transformations. It turns out that in particle physics, nature is not invariant under time reversal, parity and their product, only restricted Lorentz transformations are symmetries of nature. Thus each element $\Lambda \in \mathbb{R}^{1,3} \rtimes SO^+(1,3)$ induces a symmetry $U(\Lambda)$. Since every element in the vicinity of identity of Lie group $\mathbb{R}^{1,3} \rtimes SO^+(1,3)$ is a square of some other element, and any element in $\mathbb{R}^{1,3} \rtimes SO^+(1,3)$ is a product of finite number of elements in the vicinity of identity, the symmetry $U(\Lambda)$ is actually a unitary operator. Clearly

$$U(\Lambda_1)U(\Lambda_2) = U(\Lambda_1\Lambda_2)$$

on the space of unit rays, which is equivalent to say

$$U(\Lambda_1)U(\Lambda_2) = e^{if(\Lambda_1,\Lambda_2)}U(\Lambda_1\Lambda_2) \tag{6.1}$$

where $f\colon (\mathbb{R}^{1,3} \rtimes SO^+(1,3)) \times (\mathbb{R}^{1,3} \rtimes SO^+(1,3)) \to \mathbb{R}$ is a function. A map from $\mathbb{R}^{1,3} \rtimes SO^+(1,3)$ to the group $U(\mathcal{H})$ of unitary operators on a Hilbert space is a called a projective representation if the identity element is mapped to the identity operator with a phase factor and 6.1 holds. The name comes from the fact that if we consider the quotient group $U(\mathcal{H})/\{cI|c\in\mathbb{C}, |c|=1\}$, then the composition of the projective representation and the quotient map

$$\mathbb{R}^{1,3} \rtimes SO^+(1,3) \to U(\mathcal{H}) \to U(\mathcal{H})/\{cI|c\in\mathbb{C}, |c|=1\}$$

is then a representation of the Poincaré group. Cleary if we multiply a projective representation by a phase factor which depends on the elements in the Poincaré group is again a projective representation, and they correspond to same representation into $U(\mathcal{H})/\{cI|c\in\mathbb{C}, |c|=1\}$. Moreover there is a one-to-one correspondence between projective representations into $U(\mathcal{H})$ up to a phase function and representations into $U(\mathcal{H})/\{cI|c\in\mathbb{C}, |c|=1\}$. We have the following theorem by Bargmann [7] and Wigner [50].

Theorem 6.3. (Wigner 1939, Bargmann 1954) *Any projective representation of restricted Poincaré group $\mathbb{R}^{1,3} \rtimes SO^+(1,3)$ to the group $U(\mathcal{H})$ of unitary operators on a Hilbert space is induced by a unitary representation of $\mathbb{R}^{1,3} \rtimes SL(2,\mathbb{C})$, this is given by sending a pre-image of an element in $\mathbb{R}^{1,3} \rtimes SO^+(1,3)$ under double cover, to the equivalence class of the unitary operator of representation of this element.*

Thus one has to study the infinite dimensional strong continuous unitary representation of the group $\mathbb{R}^{1,3} \rtimes SL(2,\mathbb{C})$. This was done by Wigner, see [50] and [6]. The original motivation of Wigner's classification was to under stand the one-particle state in quantum theory. Wigner though that for any quantum field theory, with or without interaction, there must exist a subspace which describe the state where there is only one single particle, and this subspace should look the same for each observer, that is invariant under the action of Poincaré group. Wigner's idea was that the irreducible representations of Poincaré group can be used to classify the types of the particles. The representation of $\mathbb{R}^{1,3} \rtimes SL(2,\mathbb{C})$ induces a representation of Poincaré algebra, denote the image of each generator, one get the energy-momentum operator $\{P^\mu\}$ and the generator of rotation and boosts $\{J^{\nu\rho} = -J^{\rho\nu}\}$ such that $e^{-ia_\mu P^\mu}$, $e^{i\theta_{\nu\rho}J^{\nu\rho}}$ are corresponding unitary operators. It can be shown that the following two operators commute with all elements in the image of representation of Poincaré algebra

$$P_\mu P^\mu = (P^0)^2 - (P^1)^2 - (P^2)^2 - (P^3)^2$$

$$W_\mu W^\mu$$

where $W_\lambda = \frac{1}{2}\epsilon_{\mu\nu\rho\lambda} M_{\mu\nu} P^\rho$ called the Pauli-Lubanski operator, $\epsilon_{\mu\nu\rho\lambda}$ is the Levi-Civita symbol. Since the representation is irreducible, then we know that they are constant multiple of identity operator. Moreover, these constants can be used to classify the representation

$$P_\mu P^\mu = m^2 \mathrm{Id}, W_\mu W^\mu = m^2 s(s+1)$$

where m is interpreted to be the mass, and when m is positive, s takes values in non-negative half integers, is interpreted as the spin of the particle. We only consider the case of $m>0$. In this case, Wigner showed there is a one-to-one correspondence between finite dimensional irreducible representations $D^{(s,0)}$ of $SL(2,\mathbb{C})$ and massive $(m\neq 0)$ irreducible continuous unitary representation of $\mathbb{R}^{1,3}\rtimes SL(2,\mathbb{C})$. We first briefly talk about the finite dimensional irreducible representation of $SL(2,\mathbb{C})$ (in complex vector space), more details can be found in [13].

The study of the finite dimensional irreducible representation of the spin group can be reduced to the study of its Lie algebra, for convenience we can study the complexification of this Lie algebra. We redefine the generators in the previous discussion $\{L_1, L_2, L_3, M_1, M_2, M_3\}$ by $\{L_k = i\,L_k, M_k = i\,M_k, k=1,2,3\}$, then the commutator relations of these generators are given by

$$[L_i, L_j] = i\,\epsilon_{ijk}\,L_k, [L_i, M_j] = i\,\epsilon_{ijk}\,M_k, [M_i, M_j] = -i\,\epsilon_{ijk}\,L_k$$

where ϵ_{ijk} is the Levi-Civita symbol. Define $J_\cdot^{(\pm)} = \frac{1}{2}(L_\cdot \pm i\,M_\cdot)$, then one can show that

$$[J_i^{(+)}, J_j^{(+)}] = i\,\epsilon_{ijk}\,J_k^{(+)}, [J_i^{(-)}, J_j^{(-)}] = i\,\epsilon_{ijk}\,J_k^{(-)}, [J_i^{(+)}, J_j^{(-)}] = 0$$

thus the Lie algebra can be decomposed into a direct sum of two copies of rotation Lie algebra. One can show that the complete list of inequivalent irreducible representation of Lorentz group can be labeled by two half integers $D^{(s_+,s_-)}$ where $s_\pm = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ with property $[D^{(s_+,s_-)}]^*$ and $D^{(s_-,s_+)}$ are equivalent representations.

As proved by Wigner, for $m>0$ and a half integer s, the unique irreducible continuous unitary representation of $\mathbb{R}^{1,3}\rtimes SL(2,\mathbb{C})$ can be constructed in the following way. Consider in momentum 4-space the mass shell $\Gamma_m^+ := \{p\in\mathbb{R}^{1,3} | \, p^2 = m^2, x_0 > 0\}$ and the Lorentz invariant measure $d\Omega_m(p) = \frac{dp_1 dp_2 dp_3}{\sqrt{m^2 + p_1^2 + p_2^2 + p_3^2}}$ on it. Consider the space $\bigoplus^{2s+1} L^2(\Gamma_m^+)$ of functions the form $\Psi(p,\sigma)$ where $p\in\Gamma_m^+$ and $\sigma = -s, \ldots, s$, with inner product

$$(\Phi, \Psi) = \sum_{\sigma=-s}^{s} \int_{\Gamma_m^+} \overline{\Phi(p,\sigma)}\,\Psi(p,\sigma) d\Omega_m(p)$$

and the action of $(a, A)\in\mathbb{R}^{1,3}\rtimes SL(2,\mathbb{C})$ given by

$$((a, A)\Psi)(p,\sigma) = e^{-ia\cdot p}\sum_{\sigma'} D_{\sigma\sigma'}^{(s)}(R(p,A))\Psi(A^{-1}p, \sigma')$$

where $R(p,A) = W(A^{-1}p, A) = L(p)^{-1}AL(A^{-1}p)\in SU(2)$, which is the little group of vector $(m,0,0,0)$, $L(p)$ is the Lorentz transformation satisfies $p = L(p)(m,0,0,0)$, and $D^{(s)}$ is an irreducible representation of $SU(2)$, see chapter 2 of [45]. Under the Mackey's theory of induced representation, see [42], this representation can be extended to the representation

$$((a, A)\Psi)(p,\sigma) = \sum_{\sigma'} e^{-ia\cdot p} D_{\sigma\sigma'}^{(s)}(A)\Psi(A^{-1}\,p, \sigma')$$

where $D^{(s)}_{\sigma'\sigma}$ is the irreducible representation equivalent to $D^{(s,0)}$, as an extension of irreducible representation of $SU(2)$. The covariance in the Wightman's axiom is motivated by this transformation law.

Now we are in the position to introduce the Wightman axioms for the relativistic quantum fields.

1. Space of states

- States are represented as unit rays in a separable complex Hilbert space $\mathcal{H}$.

- There is a strong continuous unitary representation of the group $\mathbb{R}^{1,3} \rtimes SL(2, \mathbb{C})$.

- (**uniqueness of vacuum**) There is a unique unit ray $\{\Omega\}$ (interpreted as vacuum) such that

$$U(a, A)\Omega = \Omega$$

for any $(a, A) \in \mathbb{R}^{1,3} \rtimes SL(2, \mathbb{C})$.

- (**spectrum condition**) The generators of space-time translations (P^0, P^1, P^2, P^3), interpreted as the energy-momentum operator, has spectrum in closed forward light cone $\overline{V_+}$.

2. Observables and covariance

- A set of operator valued distributions $\{\varphi^{(k)}_n | k, n \in \mathbb{N}\}$, where k labels the type of the field which can be at most countable and n labels the components of the field which can only take finite number of values, and a dense subspace D where all the operators $\varphi^{(k)}_n(f)$ and $\varphi^{(k)*}_n(f) = \varphi^{(k)}_n(\bar{f})^*$ are defined, for all $n \in \mathbb{N}$ and $f \in \mathcal{S}(\mathbb{R}^4)$.

- The vacuum Ω is contained in D.

- The domain D is invariant under the action of $U(A, a)$, $\varphi^{(k)}_n(f)$ and $\varphi^{(k)*}_n(f)$, for all $(a, A) \in \mathbb{R}^{1,3} \rtimes SL(2, \mathbb{C})$, $n \in \mathbb{N}$ and $f \in \mathcal{S}(\mathbb{R}^4)$.

- The covariant transformation of fields operator under the action of (a, A) is given by

$$U(a, A)\varphi^{(k)(*)}_n(f)U(a, A)^{-1} = \sum D^{(k)}_{nm}(A^{-1})\varphi^{(k)(*)}_m((a, A)f)$$

where $D^{(k)}_{nm}(A)$ are matrices of a finite dimensional irreducible representation of the group $SL(2, \mathbb{C})$ with $\varphi^{(k)}_n$ as its components, and $(a, A)f = f(A^{-1}(x - a))$. If the representation $D(A)$ is a representation of group $\mathrm{SO}^+(1, 3)$, then the fields are called a tensor field, otherwise the fields are called spinor fields. This transformation law is linear in the test function.

- The vacuum Ω is a cyclic vector, which means the linear span D_0 of the set $\{\varphi^{(k_1)(*)}_{i_1}(f_1) \ldots \varphi^{(k_m)(*)}_{i_m}(f_m)\,\Omega |\, m \in \mathbb{N}, i_1, \cdots, i_m \in \mathbb{N}, f_1, \ldots, f_m \in \mathcal{S}(\mathbb{R}^4)\}$ is dense.

3. Locality or Microcausality

- For any two test functions $f, g \in \mathbb{R}^4$ whose supports consists only space-like separated points, the operators $\varphi^{(k)(*)}_n(f)$ and $\varphi^{(k')(*)}_m(g)$ satisfies

$$\varphi^{(k)(*)}_n(f)\varphi^{(k')(*)}_m(g) - \sigma(k, k')\varphi^{(k')(*)}_m(g)\varphi^{(k)(*)}_n(f) = 0$$

where $\sigma(k, k') = 1$ if one of k and k' is representation with integer spin, and $\sigma(k, k') = -1$ if both k and k' are representation of odd spin.

We say that a field theory is a Hermitian scalar boson field, if real test functions are mapped to the symmetric operators, the representation of the Lorentz group is taken to be the trivial representation, and the sign in the microcausality is taken to be 1. We make a few comments on these axioms.

Remark 6.4.

1. The original formulation of locality in [44], the number $\sigma(k,k')$ is not assumed to be a constant depends only on the type of the fields, but afterwards, they proved the famous spin-statistics theorem which says one has to choose the sign for components of fields in a irreducible representation in the way we assumed, and for different types of fields, one can do the Klein transformation to make the sign agree with our choice, so there is no loss of generality. See page 328 of [10] for more discussion on this issue.

2. The dense domain D in the assumption of observables, is a technical assumption. But for the symmetric operators defined on a dense domain, it is a well-known fact that there may not be a unique self-adjoint extension. Thus this axiom may produce difficulties when constructing examples of Wightman fields.

3. It is clear that we can formulate the axioms for any space dimension d, then one has to study the representation theory of the group $\mathrm{SO}^+(1,d)$, in order to formulate the correct transformation laws and commutator relations. For most literature on the subject of axiomatic field theories, people usually treat the case of single hermitian scalar field, which avoids these difficulties.

4. For the gauge theory, this set of axioms has to be modified. For the free electro-magnetic field, there is a negative result by Ferrari, Picasso and Strocchi [18], which says the covariant theory for the four-vector potential as operator-valued distributions, whose curl is also covariant, interpreted as the electromagnetic field, and and satisfies the classical free Maxwell equations, does not exists. One can not just consider the electromagnetic tensor as covariant observables as in classical electrodynamics, where the four potential is regarded as a mathematical trick. In quantum theory, the four potential has physical effect, for example the famous Aharonov-Bohm effect [2], where $\exp\left(ie\oint_C \mathbf{A}(x)d\mathbf{x}\right)$ is an observable, but this includes less information of the gauge field $\mathbf{A}(x)$. For an analogue in the case of non-abelian gauge theory, see Yang and Wu [52]. These effects has been observed in experiments. Glimm and Lee proposed a possible system of axioms for the quantum gauge theory, see [22].

5. There is another approach called rigged Hilbert space approach to quantum theory different from the Hilbert space theory by von Neumann. The motivation of this approach is to make the Dirac's formalism rigorous, and the crucial feature is one can have non-normalizable eigenstate. For this approach to quantum mechanics, see [40]. Bogoliubov and his collaborators modified the Wightman axiom into a system using this rigged Hilbert space approach in [9], but they did not use it in [10]. Prigogine and Antoniou argued that this is the suitable frame work to describe the irreversible system in quantum theory [3].

6. We assumed that the Hilbert space of states are separable, but the constructions in physics usually result in a non-separable Hilbert space, for example if you take tensor product an infinite number of Hilbert space, then you will get a non-separable Hilbert space, but it can be shown that only a separable subspace is meaningful to represents the physical states and define the operators.

The square of the energy-momentum operator

$$(P^0)^2-(P^1)^2-(P^2)^2-(P^3)^2$$

is called the mass operator. A quantum field theory is said to has a **mass gap**, if there is a positive number Δ, such that there is no eigenvalue between 0 and Δ in the spectrum of this operator.

The correlation of Wightman fields, given by

$$W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}(f_1,\cdots,f_n):=\langle\Psi_0,\varphi_{m_1}^{(k_1)(*)}(f_1)\cdots\varphi_{m_n}^{(k_n)(*)}(f_n)\,\Psi_0\rangle$$

for any $f_1,\cdots,f_n\in\mathcal{S}(\mathbb{R}^4)$ is clearly a continuous multilinear functional, and thus by nuclear theorem, this defines a tempered distribution on $\mathbb{R}^{4n}$. The collection $\left\{W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}\right\}$ are called the Wightman distributions.

Proposition 6.5. (Hermiticity) *We have for all test functions*

$$W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}(f_1,\cdots,f_n)=\overline{W_{m_n,\cdots,m_1}^{k_n^{-(*)},\cdots,k_1^{-(*)}}(f_n,\cdots,f_1)}$$

where the notation $-(*)$ means if we have index k_i, then we take $k_i^{-(*)}=k_i^*$, if we have k_i^*, we then take $k_i^{-(*)}=k_i$.

Proof. Clearly we have

$$\begin{aligned}
& W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}(f_1,\cdots,f_n)\\
=\ & \langle\Psi_0,\varphi_{m_1}^{(k_1)(*)}(f_1)\cdots\varphi_{m_n}^{(k_n)(*)}(f_n)\,\Psi_0\rangle\\
=\ & \langle\varphi_{m_n}^{(k_n)(-*)}(f_n)\cdots\varphi_{m_1}^{(k_1)(-*)}(f_1)\Psi_0,\Psi_0\rangle\\
=\ & \overline{\langle\Psi_0,\varphi_{m_1}^{(k_1)(-*)}(f_1)\cdots\varphi_{m_n}^{(k_n)(-*)}(f_n)\Psi_0\rangle}\\
=\ & \overline{W_{m_n,\cdots,m_1}^{k_n^{-(*)},\cdots,k_1^{-(*)}}(f_n,\cdots,f_a)}
\end{aligned}$$

which gives the result. □

Proposition 6.6. (Positivity) *For any finite sequence of test functions* $\left\{f_{m_1,\cdots,m_i}^{k_1^{(*)},\cdots,k_i^{(*)}}\,\middle|\,f_{m_1,\cdots,m_i}^{k_1^{(*)},\cdots,k_i^{(*)}}\in\mathcal{S}(\mathbb{R}^{4i})\right\}$, *we have*

$$\sum_{i,j}\sum_{\substack{k_1'^{(*)},\cdots,k_j'^{(*)}\\ m_1',\cdots,m_j'}}\sum_{\substack{k_1^{(*)},\cdots,k_i^{(*)}\\ m_1,\cdots,m_i}} W_{m_j',\cdots,m_1',m_1,\cdots,m_i}^{k_j'^{-(*)},\cdots,k_1'^{-(*)},k_1^{(*)},\cdots,k_i^{(*)}}\left(\overline{\check{f}_{m_1',\cdots,m_j'}^{k_1'^{(*)},\cdots,k_j'^{(*)}}}\,f_{m_1,\cdots,m_i}^{k_1^{(*)},\cdots,k_i^{(*)}}\right)\geqslant 0$$

where $\check{f}^{k_1'^{(*)},\cdots,k_j'^{(*)}}_{m_1',\cdots,m_j'}$ *is the function* $f^{k_1'^{(*)},\cdots,k_j'^{(*)}}_{m_1',\cdots,m_j'}(x_j,x_{j-1},\ldots,x_1)$.

Proof. If all test function $f^{k_1^{(*)},\cdots,k_i^{(*)}}_{m_1,\cdots,m_i}$ as the form $f^{k_1^{(*)}}_{m_1}\cdots f^{k_i^{(*)}}_{m_i}$ .This is just from the fact that the norm of the vector

$$\sum_i \sum_{\substack{k_1^{(*)},\cdots,k_i^{(*)}\\ m_1,\cdots,m_i}} \varphi^{(k_1)(*)}_{m_1}(f_1)\cdots\varphi^{(k_i)(*)}_{m_i}(f_i)\Psi_0$$

is non-negative. And note that any test function $f^{k_1^{(*)},\cdots,k_i^{(*)}}_{m_1,\cdots,m_i}$can be approximated by a sequence of the form

$$\sum_l f^{k_1^{(*)}}_{m_1,l}\cdots f^{k_i^{(*)}}_{m_i,l}$$

where the convergence is in the space if Schwartz function. The statement is clearly true for such elements in the approximating sequence. □

Proposition 6.7. (Covariance) *We have*

$$\sum_{n_1,\ldots,n_l} D^{(k_1)}_{m_1n_1}(A^{-1})\ldots D^{(k_l)}_{m_ln_l}(A^{-1})W^{k_1^{(*)},\cdots,k_l^{(*)}}_{n_1,\cdots,n_l}((a,A)f_1,\cdots,(a,A)f_l)$$
$$=W^{k_1^{(*)},\cdots,k_l^{(*)}}_{m_1,\cdots,m_l}(f_1,\cdots,f_l)$$

for any $(a,A)\in\mathbb{R}^{1,3}\rtimes SL(2,\mathbb{C})$, $l,k_1,\ldots,k_l\in\mathbb{N}$ *and* $0\leqslant m_i\leqslant 2s(k_i)+1, 1\leqslant i\leqslant l$.

Proof. We use covariance property in the axioms of Wightman fields

$$\begin{aligned}
& W^{k_1^{(*)},\cdots,k_l^{(*)}}_{m_1,\cdots,m_l}(f_1,\cdots,f_l)\\
=\ & \langle\Psi_0,\varphi^{(k_1)(*)}_{m_1}(f_1)\cdots\varphi^{(k_l)(*)}_{m_l}(f_l)\,\Psi_0\rangle\\
=\ & \langle U(a,A)^{-1}\Psi_0,\varphi^{(k_1)(*)}_{m_1}(f_1)\cdots\varphi^{(k_l)(*)}_{m_l}(f_l)\,U(a,A)^{-1}\Psi_0\rangle\\
=\ & \langle\Psi_0,U(a,A)\varphi^{(k_1)(*)}_{m_1}(f_1)U(a,A)^{-1}U(a,A)\cdots U(a,A)\,\varphi^{(k_l)(*)}_{m_l}(f_l)\,U(a,A)^{-1}\Psi_0\rangle\\
=\ & \left\langle\Psi_0,\sum_{n_1}D^{(k_1)}_{m_1n_1}(A^{-1})\varphi^{(k_1)(*)}_{n_1}((a,A)f_1)\cdots\sum_{n_l}D^{(k_l)}_{m_ln_l}(A^{-1})\varphi^{(k_l)(*)}_{m_l}((a,A)f_l)\,\Psi_0\right\rangle\\
=\ & \sum_{n_1,\ldots,n_l}D^{(k_1)}_{m_1n_1}(A^{-1})\ldots D^{(k_l)}_{m_ln_l}(A^{-1})\langle\Psi_0,\varphi^{(k_1)(*)}_{n_1}((a,A)f_1)\cdots\varphi^{(k_l)(*)}_{m_l}((a,A)f_l)\,\Psi_0\rangle\\
=\ & \sum_{n_1,\ldots,n_l}D^{(k_1)}_{m_1n_1}(A^{-1})\ldots D^{(k_l)}_{m_ln_l}(A^{-1})W^{k_1^{(*)},\cdots,k_l^{(*)}}_{n_1,\cdots,n_l}((a,A)f_1,\cdots,(a,A)f_l)
\end{aligned}$$

which gives the result. □

Proposition 6.8. (Locality or Microcausality) *If the supports of test functions f_l, f_{l+1} consist only space like points, then*

$$W_{m_1,\cdots,m_l,m_{l+1},\ldots,m_n}^{k_1^{(*)},\cdots,k_l^{(*)},k_{l+1}^{(*)},\ldots,k_n^{(*)}}(f_1,\cdots,f_l,f_{l+1},\ldots,f_n)$$

$$=(-1)^{\sigma(k_l,k_{l+1})}W_{m_1,\cdots,m_{l+1},m_l,\ldots,m_n}^{k_1^{(*)},\cdots,k_{l+1}^{(*)},k_l^{(*)},\ldots,k_n^{(*)}}(f_1,\cdots,f_{l+1},f_l,\ldots,f_n)$$

for all possible indices.

Proof. Using locality in the axioms of Wightman fields

$$\begin{aligned}
&W_{m_1,\cdots,m_l,m_{l+1},\ldots,m_n}^{k_1^{(*)},\cdots,k_l^{(*)},k_{l+1}^{(*)},\ldots,k_n^{(*)}}(f_1,\cdots,f_l,f_{l+1},\ldots,f_n)\\
=&\ \langle\Psi_0,\varphi_{m_1}^{(k_1)(*)}(f_1)\cdots\varphi_{m_l}^{(k_l)(*)}(f_l)\varphi_{m_{l+1}}^{(k_{l+1})(*)}(f_{l+1})\cdots\varphi_{m_l}^{(k_l)(*)}(f_l)\,\Psi_0\rangle\\
=&\ \sigma(k,k')\langle\Psi_0,\varphi_{m_1}^{(k_1)(*)}(f_1)\cdots\varphi_{m_{l+1}}^{(k_{l+1})(*)}(f_{l+1})\varphi_{m_l}^{(k_l)(*)}(f_l)\cdots\varphi_{m_l}^{(k_l)(*)}(f_l)\,\Psi_0\rangle\\
=&\ (-1)^{\sigma(k_l,k_{l+1})}W_{m_1,\cdots,m_{l+1},m_l,\ldots,m_n}^{k_1^{(*)},\cdots,k_{l+1}^{(*)},k_l^{(*)},\ldots,k_n^{(*)}}(f_1,\cdots,f_{l+1},f_l,\ldots,f_n)
\end{aligned}$$

which gives the result. □

Proposition 6.9. (Spectrum Property) *Under the change of variables*

$$\xi_1=x_1-x_2,\ldots,\xi_{n-1}=x_{n-1}-x_n,\xi_n=x_n$$

where $x_1,\ldots,x_n\in\mathbb{R}^4$, each tempered distribution $W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}$ depends only on $\xi_1,\ldots,\xi_{n-1}$, that is

$$\frac{\partial W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}}{\partial\xi_n}=0$$

then there is a tempered distribution $M_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}\in\mathcal{S}'(\mathbb{R}^{4(n-1)})$, such that

$$W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}=M_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}\otimes 1$$

where 1 *is a constant function* 1 *on $\mathbb{R}^4$. Moreover, the Fourier transform $\tilde{M}_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}$ of the tempered distribution $M_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}$ is supported in the* $(n-1)-$fold *product of closed forward light cone $\overline{V_+}\times\cdots\times\overline{V_+}$.*

Proof. Since the Wightman distributions are translation invariant, due to the translation invariance property of the Wightman fields, the first statement is clear. For the second statement about the support, we first show that for all $\Psi,\Phi\in\mathcal{H}$, we have

$$\int_{\mathbb{R}^4}e^{ip\cdot a}\,da\,\langle\Psi,U(a,1)\,\Phi\rangle=0$$

for $p\notin\overline{V_+}$. To show this, first we observe that this identity is continuous about two vectors $\Psi,\Phi\in\mathcal{H}$, then we only need to show this in the dense domain where the unitary translation can be written as

$$U(a,1)=e^{-ia_\mu P^\mu}$$

for where $\{P^\mu\}$ is the energy-momentum operator. Then due to spectrum theorem

$$\begin{aligned} & \int_{\mathbb{R}^4} e^{ip\cdot a}\, da\, \langle \Psi, U(a,1)\, \Phi\rangle \\ = & \int_{\mathbb{R}^4} e^{ip\cdot a}\, da\, \langle \Psi, e^{-ia_\mu P^\mu}\, \Phi\rangle \\ = & \int_{\mathbb{R}^4} e^{ip\cdot a}\, da\, \langle \Psi, \int_{\overline{V_+}} e^{-ia_\mu \tilde{p}^\mu}\, dE(\tilde{p})\, \Phi\rangle \\ = & \int_{\overline{V_+}} \int_{\mathbb{R}^4} e^{ia_\mu (p^\mu - \tilde{p}^\mu)} da\, d\langle \Psi, E(\tilde{p})\, \Phi\rangle \\ = & \; 0 \end{aligned}$$

since we have assumed the spectrum of energy-momentum operator is in $\overline{V_+}$ and $p \notin \overline{V_+}$.

To show $\tilde{M}^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}$ has support inside $\overline{V_+} \times \cdots \times \overline{V_+}$, first observe that for $p \notin \overline{V_+}$, we have

$$0 = \int_{\mathbb{R}^4} e^{ip\cdot a}\, da\, \langle \Psi_0, \varphi^{(k_1)(*)}_{m_1}(f_1) \cdots \varphi^{(k_l)(*)}_{m_l}(f_l) U(-a,1)\, \varphi^{(k_{l+1})(*)}_{m_l+1}(f_{l+1}) \cdots \varphi^{(k_n)(*)}_{m_n}(f_n)\, \Psi_0\rangle$$

since

$$\begin{aligned} & \langle \Psi_0, \varphi^{(k_1)(*)}_{m_1}(f_1) \cdots \varphi^{(k_l)(*)}_{m_l}(f_l) U(-a,1) \varphi^{(k_{l+1})(*)}_{m_l+1}(f_{l+1}) \cdots \varphi^{(k_n)(*)}_{m_n}(f_n)\, \Psi_0\rangle \\ = & \langle \Psi_0, \varphi^{(k_1)(*)}_{m_1}(f_1) \cdots \varphi^{(k_l)(*)}_{m_l}(f_l) \varphi^{(k_{l+1})(*)}_{m_l+1}((-a,1)\, f_{l+1}) \cdots \varphi^{(k_n)(*)}_{m_n}((-a,1)\, f_n)\, \Psi_0\rangle \end{aligned}$$

we have

$$0 \;=\; \int_{\mathbb{R}^4} e^{ip\cdot a}\, da W^{k_1^{(*)},\cdots,k_l^{(*)},k_{l+1}^{(*)},\ldots,k_n^{(*)}}_{m_1,\cdots,m_l,m_{l+1},\ldots,m_n}(f_1,\cdots,f_l,(-a,1)\, f_{l+1},\ldots,(-a,1)\, f_n)$$

this implies

$$0 = \int_{\mathbb{R}^4} e^{ip\cdot a}\, da\, M^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}(\xi_1,\ldots,\xi_l + a,\ldots,\xi_{n-1})$$

which means if $p_l \notin \overline{V_+}$, then

$$\begin{aligned} & \tilde{M}^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}(p_1,\ldots,p_{n-1}) \\ = & \int_{\mathbb{R}^4} \cdots \int_{\mathbb{R}^4} d\xi_1 \ldots d\xi_l \ldots d\xi_{n-1} e^{i\sum p_j\cdot\xi_j} M^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}(\xi_1,\ldots,\xi_l,\ldots,\xi_{n-1}) \\ = & \int_{\mathbb{R}^4} \cdots \int_{\mathbb{R}^4} d\xi_1 \ldots d\xi_{n-1} e^{i\sum_{j\neq l} p_j\cdot\xi_j} \int_{\mathbb{R}^4} e^{ip_l\cdot\xi_l} d\xi_l M^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}(\xi_1,\ldots,\xi_l,\ldots,\xi_{n-1}) \\ = & \; 0 \end{aligned}$$

and this concludes the result. □

Note that we have used a different sign convention in Fourier transform, this convention is usually taken in physics.

Proposition 6.10. (Cluster Property) *Suppose $a \in \mathbb{R}^{1,3}$ is a space-like vector, then*

$$W^{k_1^{(*)},\cdots,k_l^{(*)},k_{l+1}^{(*)},\ldots,k_n^{(*)}}_{m_1,\cdots,m_l,m_{l+1},\ldots,m_n}(f_1,\cdots,f_l,(\lambda a,I)f_{l+1},\ldots,(\lambda a,I)f_n)\xrightarrow{\lambda\to+\infty}$$

$$W^{k_1^{(*)},\cdots,k_l^{(*)}}_{m_1,\cdots,m_l}(f_1,\cdots,f_l)W^{k_{l+1}^{(*)},\cdots,k_n^{(*)}}_{m_{l+1},\cdots,m_n}(f_{l+1},\cdots,f_n)$$

for all possible indices and test functions.

Proof. By definition and covariance, clearly

$$\begin{aligned}
&W^{k_1^{(*)},\cdots,k_l^{(*)},k_{l+1}^{(*)},\ldots,k_n^{(*)}}_{m_1,\cdots,m_l,m_{l+1},\ldots,m_n}(f_1,\cdots,f_l,(\lambda a,I)f_{l+1},\ldots,(\lambda a,I)f_n)\\
=\;&\langle\Psi_0,\varphi^{(k_1)(*)}_{m_1}(f_1)\cdots\varphi^{(k_l)(*)}_{m_l}(f_l),\varphi^{(k_{l+1})(*)}_{m_{l+1}}((\lambda a,I)f_{l+1}),\ldots,\varphi^{(k_n)(*)}_{m_n}((\lambda a,I)f_n)\,\Psi_0\rangle\\
=\;&\langle\Psi_0,\varphi^{(k_1)(*)}_{m_1}(f_1)\cdots\varphi^{(k_l)(*)}_{m_l}(f_l),U(\lambda a,I)\varphi^{(k_{l+1})(*)}_{m_{l+1}}(f_{l+1})U(\lambda a,I)^{-1},\ldots,U(\lambda a,\\
&I)\varphi^{(k_n)(*)}_{m_n}(f_n)\,U(\lambda a,I)^{-1}\Psi_0\rangle\\
=\;&\langle\Psi_0,\varphi^{(k_1)(*)}_{m_1}(f_1)\cdots\varphi^{(k_l)(*)}_{m_l}(f_l),U(\lambda a,I)\varphi^{(k_{l+1})(*)}_{m_{l+1}}(f_{l+1}),\ldots,\varphi^{(k_n)(*)}_{m_n}(f_n)\Psi_0\rangle\\
=\;&\langle(\varphi^{(k_l)-(*)}_{m_l}(f_l)\cdots\varphi^{(k_1)-(*)}_{m_1}(f_1)\,\Psi_0),U(\lambda a,I)(\varphi^{(k_{l+1})(*)}_{m_{l+1}}(f_{l+1}),\ldots,\varphi^{(k_n)(*)}_{m_n}(f_n)\Psi_0)\rangle
\end{aligned}$$

In general one can show that

$$\lim_{\lambda\to\infty}\langle\Phi,U(\lambda a,I)\,\Psi\rangle=\langle\Phi,\Psi_0\rangle\langle\Psi_0,\Psi\rangle$$

we do not produce the proof here, see [10]. It is clear the cluster property for the Wightman distribution follows from this result. □

6.3 Wightman Reconstruction Theorem

Now we collect all the properties from previous section and formulate the axioms for Wightman distributions, see page 117 of [44] and page 333 of [10].

Wightman distributions a collection of tempered distributions $\left\{W^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}\in\mathcal{S}'(\mathbb{R}^{4n})\middle| n\in\mathbb{N},0\leqslant m_i\leqslant 2s(k_i)+1<+\infty,1\leqslant i\leqslant n\right\}$ and $W^{[0]}$ without any index is assumes to be 1, where $s(k_i)$ is understood to be the spin described by k_i-th field, with following properties:

1. **(Hermiticity)** We have

$$W^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}(f_1,\cdots,f_n)=\overline{W^{k_n^{-(*)},\cdots,k_1^{-(*)}}_{m_n,\cdots,m_1}(f_n,\cdots,f_1)}$$

where the notation $-(*)$ means if we have index k_i, then we take $k_i^{-(*)}=k_i^*$, if we have k_i^*, we then take $k_i^{-(*)}=k_i$.

2. (Positivity) For any finite sequence of test functions $\left\{ f_{m_1,\cdots,m_i}^{k_1^{(*)},\cdots,k_i^{(*)}} \middle| f_{m_1,\cdots,m_i}^{k_1^{(*)},\cdots,k_i^{(*)}} \in \mathcal{S}(\mathbb{R}^{4i}) \right\}$, we have

$$\sum_{i,j} \sum_{\substack{k_1'^{(*)},\cdots,k_j'^{(*)} \\ m_1',\cdots,m_j'}} \sum_{\substack{k_1^{(*)},\cdots,k_i^{(*)} \\ m_1,\cdots,m_i}} W_{m_j',\cdots,m_1',m_1,\cdots,m_i}^{k_j'^{-(*)},\cdots,k_1'^{-(*)},k_1^{(*)},\cdots,k_i^{(*)}} \left(\overline{\check{f}_{m_1',\cdots,m_j'}^{k_1'^{(*)},\cdots,k_j'^{(*)}}} \otimes f_{m_1,\cdots,m_i}^{k_1^{(*)},\cdots,k_i^{(*)}} \right) \geqslant 0$$

where $\check{f}_{m_1',\cdots,m_j'}^{k_1'^{(*)},\cdots,k_j'^{(*)}}$ is the function $f_{m_1',\cdots,m_j'}^{k_1'^{(*)},\cdots,k_j'^{(*)}}(x_j, x_{j-1}, \ldots, x_1)$.

3. (Covariance) We have

$$\sum_{n_1,\ldots,n_l} D_{m_1 n_1}^{(k_1)}(A^{-1}) \ldots D_{m_l n_l}^{(k_l)}(A^{-1}) W_{n_1,\cdots,n_l}^{k_1^{(*)},\cdots,k_l^{(*)}}((a, A) f_1, \cdots, (a, A) f_l)$$
$$= W_{m_1,\cdots,m_l}^{k_1^{(*)},\cdots,k_l^{(*)}}(f_1, \cdots, f_l)$$

for any $(a, A) \in \mathbb{R}^{1,3} \rtimes SL(2, \mathbb{C})$, $l, k_1, \ldots, k_l \in \mathbb{N}$ and $0 \leqslant m_i \leqslant 2s(k_i) + 1, 1 \leqslant i \leqslant l$.

4. (Locality or Microcausality) If the supports of test functions f_l, f_{l+1} consist only space like points, then

$$W_{m_1,\cdots,m_l,m_{l+1},\ldots,m_n}^{k_1^{(*)},\cdots,k_l^{(*)},k_{l+1}^{(*)},\ldots,k_n^{(*)}}(f_1, \cdots, f_l, f_{l+1}, \ldots, f_n)$$

$$= (-1)^{\sigma(k_l,k_{l+1})} W_{m_1,\cdots,m_{l+1},m_l,\ldots,m_n}^{k_1^{(*)},\cdots,k_{l+1}^{(*)},k_l^{(*)},\ldots,k_n^{(*)}}(f_1, \cdots, f_{l+1}, f_l, \ldots, f_n)$$

for all possible indices.

5. (Cluster Property) Suppose $a \in \mathbb{R}^{1,3}$ is a space-like vector, then

$$W_{m_1,\cdots,m_l,m_{l+1},\ldots,m_n}^{k_1^{(*)},\cdots,k_l^{(*)},k_{l+1}^{(*)},\ldots,k_n^{(*)}}(f_1, \cdots, f_l, (\lambda a, I) f_{l+1}, \ldots, (\lambda a, I) f_n) \xrightarrow{\lambda \to +\infty}$$

$$W_{m_1,\cdots,m_l}^{k_1^{(*)},\cdots,k_l^{(*)}}(f_1, \cdots, f_l) W_{m_{l+1},\cdots,m_n}^{k_{l+1}^{(*)},\cdots,k_n^{(*)}}(f_{l+1}, \cdots, f_n)$$

for all possible indices and test functions.

6. (Spectrum Property) Under the change of variables

$$\xi_1 = x_1 - x_2, \ldots, \xi_{n-1} = x_{n-1} - x_n, \xi_n = x_n$$

where $x_1, \ldots, x_n \in \mathbb{R}^4$, each tempered distribution $W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}$ depends only on $\xi_1, \ldots, \xi_{n-1}$, that is

$$\frac{\partial W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}}}{\partial \xi_n} = 0$$

then there is a tempered distribution $M_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}} \in \mathcal{S}'(\mathbb{R}^{4(n-1)})$, such that

$$W_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}} = M_{m_1,\cdots,m_n}^{k_1^{(*)},\cdots,k_n^{(*)}} \otimes 1$$

where 1 is a constant function 1 on $\mathbb{R}^4$. Moreover, the Fourier transform $\tilde{M}^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}$ of the tempered distribution $M^{k_1^{(*)},\cdots,k_n^{(*)}}_{m_1,\cdots,m_n}$ is supported in the $(n-1)-$fold product of closed forward light cone $\overline{V_+}\times\cdots\times\overline{V_+}$.

The celebrated Wightman reconstruction theorem in the following, show there is an one-to-one correspondence between Wightman fields and Wightman distributions. Thus one can construct a quantum field theory by proposing its correlation functions.

Theorem 6.11. (Wightman Reconstruction Theorem) *For a given set of Wightman distributions satisfying axioms 1 to 6, there exists a unique Wightman quantum field theory up to unitary equivalence.*

For connivence, we produce the proof for the case of single Hermitian scalar field, the general case is true with more technical analysis.

Proof. Basically this is a GNS type construction. First we construct the Hilbert space. Consider the vector space H of sequence $f=(f_0,f_1,\ldots)$ where $f_i\in\mathcal{S}(\mathbb{R}^{4i})$ with only a finite number of nonzero components. The vacuum vector is $\Psi_0=(1,0,0,\ldots)$. Define skew-linear form by

$$\langle f,g\rangle:=\sum_{i,j=0}^{\infty}W_{i+j}(\bar{\check{f}}_i\otimes g_j)$$

which is clearly linear in g, conjugate-linear in f. It is skew symmetric by hermiticity and non-negative definite by positivity. Representation of Poincaré group is given by

$$U(a,\Lambda)f=(f_0,(a,\Lambda)f_1,(a,\Lambda)f_2,\ldots)$$

then by covariance assumption, the skew linear form is preserved, and clearly the vacuum is an invariant vector.

Now we define the field. For any test function $h\in\mathcal{S}$, the operator $\varphi(h)$ is given by

$$\varphi(h)f=(0,h\otimes f_0,h\otimes f_1,\ldots)$$

with transform law $U(a,\Lambda)\varphi(h)U(a,\Lambda)^{-1}=\varphi((a,\Lambda)h)$ easily verified. It's easy to see that φ is a operator valued distribution, and real, which means $\varphi(\bar{h})=\varphi(h)^*$.

Define the subspace $H_0:=\{f\in H|\langle f,f\rangle=0\}$, which is clearly leave invariant by acting $U(a,\Lambda)$ and $\varphi(h)$. Thus one can complete the space H/H_0 to get the physical Hilbert space $\mathcal{H}$, and dense domain $D=H/H_0$, thus $\mathcal{H}$ is separable. The element in D corresponding to f is denoted by Ψ_f. Clearly this domain is invariant under the action of field operator and the unitary representation of Poincaré group.

The representation of Poincaré group on H induced a representation on H/H_0, hence can be extended to a strong continuous unitary representation on $\mathcal{H}$. The strong continuity can be seen by

$$\begin{aligned}&\langle\Psi_f-U(a,\Lambda)\Psi_f,\Psi_f-U(a,\Lambda)\Psi_f\rangle\\=&\langle f-U(a,\Lambda)f,f-U(a,\Lambda)f\rangle\\=&\sum_{i,j}W_{i+j}\left(\overline{(\check{f}_i-(a,\Lambda)\,\check{f}_i)}\otimes(f_j-(a,\Lambda)f_j)\right)\end{aligned}$$

which is clearly convergent to zero as $(a,\Lambda)\to(0,I)$, since the Poincaré group acts on the space of test functions continuously. This continuity can be extended to all $\mathcal{H}$ easily by noting that D is dense.

Denote the corresponding Ψ_0 in $\mathcal{H}$ also by Ψ_0, which is the vacuum operator. To show it is the unique invariant vector, consider another one Ψ_0', if exist, orthogonal to Ψ_0 and normalized without loose generality. If it has the form Ψ_f for some $f\in H$, then for all space-like vector a and cluster property, we have

$$\begin{aligned}
&\langle\Psi_f,\Psi_f\rangle\\
=&\lim_{\lambda\to\infty}\langle\Psi_f,U(\lambda a,I)\,\Psi_f\rangle\\
=&\lim_{\lambda\to\infty}\sum_{i,j=0}^{\infty}W_{i+j}(\bar{\check{f}}_i\otimes(\lambda a,I)f_j)\\
=&\sum_{i,j=0}^{\infty}W_i(\bar{\check{f}}_i)W_j(f_j)\\
=&\langle\Psi_0',\Psi_0\rangle\langle\Psi_0,\Psi_0'\rangle=0
\end{aligned}$$

and for $\Psi_0'\in\mathcal{H}$ which does not have this form, one can approximate if by a normalized Ψ_f, then one can still get $\langle\Psi_0',\Psi_0'\rangle=0$. Thus this vacuum is unique. It is not hard to see that this vacuum is cyclic, and the vacuum expectation is just the Wightman distributions in the assumption.

We need to show that the spectrum of the energy-momentum lies in the closed forward light cone. Due to spectrum property, we have

$$\int_{\mathbb{R}^4}e^{ip\cdot a}\,da\,\langle\Psi_f,U(a,1)\,\Psi_g\rangle=0$$

for any $f,g\in H$, and $p\notin\overline{V_+}$, which means p is not in the spectrum of the energy-momentum.

This completes the proof of existence part. We need to how the uniqueness up to unitary transformation. Suppose there is another Wightman fields with the same set of correlation functions, the corresponding informations are $\mathcal{H}',\Psi_0',U'(a,\Lambda),\varphi'$. Then the map V from $\mathcal{H}$ to $\mathcal{H}'$, defined by

$$V\Psi_f=f_0\Psi_0'+\varphi'(f_1)\Psi_0'+\varphi'(f_2^{(1)})\varphi'(f_2^{(2)})\Psi_0'+\cdots$$

where $f=(f_0,f_1,f_2^{(1)}\otimes f_2^{(2)},f_3^{(1)}\otimes f_3^{(2)}\otimes f_3^{(3)},\ldots)$, note that vectors of this form are dense in D and $\mathcal{H}$, and V preserves the inner product, hence extends to a unitary operator by cyclicity of Ψ_0'. The relations

$$\varphi(h)=V^{-1}\varphi'(h)V,\,U(a,\Lambda)=V^{-1}U'(a,\Lambda)V$$

are easy to verify. □

6.4 Osterwalder-Schrader Axioms for Schwinger Functions and Euclidean Quantum Field Theory

Before talking about the Schwinger functions, we discuss heuristically what we are doing. In quantum theory, according to Born's rule, one is interested in the quantity $\langle\,\varphi_f\,|U\,(t_1-t_0)|\,\varphi_i\,\rangle$, which is the probability of finding the system in the final configuration $|\,\varphi_f\,\rangle$ after some time from initial configuration $|\,\varphi_i\,\rangle$, here the time evolution operator is given by the exponential of Hamiltonian $U\,(t_1-t_0)=e^{-i(t_1-t_0)\frac{H}{\hbar}}$. The essence of Feynman path integral is to represent this transition probability into a probability theory on the space of all historical configurations between two times, that is

$$\langle\varphi_f\,|U\,(t_1-t_0)|\,\varphi_i\,\rangle=\langle\varphi_f\left|e^{-i(t_1-t_0)\frac{H}{\hbar}}\right|\varphi_i\,\rangle=\int_{\substack{\varphi(t_0)=\varphi_i\\ \varphi(t_1)=\varphi_f}}e^{\frac{i}{\hbar}S[\varphi]}\,\mathcal{D}[\varphi]$$

where $S[\varphi]$ is the classical action functional, for example, if φ is the configuration of the historical position q of a particle, then

$$S[q]=\int_{t_0}^{t_1}\left[\frac{1}{2}m\dot{q}(s)-V(q(s))\right]ds$$

or if φ is the configuration of the historical distribution φ of a scalar field, then

$$S[\varphi]=\int_{t_0}^{t_1}\int_{\mathbb{R}^3}\left[\frac{1}{2}(\partial_t^2-\Delta)\varphi(t,x,y,z)-\frac{1}{2}m^2\varphi^2-V(\varphi(t,x,y,z))\right]ds\,dx\,dy\,dz.$$

The goodness of path integral representation is, the expectation value of time ordered product operators (operators in Heisenberg picture), can be turned into some kind of probabilistic expectation value over histories, that is

$$\begin{aligned}&\int_{\substack{\varphi(t_0)=\varphi_i\\ \varphi(t_{n+1})=\varphi_f}}\varphi(t_n)\cdots\varphi(t_1)e^{\frac{i}{\hbar}S[\varphi]}\,\mathcal{D}[\varphi]\\ =\;&\int d\varphi_n\ldots\int d\varphi_1\langle\varphi_f\,|U(t_{n+1}-t_n)|\varphi_n\rangle\varphi_n\langle\varphi_n|\cdots|\varphi_1\rangle\varphi_1\langle\varphi_1|U(t_1-t_0)|\varphi_i\rangle\\ =\;&\langle\varphi_f\,|U(t_{n+1})\Phi(t_n)\cdots\Phi(t_1)U(-t_0)|\varphi_i\rangle\\ =\;&\langle t=0,\varphi_f\,|\Phi(t_n)\cdots\Phi(t_1)|\,0,\varphi_i\rangle\end{aligned}$$

where $t_{n+1}>t_n>t_{n-1}>\cdots>t_1>t_0$ and Φ is the quantum analog of φ, usually given by canonical quantization. The state $|\,t=0,\varphi_i\rangle$ means the time zero state which evolves to labeled by configuration φ_i in time t_0.

Usually the Hamiltonian is positive $H\geqslant 0$ in the sense $\langle\psi|H|\,\psi\rangle\geqslant 0$ for all possible state $|\,\psi\,\rangle$, then clear if we replace the positive time t in $e^{-it\frac{H}{\hbar}}$ by $-i\tau$ for $\tau>0$, then

$$e^{-it\frac{H}{\hbar}}\to e^{-\tau\frac{H}{\hbar}}$$

should also make sense, and the path integral formula turns into

$$\langle\,\varphi_f\,|U_E(\tau_1-\tau_0)|\varphi_i\,\rangle=\langle\varphi_f\left|e^{-(\tau_1-\tau_0)\frac{H}{\hbar}}\right|\varphi_i\,\rangle=\int_{\substack{\varphi(\tau_0)=\varphi_i\\ \varphi(\tau_1)=\varphi_f}}e^{-\frac{1}{\hbar}S_E[\varphi]}\,\mathcal{D}[\varphi]$$

where $S_E[\varphi]$ is the Euclidean action, which is positive usually. We have an analog formula for expectations

$$\langle \tau = 0, \varphi_f \,|\Phi(\tau_n) \cdots \Phi(\tau_1)|\, \tau = 0, \varphi_i\rangle = \int_{\substack{\varphi(\tau_0)=\varphi_i \\ \varphi(\tau_{n+1})=\varphi_f}} \varphi(\tau_n) \cdots \varphi(\tau_1) e^{-\frac{1}{\hbar}S_E[\varphi]} \mathcal{D}[\varphi]$$

with $\tau_{n+1} > \tau_n > \tau_{n-1} > \cdots > \tau_1 > \tau_0$.

We observe the similarity between the Euclidean path integral measure and Boltzmann distribution in equilibrium statistical mechanics, where the Euclidean time τ and the temperature T should be related by $\tau = \frac{\hbar}{kT}$. According to quantum statistical mechanics, if we assume the spectrum of H is discrete for convenience, say $E_0 < E_1 < \cdots < E_n < \cdots$, the ensemble average at temperature T of an operator A should be given by

$$\langle A\rangle = \frac{\mathrm{Tr}\big(e^{-\frac{1}{kT}H} A\big)}{\mathrm{Tr}\big(e^{-\frac{1}{kT}H}\big)} = \frac{\sum_{i=0}^{\infty} e^{-\frac{1}{kT}E_i} \langle i|A|\, i\rangle}{\sum_{i=0}^{\infty} e^{-\frac{1}{kT}E_i}}$$

where k is the Boltzmann constant. Clearly if we consider the limit $T \to 0$, we have

$$\frac{\sum_{i=0}^{\infty} e^{-\frac{1}{kT}E_i} \langle i|A|\, i\rangle}{\sum_{i=0}^{\infty} e^{-\frac{1}{kT}E_i}} \xrightarrow{T\to 0} \langle 0|A|\, 0\rangle$$

and this limit corresponds to the limit $\tau \to +\infty$. Thus the equation

$$\langle \Phi(\tau_n) \cdots \Phi(\tau_1)\rangle = \frac{\int d\varphi' \int_{\substack{\varphi(\tau_0)=\varphi_i \\ \varphi(\tau_{n+1})=\varphi_f \\ \varphi(0)=\varphi'}} \varphi(\tau_n) \cdots \varphi(\tau_1) e^{-\frac{1}{\hbar}S_E[\varphi]} \mathcal{D}[\varphi] d\varphi'}{\int d\varphi' \int_{\substack{\varphi(\tau_0)=\varphi_i \\ \varphi(\tau_{n+1})=\varphi_f \\ \varphi(0)=\varphi'}} e^{-\frac{1}{\hbar}S_E[\varphi]} \mathcal{D}[\varphi] d\varphi'}$$

subjects to the limit $\tau_0 \to -\infty, \tau_{n+1} \to +\infty$, one has

$$\langle 0|\Phi(\tau_n) \cdots \Phi(\tau_1)|\, 0\rangle = \frac{\int \varphi(\tau_n) \cdots \varphi(\tau_1) e^{-\frac{1}{\hbar}S_E[\varphi]} \mathcal{D}'[\varphi]}{\int e^{-\frac{1}{\hbar}S_E[\varphi]} \mathcal{D}'[\varphi]}$$

where the path integral measure is over the space of all configurations with some decay property (one can also pose the periodic condition on Euclidean time, and then study the limit that the period goes to infinity). Note that the right hand can be interpreted as the moment in probability theory, and the left hand side is the vacuum expectations. Clearly if we change back from the Euclidean time to Minkowski time, we get the corresponding formula of vacuum expectation of time ordered product and averaging over path integral measure, and this explain what we are really describing about with the Wightman distributions.

Now we describe how to construct the Schwinger functions by analytic continuation of Wightman distributions. Such Schwinger functions are the Euclidean correlation functions we just described. For simplicity we do it for a single scalar Boson field.

The analytic continuation takes three steps. The first step is to continue the Wightman distributions to complex variables in the tube. This is based on the following theorem, for the proof, see theorem 3.5 of [44]. We continue to use the notation in the axioms of Wightman distributions.

Theorem 6.12. *There are holomorphic functions* $W_n(z_1,\ldots,z_n)$ *and* $M_{n-1}(\tilde{z}_1,\ldots,\tilde{z}_{n-1})$*, where* $z_i=(z_i^0,z_i^1,z_i^2,z_i^3)$ *and denote* $\tilde{z}_j=x_j-i\,y_j$*, such that*

$$W_n(z_1,\cdots,z_n)=M_{n-1}(z_1-z_2,\ldots,z_{n-1}-z_n)$$

defined on the tube $T_{n-1}=\{-\mathrm{Im}\,(z_i-z_{i+1})\in V_+|\, i=1,\cdots,n-1\}$*, and polynomially bounded, such that the boundary value* $M_{n-1}(\tilde{z}_1,\ldots,\tilde{z}_{n-1})$ *is the distribution* M_{n-1}*, i.e*

$$\lim_{y_j\to 0} M_{n-1}(x_1-i\,y_1,\cdots,x_{n-1}-i\,y_{n-1})=M_{n-1}$$

in the sense of tempered distribution.

Next step, we use the covariance property of the Wightman distributions. We need to introduce the complex Lorentz group, which is the connected component of the identity of the group of complex matrices that preserve the complex bilinear form

$$\langle z_1,z_2\rangle=z_1^0z_2^0-z_1^1z_2^1-z_1^2z_2^2-z_1^3z_2^3$$

on the space $\mathbb{C}^{1,3}$. Denote this group by $L(\mathbb{C})$. The fundamental group of $L(\mathbb{C})$ is also $\mathbb{Z}_2$ and its covering group is given by $SL(2,\mathbb{C})\times SL(2,\mathbb{C})$. A point in the space $\mathbb{C}^{1,3}$ is called a Euclidean point if it as the form $(-ix^0,x^1,x^2,x^3)$ where $x^0,x^1,x^2,x^3\in\mathbb{R}$, the name come from the fact that when then complex bilinear form restricts to such real subspace, we have Euclidean inner product (up to a minus sign). We denote these points by E.

Since we are dealing with Boson fields here, it is enough to consider the Lorentz group and its complexification. Since we know that the tempered distributions W_n transforms as following

$$W_n(x_1,\cdots,x_n)=W_n(\Lambda\,x_1,\cdots,\Lambda\,x_n)$$

for any $\Lambda\in SO^+(1,3)$, by the uniqueness of the analytic continuation we have

$$W_n(z_1,\cdots,z_n)=W_n(\Lambda\,z_1,\cdots,\Lambda\,z_n)$$

and

$$M_{n-1}(\tilde{z}_1,\cdots,\tilde{z}_{n-1})=M_{n-1}(\Lambda\,\tilde{z}_1,\cdots,\Lambda\,\tilde{z}_{n-1})$$

but we can see here that the Lorentz transformations preserve the tube T_{n-1}, now one can use this identity to extend the action of Lorentz group to complex Lorentz group. Then one can define M_{n-1} on the so called extended tube

$$T_{n-1}^e:=\bigcup_{\Lambda\in L(\mathbb{C})}\Lambda T_{n-1}$$

and this extension is single valued, see theorem 2-11 of [44]. The points in the extended tube can be characterized by the following theorem due to Jost.

Proposition 6.13. *A point* $(\tilde{z}_1, \cdots, \tilde{z}_{n-1})$ *such that* $\tilde{z}_1, \cdots, \tilde{z}_{n-1} \in \mathbb{R}^{1,3}$ *is in* T_{n-1}^e *if and only if all points of the form*

$$\sum_{i=1}^{n-1} \lambda_i \tilde{z}_i, \quad \lambda_i \geq 0 \text{ and } \sum_{i=1}^{n-1} \lambda_i > 0$$

are space-like.

These points are called the Jost points. The last step we use the locality condition, in our case of single scalar Boson field, we have

$$W_n(x_1, \ldots, x_n) = W_n(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

for any permutation σ. Thus one can define the value of $W_n(z_{\sigma(1)}, \ldots, z_{\sigma(n)})$ by $W_n(z_1, \cdots, z_n)$, and hence one can define M_{n-1} when $W_n(z_{\sigma(1)}, \ldots, z_{\sigma(n)})$ is define. It is a fact that this analytic continuation is well-defined and also single-valued. The intersection of Euclidean points E^n and W_n's holomorphic domain is the following set

$$\mathbb{R}_{\neq}^{4n} := \{(x_1, x_2, \cdots, x_n) | \, x_i \in \mathbb{R}^4, x_i \neq x_j \text{ if } i \neq j\} \subset E^n$$

and the Schwinger function S_n is defined to be the restriction of W_n on this subspace, note that it is a polynomially bounded analytic function, hence can also be view as a tempered distribution.

For the general case of analytic continuation of Wightman distributions for tensor or spinor fields, this procedure also works.

Now we introduce the Osterwalder-Schrader axioms for scalar boson fields, see [35] and [36]. We introduce the sets

$$\mathbb{R}_{\neq}^{4n} := \{(x_1, x_2, \cdots, x_n) | \, x_i \in \mathbb{R}^4, x_i \neq x_j \text{ if } i \neq j\}$$

$$\mathbb{R}_{<}^{4n} := \{(x_1, x_2, \cdots, x_n) | \, x_i \in \mathbb{R}^4, 0 < x_{1,0} < x_{2,0} < \cdots < x_{n,0}\}$$

$$\mathbb{R}_{-}^{4(n-1)} := \{(\xi_1, \cdots, \xi_{n-1}) | \, \xi_i \in \mathbb{R}^4, \xi_{i,0} < 0, i = 1, 2, \cdots, n-1\}$$

and the time reflection operator

$$\Theta f(x_0, x_1, x_2, x_3) = f(-x_0, x_1, x_2, x_3)$$

$$\Theta(x_0, x_1, x_2, x_3) = (-x_0, x_1, x_2, x_3)$$

$$\Theta f(x_1, \ldots, x_n) = f(\Theta x_1, \ldots, \Theta x_n)$$

Schwinger distribution is a collection of tempered distributions

$$\{S_n \in \mathcal{S}'(\mathbb{R}_{\neq}^{4n}) | n \in \mathbb{N}\}$$

and $S_0 := 1$, with following properties:

1. (Linear Growth) There exists an integer $s \in \mathbb{N}$, and a sequence $\{\sigma_n\}$ of positive numbers, such that

$$\sigma_n \leqslant C(n!)^{C'}$$

for some constants C, C' independent of n, and

$$|S_n(f)| \leqslant \sigma_n \|f\|_{n \times s}$$

for all $f\in\mathbb{R}^{4n}_{\neq}$ and $n\in\mathbb{N}$.

2. (Euclidean Invariance) We have

$$S_n((a,A)f_1,\cdots,(a,A)f_n)$$
$$=S_n(f_1,\cdots,f_n)$$

for any $(a,A)\in\mathbb{R}^4\rtimes \mathrm{SO}(4)$, and $f_i\in\mathcal{S}(\mathbb{R}^{4n}_{\neq})$.

3. (Reflection Positivity) For any finite sequence of test functions $\{f_i|f_i\in\mathcal{S}(\mathbb{R}^{4i})$ supported in $\mathbb{R}^{4i}_{<}\}$, we have

$$\sum_{i,j} S_{i+j}(\overline{\Theta\check{f}_j}\otimes f_i)\geqslant 0.$$

.

4. (Symmetry) For all test functions

$$S_n(f_1,\cdots,f_l,f_{l+1},\ldots,f_n)$$
$$=S_n(f_1,\cdots,f_{l+1},f_l,\ldots,f_n)$$

hence for all permutations of $\{1,\ldots,n\}$ and $f_1\otimes\cdots\otimes f_n$ supported in $\mathbb{R}^{4n}_{\neq}$.

5. (Cluster Property) Suppose $a\in\mathbb{R}^4$ is a non-zero vector of the form $(0,a_1,a_2,a_3)$, then

$$S_{m+n}(f_m,(\lambda a,I)f_n)\xrightarrow{\lambda\to+\infty}$$
$$S_m(f_m)S_n(f_n)$$

for all $f_m\in\mathbb{R}^{4m}_{<}$ and $f_n\in\mathbb{R}^{4n}_{<}$.

It can be shown that the Schwinger functions defined by analytic continuation of Wightman distributions satisfy all these axioms except the linear growth condition, in [36] they showed that if one assumes the growth of Wightman distributions, one can produce these linear growth conditions on Schwinger functions, and then they showed these two systems of axioms are equivalent.

The property of reflection positivity is simply a result of positivity of Wightman distributions together with the fact that for Minkowski time we have

$$\varphi(t,x)=e^{itP^0-ix_1P^1-ix_2P^2-ix_3P^3}\varphi(0,0)e^{-itP^0+ix_1P^1+ix_2P^2+ix_3P^3}$$

and after using Wick rotation $t=-i\tau$ we have

$$\varphi(\tau,x)=e^{\tau P^0-ix_1P^1-ix_2P^2-ix_3P^3}\varphi(0,0)e^{-\tau P^0+ix_1P^1+ix_2P^2+ix_3P^3}$$

then we have

$$\varphi^*(\tau,x)=\varphi(-\tau,x)$$

for any Euclidean time τ, note that conjugate transformation of a Hermitian scalar Boson field in Euclidean theory is different from Minkowski theory. This is a crucial discovery to recover the Wightman fields from Schwinger functions, and this property is found to be fruitful to have many applications in other problems, see [28] and references therein.

Note that this axiom can also be formulated in any dimension $d > 1$. The first condition on linear growth was appeared in [36], since they found some errors in the argument of the first paper [35], and this condition makes the Osterwalder-Schrader axioms stronger than the Wightman axioms. In the page 397 of the book [10], they gave a different condition corresponding to the spectrum condition in the Wightman axioms and they considered more general fields of spinor or tensor character. And then they proved that their axioms are equivalent to the Wightman distribution.

Theorem 6.14. (Osterwalder-Schrader Reconstruction) *There exist a unique Wightman quantum field theory whose Schwinger functions agree with the given set with properties listed above.*

The proof is rather technical, we do not produce it here, see [36]. Note that in there papers, they remarked that this theorem also works for the general spin or tensor fields.

There is another theorem which could also have the name Osterwalder-Schrader reconstructions theorem, just like the Wightman reconstruction theorem, which says that these assumptions of Schwinger functions can produce a set of operator-valued distributions on some Hilbert space, with a similar set of axioms as in Wightman fields, but are Euclidean invariant.

A different set of Osterwalder-Schrader axioms is given Jaffe and Glimm [21], which characterize the probability measure on space of tempered distributions $\mathcal{S}'(\mathbb{R}^d)$ (see the footnotes on page 91 of [21]).

A Euclidean quantum field theory is a probability measure $d\mu$ on $\mathcal{S}'(\mathbb{R}^d)$, such that following conditions are satisfied:

1. (Analyticity) Define the generating functional $S[f]$ by

$$S[f] := \int_{\mathcal{S}'(\mathbb{R}^d)} e^{\phi(f)} d\mu$$

then for every finite set of test functions $f_j \in \mathcal{S}(\mathbb{R}^d)$, $j = 1, 2, \ldots, n$, and complex numbers $z = \{z_1, \ldots, z_n\} \in \mathbb{C}^n$, the function

$$z \mapsto S\left[\sum_{j=1}^{n} z_j f_j\right]$$

is an entire function on $\mathbb{C}^n$.

2. (Regularity) There is a constant p with $1 \leq p \leq 2$ and c, such that for all $f \in \mathcal{S}(\mathbb{R}^d)$, we have

$$|S[f]| \leq e^{c(\|f\|_{L^1} + \|f\|_{L^p}^p)}$$

If $p = 2$, then we also assume there exists a locally integrable function $S_2(x_1, x_2)$ on $\mathbb{R}^d \times \mathbb{R}^d$ such that

$$\int_{\mathcal{S}'(\mathbb{R}^d)} \phi(f_1)\phi(f_2) d\mu = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f_1(x_1) f_2(x_2) S_2(x_1, x_2) dx_1 dx_2$$

for all f_1, $f_2 \in \mathcal{S}(\mathbb{R}^d)$.

3. (Euclidean Invariance) The Euclidean group $\mathbb{R}^d \rtimes \mathrm{SO}(d)$ act on $\mathcal{S}(\mathbb{R}^d)$ by $\{(a,R)f\}(x) = f(R^{-1}(x-a))$, then we assume that $S[f]$ is invariant under Euclidean symmetries, that is $S[f] = S[(a,R)f]$ for all $(a,R) \in \mathbb{R}^d \rtimes \mathrm{SO}(d)$.

4. (Reflection Positivity) For any sequence $f_1, \ldots, f_n$ of test functions supported on the upper half plane $\mathbb{R}^d_+ := \{(x_0, x_1, \ldots, x_{d-1}) | x_0 > 0\}$, the matrix defined by

$$M_{ij} := S[f_i + \Theta f_j^*]$$

is positive semi-definite, where Θ is time reflection operator, i.e. $\Theta f(x_0, x_1, \ldots, x_{d-1}) = f(-x_0, x_1, \ldots, x_{d-1})$.

5. (Ergodicity) For any test function f supported on the upper half plane $\mathbb{R}^d_+$, we have

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t e^{\phi(T(-s)f)} ds = \int_{\mathcal{S}'(\mathbb{R}^d)} e^{\phi(f)} d\mu$$

where $T(-s) := ((-s, 0, \ldots, 0), I) \in \mathbb{R}^d \rtimes \mathrm{SO}(d)$.

The Schwinger function can be defined from this axioms by

$$S_n(f_1, \cdots, f_n) := \mathbb{E}[\phi(f_1) \ldots \phi(f_n)] = \int_{\mathcal{S}'(\mathbb{R}^d)} \phi(f_1) \ldots \phi(f_n) d\mu$$

for any $f_1, \ldots, f_n \in \mathcal{S}(\mathbb{R}^d)$. The expectation converges due to the analyticity condition, one can use the Cauchy integral formula for several complex variables. See chapter 19 of [21] for the proof that axioms by Jaffe and Glimm implies Wightman axioms. We should remark that there are many other choices to produce a system of axioms that can used to deduce the Wightman axioms, for example see [19] and [16].

7 The Stochastic Quantization of Abelian Higgs model

We introduced the axioms that a quantum field theory should satisfy, then one has a natural question whether there is an interesting example of interacting quantum fields satisfying these axioms. In this chapter, we first concern the problem of constructive quantum field theory by introducing the stochastic quantization method. Then we introduce the Abelian Higgs model, including its definition, difficulties, related works, and so on.

7.1 Stochastic Quantization

In the last chapter, we introduced the Euclidean quantum field theory where we described the axioms that Schwinger functions should satisfy. An important question is, how can we construct an example of Euclidean quantum field theory, that satisfying these axioms, hence one get a quantum field theory in Minkowski space time by using the Osterwalder-Schrader reconstruction theorem.

We noted the similarity between the Euclidean path integral measure $e^{-\frac{1}{\hbar}S_E[\varphi]}\mathcal{D}'[\varphi]$ the Boltzmann distribution in equilibrium statistical mechanics. In the study of statistical mechanics, such an equilibrium state is obtained by preparing some non-equilibrium statistical mechanical system and subject to some relaxation time. This process can be described by the Langevin dynamics. Thus to get the Euclidean path integral measure $e^{-\frac{1}{\hbar}S_E[\varphi]}\mathcal{D}'[\varphi]$, one imagine a non-equilibrium system, evolves according to the Langevin equation

$$\partial_t\varphi(t,x) = -\left(\frac{\delta S_E[\varphi]}{\delta\varphi}\right)|_{\varphi=\varphi(t,x)} + \xi$$

or equivalently

$$\partial_t\varphi(t,x) = -\frac{\delta\hat{S}_E[\varphi]}{\delta\varphi} + \xi, \hat{S}_E[\varphi] = \int dt\, S_E[\varphi]$$

where δ is the functional derivative, t is some fictitious time describes the evolution of non-equilibrium system, and ξ is the space-(fictitious)time white noise. Since the white noise is delta correlated in fictitious time direction, then the solution process if exists, should be a Markov process.

Let us see formally why the Euclidean path integral measure is an stationary measure of this equation. We need to compute the Fokker-Planck equation of this Langevin equation, which is the dynamical equation of the probability distribution $P(\varphi,t)$. For any functional $F[\varphi]$, we have

$$\mathbb{E}_\xi[F[\varphi(t)]] = \int F(\varphi)P(\varphi,t)D\varphi$$

where $\mathbb{E}_\xi$ is the expectation with respect to the law of white noise, since the solution of the Langevin equation should be a functional of the white noise. In order to find the time derivative of both side, we need to find the first order expansion of left hand, which is

$$
\begin{aligned}
&\mathbb{E}_\xi\left[F\left(\varphi\left(t+dt\right)\right)-F(\varphi(t))\right]\\
&=\mathbb{E}_\xi\left[\int d\,x\,\frac{\delta\,F}{\delta\,\varphi}\,\partial_t\,\varphi\,d\,t+\iint dx\,dx'\,\frac{\delta^2\,F}{\delta\,\varphi\,\delta\,\varphi'}\,(\partial_t\,\varphi\,d\,t)\,(\partial_t\,\varphi\,d\,t)\right]\\
&=\mathbb{E}_\xi\left[\int d\,x\,\frac{\delta\,F}{\delta\,\varphi}\left(-\frac{\delta\,\hat{S}_E}{\delta\,\varphi}+\xi\right)d\,t+\iint dx\,dx'\,\frac{\delta^2\,F}{\delta\,\varphi\,\delta\,\varphi'}\,(\xi\,d\,t)\,(\xi'\,d\,t)\right]\\
&=\mathbb{E}_\xi\left[-\int dx\,\frac{\delta\,F}{\delta\,\varphi}\frac{\delta\,\hat{S}_E}{\delta\,\varphi}\right]dt+\iint dx\,dx'\,\mathbb{E}_\xi\left[\frac{\delta^2\,F}{\delta\,\varphi\,\delta\,\varphi'}\right]\underbrace{\mathbb{E}_\xi\left[\xi\,d\,t\,\xi'\,d\,t\right]}_{\delta(x-x')dt}\\
&=\int dx\int D\varphi\left[-\frac{\delta\,F}{\delta\,\varphi}\,\frac{\delta\,\hat{S}_E}{\delta\,\varphi}+\frac{\delta^2\,F}{\delta\,\varphi^2}\right]P(\varphi,t)dt\\
&=\int dx\int D\varphi\,F\frac{\delta}{\delta\,\varphi}\left[\frac{\delta\,\hat{S}_E}{\delta\,\varphi}+\frac{\delta}{\delta\,\varphi}\right]P(\varphi,t)dt
\end{aligned}
$$

where we used that $\varphi(t)$ and $\xi(t)$ are independent to deduce any functional of φ up to time t should be independent with $\xi(t)$, since $\varphi(t)$ only depends on instantaneous past of $\xi(t)$ due to the Markov property. Thus

$$
\begin{aligned}
&\frac{d}{dt}\mathbb{E}_\xi[F[\varphi(t)]]\\
&= \int dx\int D\varphi\,F\frac{\delta}{\delta\,\varphi}\left[\frac{\delta\,\hat{S}_E}{\delta\,\varphi}+\frac{\delta}{\delta\,\varphi}\right]P(\varphi,t)\\
&= \int F(\varphi)\partial_t P(\varphi,t)D\varphi
\end{aligned}
$$

holds for any functional F. Then the Fokker-Planck equation is given by

$$
\partial_t P(\varphi,t)=\int dx\frac{\delta}{\delta\,\varphi}\left[\frac{\delta\,\hat{S}_E}{\delta\,\varphi}+\frac{\delta}{\delta\,\varphi}\right]P(\varphi,t)
$$

and it is clear $e^{-S_E[\varphi]}$ is a stationary solution.

The stochastic quantization method is proposed by Parisi and Wu in [37], and one should note the difference between this method and Nelson's in [32] [33] [34], where Nelson used the real time to derive the Schrödinger equation. The method of Parisi and Wu is based on an hypothetical process depending on a fictitious time. And one can use different Langevin equations to describe the Euclidean path integral measure, see [14]. There are cases where the stochastic quantization method does not work, that is the proposed Langevin dynamics does not have a limit stationary measure, see [17].

The stochastic quantization equation requires us to study the solution theory for stochastic partial differential equations, establish the meaning of the equation, the existence of local solutions, and most importantly the long time existence of the solution, since we need to take $t \to \infty$ to get the equilibrium state. This is already a hard problem in general, but the requirement that the limiting measure should satisfy the Osterwalder-Schrader axioms or its some kind of modifications makes it even harder. It has been shown by Jaffe in [27], that reflection positivity is not satisfied in the finite time non-equilibrium state of the solution, even for free scalar field theory. For the gauge theories, as we talked about in last chapter, we are not sure about how Wightman axioms should be modified to contain gauge theories, in particular we have negative results in the case of free electro-magnetic field.

7.2 The Model

The Abelian Higgs model is described by following classical Lagrangian density in Minkowski space-time

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi^\dagger \Phi)$$

where Φ is a complex scalar field, $F_{\mu\nu}$ is the curvature of a gauge field A_μ, and

$$V(\Phi^\dagger \Phi) = -m^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

$$D_\mu \Phi = \partial_\mu \Phi - ie A_\mu \Phi$$

note that when $\lambda = 0$, this model is also called the scalar quantum electrodynamics, the paper [41] studied the case $\lambda = m = 0$. The Lagrangian density is invariant under following gauge transformation

$$A \to A + \nabla f, \Phi \to e^{ief} \Phi$$

for any $f \in C^1$. Note that the first transform is just adding an exact differential df to the connection one-form $A_\mu dx^\mu$, which preserves the curvature. The second transformation is just a multiplication of one-form. The gauge covariant derivative transforms like

$$D_\mu \Phi \to \partial_\mu (e^{ief} \Phi) - ie(A_\mu + \partial_\mu f) e^{ief} \Phi = e^{ief} D_\mu \Phi$$

thus it is clear that the Lagrangian density is invaraint under these transformations.

To compute the Euclidean action, note that according to the Wick rotation, we have

$$x_0 \to -i x_0^E, \partial_0 \to i \partial_0^E, A_0 \to i A_0^E, F_{0\mu} \to i F_{0\mu}^E, D_0 \Phi \to i D_0^E \Phi$$

where the rule is keep the relations $\partial_\mu^E x^\nu = \delta_\mu^\nu$ and $A_\mu^E dx_E^\mu$ is a Euclidean one-form, now combine this with the Minkowski metric, we have

$$F_{\mu\nu} F^{\mu\nu} \to -F_{\mu\nu}^E F_E^{\mu\nu}, d^4 x \to i d^4 x_E$$

For simplicity we drop the subscript or superscript E. Thus we turn the action in Minkowski space time (multiply an i)

$$iS = i\int d^4x\Big(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\Phi)^\dagger D^\mu\Phi - V(\Phi^\dagger\Phi)\Big)$$

into the Euclidean action

$$iS \to -S_E = -\int d^4x\Big(\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\Phi)^\dagger D^\mu\Phi - V(\Phi^\dagger\Phi)\Big)$$

where the gauge covariant derivative and the electro-magnetic field tensor are still given by

$$D_\mu\Phi = \partial_\mu\Phi - ieA_\mu\Phi, F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

In the following we assume the Euclidean dimension is 2 and we assume the $V(\Phi^\dagger\Phi) = 0$ for simplicity.

Then we can write down the Langevin equations for the random field (A,Φ)

$$\begin{aligned}
\partial_t A_1 &= \partial_2^2 A_1 - \partial_1\partial_2 A_2 - \frac{ie}{2}[\bar{\Phi}\partial_1\Phi - \Phi\partial_1\bar{\Phi}] - e^2 A_1\Phi\bar{\Phi} + \xi_1\\
\partial_t A_2 &= \partial_1^2 A_2 - \partial_1\partial_2 A_1 - \frac{ie}{2}[\bar{\Phi}\partial_2\Phi - \Phi\partial_2\Phi] - e^2 A_2\Phi\,\bar{\Phi} + \xi_2\\
\partial_t\Phi &= \partial_1^2\,\Phi + \partial_2^2\,\Phi - ie\,(\partial_1 A_1 + \partial_2 A_2)\Phi - 2ie\,(A_1\partial_1 + A_2\partial_2)\Phi\\
&\quad - e^2(A_1^2 + A_2^2)\Phi + \zeta
\end{aligned}$$

where ξ_1, ξ_2 are real white noise, ζ is a complex white noise, that is they satisfy

$$\begin{aligned}
&\mathbb{E}\,[\xi_\alpha(t_1,x_1)\xi_\beta(t_2,x_2)] = \delta_{\alpha\beta}\,\delta(x_1 - x_2)\delta(t_1 - t_2)\\
&\mathbb{E}\,[\bar{\zeta}(t_1,x_1)\zeta(t_2,x_2)] = 2\delta(x_1 - x_2)\delta(t_1 - t_2)\\
&\mathbb{E}\,[\xi(t_1,x_1)\;\zeta(t_2,x_2)] = 0
\end{aligned}$$

Note that the equations for the gauge fields are not parabolic equations, our intention is to make them into parabolic ones. This is achieved by stochastic gauge fixing in physics literature, see [14], where they worked in momentum space. We first solve the equations

$$\begin{aligned}
\partial_t B_1 &= (\partial_1^2 + \partial_2^2)B_1 - \frac{ie}{2}[\bar{\psi}\partial_1\psi - \psi\partial_1\bar{\psi}] - e^2 B_1\psi\bar{\psi} + \xi_1\\
\partial_t B_2 &= (\partial_1^2 + \partial_2^2)B_2 - \frac{ie}{2}[\bar{\psi}\partial_2\psi - \psi\partial_2\bar{\psi}] - e^2 B_2\psi\bar{\psi} + \xi_2\\
\partial_t\psi &= (\partial_1^2 + \partial_2^2)\psi - 2ie(B_1\partial_1 + B_2\partial_2)\psi - e^2(B_1^2 + B_2^2)\psi\\
&\quad + e^{ie\int_0^t(\partial_1 B_1 + \partial_2 B_2)ds}\zeta
\end{aligned}$$

and then transform them into original variables

$$\begin{aligned}
A_1 &= B_1 - \int_0^t \partial_1(\partial_1 B_1 + \partial_2 B_2)ds\\
A_2 &= B_2 - \int_0^t \partial_2\,(\partial_1 B_1 + \partial_2 B_2)ds\\
\Phi &= e^{-ie\int_0^t(\partial_1 B_1 + \partial_2 B_2)}\,\psi
\end{aligned}$$

where one can formally check this is indeed a solution of original equations. Note that $e^{ie\int_0^t(\partial_1 B_1+\partial_2 B_2)ds}\zeta$ has the same law of a space-time white noise. So we arrive the equations

$$\begin{aligned}
\partial_t B_1 &= (\partial_1^2+\partial_2^2)B_1-\frac{ie}{2}[\bar{\psi}\partial_1\psi-\psi\partial_1\bar{\psi}]-e^2B_1\psi\bar{\psi}+\xi_1\\
\partial_t B_2 &= (\partial_1^2+\partial_2^2)B_2-\frac{ie}{2}[\bar{\psi}\partial_2\psi-\psi\partial_2\bar{\psi}]-e^2B_2\psi\bar{\psi}+\xi_2\\
\partial_t\psi &= (\partial_1^2+\partial_2^2)\psi-2ie(B_1\partial_1+B_2\partial_2)\psi-e^2(B_1^2+B_2^2)\psi+\zeta
\end{aligned}\tag{7.1}$$

In the paper [41], the author showed how these transformation works in the case of lattice approxiations of gauge theory, where one has one SDEs instead SPDEs. We did not find a way to show the stochastic gauge fixing works rigorously without lattice gauge theory, so this is a problem.

Assuming this can be done, we try to see where is the hard part of equations 7.1. We try to do the Da Prato trick: We set $B_1=\widetilde{B}_1+Z_1, B_2=\widetilde{B}_2+Z_2, \Phi=\phi+Y$ where

$$(\partial_t-\Delta)(Z_1,Z_2,Y)=(\xi_1,\xi_2,\zeta)$$

with $\xi_1,\xi_2,\zeta\in C^{-\frac{2}{2}-1-\varepsilon}=\mathcal{C}^{-2-\varepsilon}$ and $Z_1,Z_2,Y\in\mathcal{C}^{-\varepsilon}$ then

$$\begin{aligned}
\partial_t\widetilde{B}_1 &= \Delta\tilde{B}_1-\frac{ie}{2}\Big[\boxed{\bar{\phi}\partial_1\phi-\phi\partial_1\bar{\phi}}+\bar{Y}\partial_1Y-Y\partial_1\bar{Y}\\
&\quad+\boxed{\bar{\phi}\partial_1Y+\bar{Y}\partial_1\phi-Y\partial_1\bar{\phi}-\phi\partial_1\bar{Y}}\Big]\\
&\quad-e^2(\tilde{B}_1\phi\bar{\phi}+\tilde{B}_1\phi\bar{Y}+\tilde{B}_1\bar{\phi}Y+\tilde{B}_1Y\bar{Y}+\bar{Y}Z_1+\phi Z_1\bar{Y}++\bar{\phi}Z_1Y+Z_1Y\bar{Y})\\
\partial_t\phi &= \Delta\phi-\boxed{2ie(\tilde{B}_1\partial_1+\tilde{B}_2\partial_2)\phi-2ie(Z_1\partial_1+Z_2\partial_2)\phi}\\
&\quad-2ie\boxed{(\tilde{B}_1\partial_1+\tilde{B}_2\partial_2)Y}-2ie(Z_1\partial_1+Z_2\partial_2)Y\\
&\quad-e^2(\tilde{B}_1^2+\tilde{B}_2^2+Z_1^2+Z_2^2+2\tilde{B}_1Z_1+2\tilde{B}_2Z_2)\phi\\
&\quad-e^2(\tilde{B}_1^2+\tilde{B}_2^2+Z_1^2+Z_2^2+2\widetilde{B}_1Z_1+2\tilde{B}_2Z_2)Y
\end{aligned}$$

where we expect $\widetilde{B}_1,\widetilde{B}_2,\phi\in\mathcal{C}^{1-2\varepsilon}$. The terms in block are not well-defined. We arrived at here, and did not find a way to solve these SPDEs.

Appendix A Notation

$a \lesssim b$ for variables a, b: if there exists some positive constant c which is independent of the variables under our consideration, such that $a \leqslant c \cdot b$.

$a \simeq b$ for variables a, b: if $a \lesssim b$ and $b \lesssim a$.

$i \lesssim j$ for indices in Littlewood-Paley decomposition: if there is some integer $N \in \mathbb{N}$ independent of the the indices under our consideration, such that $i \leqslant j + N$.

$i \sim j$ for indices in Littlewood-Paley decomposition: if $i \lesssim j$ and $j \lesssim i$.

Appendix B Inequalities

In this appendix, we collect some useful inequalities.

Lemma B.1. (Hölder's inequality) *Suppose* (X,μ) *is a measure space. Given* $1\leqslant p,q,r\leqslant\infty$ *with* $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$, *for any functions* $f\in L^p(X,\mu)$, $g\in L^q(X,\mu)$, *we have* $fg\in L^r(X,\mu)$ *and*

$$\|fg\|_{L^r}\leqslant\|f\|_{L^p}\cdot\|g\|_{L^q}$$

Proof. When $p=1$ or $p=\infty$, the inequality is trivial, so we assume that $1<p<\infty$. Without loss of generality, let us assume $\|f\|_{L^p}=\|g\|_{L^q}=1$. Since the logarithm function is concave, then for any $a,b>0$ and $t\in[0,1]$ we have

$$t\log(a)+(1-t)\log(b)\leqslant\log(ta+(1-t)b)$$

thus

$$a^tb^{1-t}\leqslant ta+(1-t)b$$

Then

$$\begin{aligned}\int_X|fg|^r\,d\mu &= \int_X(|f|^p)^{\frac{r}{p}}(|g|^q)^{\frac{r}{q}}d\mu\\ &\leqslant \frac{r}{p}\int_X|f|^pd\mu+\frac{r}{q}\int_X|g|^qd\mu\\ &= \frac{r}{p}+\frac{r}{q}\\ &= 1\end{aligned}$$

So the result follows. □

The proof can be easily generalized to the case of any number of functions, so we have the following generalization.

Lemma B.2. *Suppose* (X,μ) *is a measure space. Given* $1\leqslant p_1,\ldots,p_n,r\leqslant\infty$ *with* $\frac{1}{p_1}+\ldots+\frac{1}{p_n}=\frac{1}{r}$, *for any functions* $f_i\in L^{p_i}(X,\mu)$ *where* $i\in\{1,\ldots,n\}$, *we have* $f_1\cdots f_n\in L^r(X,\mu)$ *and*

$$\|f_1\cdots f_n\|_{L^r}\leqslant\|f_1\|_{L^{p_1}}\cdots\|f_n\|_{L^{p_n}}$$

Lemma B.3. (Young's inequality for convolutions) *Given* $1\leqslant p,q,r\leqslant\infty$ *with* $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}$, *for any functions* $f\in L^p$, $g\in L^q$ *we have*

$$\|f*g\|_{L^r}\leqslant\|f\|_{L^p}\cdot\|g\|_{L^q}$$

Proof. The case of $r=\infty$ is easy, let us consider the case $r<\infty$. From the identity $\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}$ we see that $q\leqslant r$ and $p\leqslant r$, thus we can define the constants $\alpha=1-\frac{p}{r}$ and $\beta=1-\frac{q}{r}$ which are all in the interval $[0,1]$, and define the constants $p'=\frac{p}{\alpha}\geqslant1$ and $q'=\frac{q}{\beta}\geqslant1$. Clearly, we have

$$\frac{1}{p'}+\frac{1}{q'}+\frac{1}{r}=\frac{\alpha}{p}+\frac{\beta}{q}+\frac{1}{r}=\left(\frac{1}{p}-\frac{1}{r}\right)+\left(\frac{1}{q}-\frac{1}{r}\right)+\frac{1}{r}=1$$

So by using Hölder's inequality, we have

$$\begin{aligned}
|f * g(x)| &= \left|\int_{\mathbb{R}^n} f(x-y)\cdot g(y)dy\right| \\
&\leqslant \int_{\mathbb{R}^n} |f(x-y)|^{(1-\alpha)}\cdot|g(y)|^{(1-\beta)}\cdot|f(x-y)|^{\alpha}\cdot|g(y)|^{\beta}dy \\
&\leqslant \left(\int_{\mathbb{R}^n} |f(x-y)|^{(1-\alpha)r}\cdot|g(y)|^{(1-\beta)r}dy\right)^{\frac{1}{r}}\cdot \\
&\quad \left(\int_{\mathbb{R}^n} |f(x-y)|^{\alpha p'}dy\right)^{\frac{1}{p'}}\cdot\left(\int_{\mathbb{R}^n} |g(y)|^{\beta q'}dy\right)^{\frac{1}{q'}} \\
&= \left(\int_{\mathbb{R}^n} |f(x-y)|^{p}\cdot|g(y)|^{q}dy\right)^{\frac{1}{r}}\cdot\|f\|_{L^p}^{\alpha}\cdot\|g\|_{L^q}^{\beta}
\end{aligned}$$

Thus

$$\begin{aligned}
\int_{\mathbb{R}^n} |f*g(x)|^r dx &\leqslant \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)|^{p}\cdot|g(y)|^{q}dy\right)dx\cdot\|f\|_{L^p}^{\alpha r}\cdot\|g\|_{L^q}^{\beta r} \\
&= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)|^{p}\cdot|g(y)|^{q}dx\right)dy\cdot\|f\|_{L^p}^{\alpha r}\cdot\|g\|_{L^q}^{\beta r} \\
&= \|f\|_{L^p}^{p}\cdot\|g\|_{L^q}^{q}\cdot\|f\|_{L^p}^{\alpha r}\cdot\|g\|_{L^q}^{\beta r} \\
&= \|f\|_{L^p}^{r}\cdot\|g\|_{L^q}^{r}
\end{aligned}$$

Hence the result follows.

□

Appendix C Functional Analysis

C.1 Function Spaces

In this appendix, we introduce some function spaces.

Definition C.1. *For non-negative integer $k \in \mathbb{N}$, the space $C^k(\mathbb{R}^n)$ is the set of all k-times differentiable function $f: \mathbb{R}^n \to \mathbb{R}$ with continuous k-th derivative, such that*

$$\|f\|_{C^k} := \sum_{0 \leqslant |\alpha| \leqslant k} \sup_{x \in \mathbb{R}^n} |\partial^\alpha f(x)|$$

is finite.

Lemma C.2. *The space $C^k(\mathbb{R}^n)$ with norm $\|\cdot\|_{C^k}$ is a Banach space.*

The concept of Hölder continuous function with exponent $\alpha \in (0, 1]$ generalizes the concept of Lipschitz continuous function and gives a way to characterizes the regularity of continuous functions.

Definition C.3. *A continuous function $f: \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^n$ is an open subset, is called Hölder continuous with exponent $\alpha \in (0, 1]$ if there is some constant $C > 0$, such that*

$$|f(x) - f(y)| \leqslant C\|x - y\|^\alpha$$

holds for any $x, y \in \Omega$. In other words, the quantity

$$\sup_{x \neq y \in \Omega} \frac{|f(x) - f(y)|}{\|x - y\|^\alpha}$$

is finite.

Definition C.4. *For the non-negative integer $k \in \mathbb{N}$ and real number $\alpha \in (0, 1]$, the Hölder space $C^{k,\alpha}(\mathbb{R}^n)$ is defined to be the collection of C^k-functions whose k-th derivatives are Hölder continuous with exponent α such that the norm*

$$\|f\|_{C^{k,\alpha}} := \sum_{0 \leqslant |\beta| \leqslant k} \sup_{x \in \mathbb{R}^n} |\partial^\beta f(x)| + \sum_{|\beta| = k} \sup_{x \neq y \in \mathbb{R}^n} \frac{|\partial^\beta f(x) - \partial^\beta f(y)|}{\|x - y\|^\alpha}.$$

is finite.

Lemma C.5. *The space $C^{k,\alpha}(\mathbb{R}^n)$ with norm $\|\cdot\|_{C^{k,\alpha}}$ is a Banach space.*

C.2 Stone's Theorem

Stone's theorem deals with the strong continuous unitary representation of the abelian group $\mathbb{R}$, more precisely, it is a map $\mathbb{R} \to U(\mathcal{H})$ where $\mathcal{H}$ is a complex separable Hilbert space, such that

$$U(t_1)U(t_2) = U(t_1 + t_2) \text{ for any } t_1, t_2 \in \mathbb{R}$$

and $\mathbb{R} \to \mathcal{H}$ defined by $t \longmapsto U(t)\Phi$ is a continuous map for any fixed $\Phi \in \mathcal{H}$. Notice that in the case of complex separable Hilbert space, the strong continuity condition is equivalent to weak continuity, which means the map $\mathbb{R} \to \mathbb{R}$ defined by $t \mapsto \langle \Psi, U(t)\Phi \rangle$ is continuous for any fixed $\Psi, \Phi \in \mathcal{H}$.

Theorem C.6. (Stone) *Suppose* $\mathbb{R} \to U(\mathcal{H})$ *is a strong continuous unitary representation of the abelian group* $\mathbb{R}$*, then there exists a self-adjoint operator* H*, defined on a dense domain* D*, such that*

$$U(t) = e^{-itH}$$

on D*. The domain* D *is defined by*

$$\left\{ \Phi \in \mathcal{H} | \lim_{\varepsilon \to 0} \frac{U(\varepsilon)\Phi - \Phi}{-i\varepsilon} \text{ exists} \right\}.$$

Conversely, if there is a self-adjoint operator H*, defined on a dense domain* D*, then the map* $t \mapsto e^{-itH}$ *can be extended to a unique strong continuous unitary representation of the abelian group* $\mathbb{R}$*.*

C.3 Nuclear Theorem

The nuclear theorem says a continuous bilinear functional can be uniquely extended to a tempered distribution on the product space of underlying variables.

Theorem C.7. (Nuclear Theorem of Schwartz) *Suppose* B *is a continuous bilinear functional on* $\mathcal{S}(\mathbb{R}^m) \times \mathcal{S}(\mathbb{R}^n)$*, then there is a unique tempered distribution* $\varphi \in \mathcal{S}'(\mathbb{R}^{m+n})$ *such that*

$$B(f, g) = \varphi(fg)$$

for any $f \in \mathcal{S}(\mathbb{R}^m)$ *and* $g \in \mathcal{S}(\mathbb{R}^n)$*.*

It is clear that is theorem is also true for case of continuous multilinear functional. The nuclear theorem of Schwartz is also called the Schwartz kernel theorem, see page 61 of [10] and references therein.

Bibliography

[1] Helmut Abels. *Pseudodifferential and singular integral operators*. De Gruyter, 2011.

[2] Yakir Aharonov and David Bohm. Significance of electromagnetic potentials in the quantum theory. *Physical Review*, 115(3):485, 1959.

[3] Ioannis E Antoniou and Ilya Prigogine. Intrinsic irreversibility and integrability of dynamics. *Physica A: Statistical Mechanics and its Applications*, 192(3):443–464, 1993.

[4] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343. Springer Science & Business Media, 2011.

[5] Ismaël Bailleul and Frédéric Bernicot. High order paracontrolled calculus. In *Forum of Mathematics, Sigma*, volume 7. Cambridge University Press, 2019.

[6] Valentine Bargmann. Irreducible unitary representations of the lorentz group. *Annals of Mathematics*, pages 568–640, 1947.

[7] Valentine Bargmann. On unitary ray representations of continuous groups. *Annals of Mathematics*, pages 1–46, 1954.

[8] Valentine Bargmann. Note on wigner's theorem on symmetry operations. *Journal of Mathematical Physics*, 5(7):862–868, 1964.

[9] Nikolai Nikolaevich Bogolubov, AA Logunov, and IT Todorov. Introduction to axiomatic quantum field theory. *Reading, Mass*, 1975.

[10] Nikolai Nikolaevich Bogolubov, Anatoly A Logunov, AI Oksak, and I Todorov. *General principles of quantum field theory*, volume 10. Springer Science & Business Media, 2012.

[11] Niels Henrik David Bohr and L. Rosenfeld. Zur frage der messbarkeit der elektromagnetischen feldgrössen. 1933.

[12] Niels Bohr and Leon Rosenfeld. Field and charge measurements in quantum electrodynamics. *Physical Review*, 78(6):794, 1950.

[13] BGG Chen, David Derbes, David Griffiths, Brian Hill, Richard Sohn, and Yuan-Sen Ting. Lectures of sidney coleman on quantum field theory. 2019.

[14] Poul H Damgaard and Helmuth Hüffel. Stochastic quantization. *Physics Reports*, 152(5-6):227–398, 1987.

[15] Cécile DeWitt-Morette, Margaret Dillard-Bleick, and Yvonne Choquet-Bruhat. *Analysis, manifolds and physics*. North-Holland, 1978.

[16] J-P Eckmann and Henri Epstein. Time-ordered products and schwinger functions. *Communications in Mathematical Physics*, 64(2):95–130, 1979.

[17] Franco Ferrari and Helmuth Hüffel. On the stochastic quantization of the chern-simons theory. *Physics Letters B*, 261(1-2):47–50, 1991.

[18] R Ferrari, LE Picasso, and F Strocchi. Some remarks on local operators in quantum electrodynamics. *Communications in Mathematical Physics*, 35(1):25–38, 1974.

[19] Jürg Fröhlich. Verification of axioms for euclidean and relativistic fields and haag's theorem in a class of $P(\backslash varphi)$_2-models. In *Annales de l'IHP Physique théorique*, volume 21, pages 271–317. 1974.

[20] Izrail'Moiseevich Gel'fand. *Generalized Functions, 5 volumes*. Academic Press, New York, 1964 - 1968.

[21] James Glimm and Arthur Jaffe. *Quantum physics: a functional integral point of view*. Springer Science & Business Media, 2012.

[22] James Glimm and Min Chul Lee. Axioms for quantum gauge fields. *ArXiv preprint arXiv:2112.08575*, 2021.

[23] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular pdes. In *Forum of Mathematics, Pi*, volume 3. Cambridge University Press, 2015.

[24] Massimiliano Gubinelli and Nicolas Perkowski. Lectures on singular stochastic pdes. *ArXiv preprint arXiv:1502.00157*, 2015.

[25] Martin Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, 2014.

[26] Werner Heisenberg. Über energieschwankungen in einem strahlungsfeld. In *Original Scientific Papers/Wissenschaftliche Originalarbeiten*, pages 116–122. Springer, 1989.

[27] Arthur Jaffe. Stochastic quantization, reflection positivity, and quantum fields. *Journal of Statistical Physics*, 161(1):1–15, 2015.

[28] Arthur Jaffe. Reflection positivity then and now. *ArXiv preprint arXiv:1802.07880*, 2018.

[29] Arthur M Jaffee. High-energy behavior in quantum field theory. i. strictly localizable fields. *Physical Review*, 158(5):1454, 1967.

[30] Svante Janson et al. *Gaussian hilbert spaces*. Number 129. Cambridge university press, 1997.

[31] Leonid Koralov and Yakov G Sinai. *Theory of probability and random processes*. Springer Science & Business Media, 2007.

[32] Edward Nelson. Derivation of the schrödinger equation from newtonian mechanics. *Physical review*, 150(4):1079, 1966.

[33] Edward Nelson. *Dynamical theories of Brownian motion*, volume 106. Princeton university press, 2020.

[34] Edward Nelson. Quantum fluctuations. In *Quantum Fluctuations*. Princeton University Press, 2021.

[35] Konrad Osterwalder and Robert Schrader. Axioms for euclidean green's functions. *Communications in mathematical physics*, 31(2):83–112, 1973.

[36] Konrad Osterwalder and Robert Schrader. Axioms for euclidean green's functions ii. *Communications in Mathematical Physics*, 42(3):281–305, 1975.

[37] Georgio Parisi, Yong Shi Wu et al. Perturbation theory without gauge fixing. *Sci. Sin*, 24(4):483–496, 1981.

[38] Nicolas Perkowski. Spdes, classical and new. *Dimension*, 2:105, 2020.

[39] Michael Reed. *Methods of modern mathematical physics: Functional analysis*. Elsevier, 2012.

[40] JE Roberts. The dirac bra and ket formalism. *Journal of Mathematical Physics*, 7(6):1097–1104, 1966.

[41] Hao Shen. Stochastic quantization of an abelian gauge theory. *Communications in Mathematical Physics*, 384(3):1445–1512, 2021.

[42] Norbert Straumann. Unitary representations of the inhomogeneous lorentz group and their significance in quantum physics. In *Springer Handbook of Spacetime*, pages 265–278. Springer, 2014.

[43] Ray F Streater. Outline of axiomatic relativistic quantum field theory. *Reports on Progress in Physics*, 38(7):771, 1975.

[44] Raymond Frederick Streater and Arthur S Wightman. *PCT, spin and statistics, and all that*, volume 52. Princeton University Press, 2000.

[45] Steven Weinberg. *The quantum theory of fields: Foundations*. Cambridge University Press, 2002.

[46] John Archibald Wheeler and Wojciech Hubert Zurek. *Quantum theory and measurement*, volume 15. Princeton University Press, 2014.

[47] Arthur S Wightman. La théorie quantique locale et la théorie quantique des champs. In *Annales de l'IHP Physique théorique*, volume 1, pages 403–420. 1964.

[48] Arthur S Wightman. How it was learned that quantized fields are operator-valued distributions. *Fortschritte der Physik/Progress of Physics*, 44(2):143–178, 1996.

[49] Arthur S Wightman and L Garding. Fields as operator-valued distributions in relativistic quantum theory. *Arkiv Fys.*, 28, 1965.

[50] Eugene Wigner. On unitary representations of the inhomogeneous lorentz group. *Annals of mathematics*, pages 149–204, 1939.

[51] Eugene Paul Wigner. Gruppentheorie und ihre anwendung auf die quantenmechanik der atomspektren. 1931.

[52] Tai Tsun Wu and Chen Ning Yang. Concept of nonintegrable phase factors and global formulation of gauge fields. *Physical Review D*, 12(12):3845, 1975.