"The Analytic Continuations of Distributions"

Dissertation

Submitted to the School of Mathematics of

The University of Nairobi

(Division of Pure Mathematics)

by

ONGARO, JARED NYANG'AU

In partial fulfillment of the requirements

for the Degree of

Master of Science in Pure Mathematics.

Declaration

I the undersigned declare that this dissertation is my original work and to the best of my knowledge has not been presented for the award of a degree in any other University.

ONGARO JARED NYANG'AU

Reg. No. I56/71571/2007

.... 18/08/09 ....

Signature Date

Declaration by Supervisors

This dissertation has been submitted for examination with my approval as a supervisor

CLAUDIO ACHOLA

.... 20/08/09 ....

Signature Date

JAMES NKUUBI

.... 24/08/09 ....

Signature Date

To my girlfriend; Lilian Kerubo

who has adjusted extremely well to having a mathematician as boyfriend.

On Analytic Continuations of Distributions

ONGARO JARED NYANG'AU

August 7, 2009

Abstract

In this dissertation, we develop the elements necessary at providing the theoretical formulation of the solution to the Gel'fand(1954) problem. All the results are well-known and our contribution is only at the level of the presentation.

Suppose K is a field of characteristic 0; for instance, take $K = \mathbb{R}$. Given $p(x) \in \mathbb{R}[x_1, \cdots, x_n]$, $\lambda \in \mathbb{C}$, $|p(x)|^\lambda$ is well defined for $Re(\lambda) \geq 0$ on $\mathbb{R}^n$. Can we analytically continue $|p(x)|$ to the entire complex plane? This is the problem of I. M. Gel'fand, posed at the International Congress of Mathematicians, Amsterdam 1954.

In 1968, Bernstein and S. I. Gel'fand [7] and Atiyah [11], independently provided a solution with proofs of this fact based on Hironaka's theorem about resolution of singularities, a very deep and difficult result. However, in 1972, Bernstein produced a beautiful and a completely algebraic proof of the result. The idea of the proof is; to analytically continue $|p(x)|$ we can use Bernstein-Sato polynomial $b(\lambda)$, which has the property that

$$b(\lambda)p(x)^{\lambda-1} = (\text{some polynomial differential operator})p(x)^{\lambda}.$$

We pick up poles from the zeros of $b(\lambda)$ (the poles of the analytic continuation will be at integer shifts of the zeros of $b(\lambda)$, since we have to repeat the process). To compute this polynomial is an involving process as seen in example 3.3.8. Fortunately, we do not always need to know $b(\lambda)$ explicitly. Bernstein showed how to prove that it exists, without actually producing it.

The problem is best considered in the context of distributions. Thus, in this dissertation the basic concepts of distribution are used and the modern theory of distribution as introduced by Laurent Schwartz is explained. Some powerful algebraic machinery: D-module theory is then introduced. Finally, we shown how a problem of analysis with complicated solution can be solved through algebraic approach with relative ease.

Table of Contents

Acknowledgements

I am proud to have benefited from the patience, quiet wisdom of my core supervisor here at Nairobi, Claudio Achola. His encyclopedic knowledge, coupled with his considerable shrewdness and topped by his understanding and kindness, make him a role model both professionally and personally. Many thanks to Claudio (again) was my research mentor at the University of Nairobi; like all excellent teachers, not only his knowledge but his tastes proved contagious, and I am a better mathematician for it. He was often able to explain things from a point of view which rendered a difficult problem much clearer. This was a great help when trying to wade through a junk of well written but at times very complicated texts on the subject. I can't describe what he has given me, I will always admire Claudio, for his mathematical mind, but aside from that, for his generosity, kindness and his way of life. I have never learn't more on anything from anybody else, and I doubt I ever will.

My second supervisor, lecturer and a friend James Nkuubi: James never stopped encouraging me in this work. Not only his commitment in mathematics, but, flexibility and availability makes him a person I couldn't avoid working with at this level. Thanks! James, through this work I have learned to appreciate the power and beauty of D-modules.

Many Thanks to Professor Rikard Bögvad for suggesting the topic and structure of the dissertation when we met,(April '09) during the EAUMP school on Set theory and Logic in Nairobi and making sure none of my mails went unreplied. He showed continued interest in my work and has given me valuable suggestions for corrections and improvements. I can only hope I didn't disappoint!. So to the entire Algebraic analysis group in Nairobi-Stockholm Claudio and Rikard, thank you!. To Madam Anne who will have been my third supervisor would it have been not what it was?! The tea, peanuts and hot dogs at the Chiromo club was refreshing and proved to be very important due to discussion which we engaged in with the senior members of the club: will truly miss the tea. People like Dr. Patrick Weke, could have said more if I was particularly more organised like him. Patrick, that is the way it should be.

My work become easier after I acquired a grant to buy a Laptop from African Millennium Mathematics Initiative(AMMSI), special thanks to the Director Professor Wandera Ogana. I wish to thank the School of Mathematics and the entire University of Nairobi, particularly, thanks to Professor John Okoth Owino for Initiating unfamiliar programme which I really benefited from. Things were not that easy after his departure as director of the school, but we were lucky that his passion for mathematics, made him rest our case to then incoming director; who renewed our scholarship without much ado. I acknowledge you; Dr. Hudson Were. To my opinion, the Division of Pure mathematics was organised thank to the leadership of Professor Khalagai. I should also point out that the much analysis you will encounter in this dissertation credit to him; He always taught expressing a lot of mathematical understanding of the underlying concepts.

Perhaps this was where to start my acknowledgments; my family, who are responsible for so much of who I am today. I am grateful to them for all of their love and support, and for only occasionally asking "what is all this good for anyway?". I love you all immensely. I could never have completed this dissertation without help in both mathematics and support from my greatest inspiration of all times ever since we joined university; my dearest friend Benjamin Kikwai, who helped inspire me. To all my other classmates Edwin, Presly, Dennis, Faith and Ester who put up with my frustrations and who continuously supported me; you are God sent. I would not have made it through the past two years without your encouragement, I thank you.

Finally, I would like to thank my high school mathematics teacher Peter Kirenge for getting me interested in mathematics to begin with. Mr. Kirenge, this is what that algebra became! Similarly, I would like to acknowledge the absolute important role of Professor G.P. Pokhariyal for his generous spirits and for having open enough minds. Pokhariyal, in response to my request to avoid doing any Geometry courses at master level replied: 'I'm not going to let you out of here with a pure mathematics degree without seeing some proper Geometry!'. I hope this dissertation shows that I have learnt the basics of geometry, and Pokhariyal will let me graduate. He is such a strict and an effective professor who makes sure every i was dotted, thank you.

ONGARO NYANG'AU

Nairobi-Kenya

August-2009.

Notation

These are some of the commonly used notation in our work.

$\mathbb{C}$: Field of Complex numbers

$\mathbb{R}$: Field of Real numbers

$\mathbb{N}$: set of Positive numbers

K: Field of characteristic 0

$K[X] = K[x_1, \cdots, x_n]$: Polynomial ring over K

$A_n(K)$: n^{th} Weyl algebra

$\mathcal{D} = C_0^\infty(\mathbb{R}^n)$: Space of all test functions

$\mathcal{D}'$: Space of distributions on $\mathbb{R}^n$

δ_{ij}: Kronecker delta

$\{B_j\}$: Bernstein Filtration

$X = (x_1, \cdots, x_n)$: Point in $\mathbb{R}^n$

Chapter 1

Introduction

1.1 Historic Remarks

D-module is a module over a ring D of differential operators. The major interest of such D-modules is as an approach to the theory of linear partial differential equations. Since around 1970, D-module theory has been built up, mainly as a response to the ideas of Mikio Sato on algebraic analysis and as an expansion on the work of Sato and Joseph Bernstein on the Bernstein-Sato polynomials. Therefore, there has been a vast amount of work on the theory of D-modules over the past 30 years, particularly due to the contributions of I.N. Bernstein, J. E. Björk, M. Kashiwara, T. Kawai, B. Malgrange, Z. Mebkhout, and others.

Question 1.1.1. *Why D-modules ?*

As Coutinho points out in his book [1], the question is particularly easy to answer. Hardly any area of Mathematics has been left untouched by this theory: from Number Theory to Mathematical Physics and from Singularity Theory to Representation of Algebraic Groups, to mention only a bunch. Indeed, the theory of D-modules sits across the traditional division into Algebra, Analysis and Geometry and this fact gives to the theory a rare beauty.

1.2 Motivation for D-Modules

Let K be a field of characteristic 0. For some fixed $n, K[X]$ denotes the polynomial ring in n variables. A_n is the sub algebra of $End_K K[X]$ generated by the operators

$$f \longrightarrow x_i f \quad \text{and} \quad f \longrightarrow \frac{\partial f}{\partial x_i}$$

which are denoted by x_i and ∂_i. We note that; $A_n[x_1, \cdots, x_n, \partial_1, \cdots, \partial_n]$.

Remark 1.2.1. *To make the notation less cluttered we will often drop the subscripts for the generators of A_1 and write them simply as x and ∂.*

The Weyl algebra is simple, i.e. the only two-sided ideals of A_n are 0 and A_n itself. However, A_n is very rich in one-sided ideals. We denote by $\langle P_1, \cdots, P_m \rangle$ the left ideal of A_n generated by the operators $P_1, \cdots, P_m \in A_n$.

The left ideal $I = \langle P_1, \cdots, P_m \rangle$ is exactly the object one would want to examine in order to study the system of linear partial differential equation(PDEs)

$$P_1 \cdot f = P_2 \cdot f = \cdots = P_m \cdot f = 0, \tag{1.1}$$

where f is the unknown function. In deed, the operators of I produce equations that follow algebraically from the equations 1.1. Now , we can quotient the ring A_n by these equations, to get a D-module.

Example 1.2.2. *To study the equation $\frac{d^3y}{dx^3} - \frac{dy}{dx} = 0$, we would take $M := A_1 / \langle \partial^3 - \partial \rangle$. while to study $(\frac{dy}{dx} - y)^2 = 0$ we would also take $M := A_1 / \langle (\partial - 1)^2 \rangle$.*

The solution of our differential equation is just a D-module homomorphism from M to a D-module S of functions. Note that while we are working with algebraic differential operators, space of functions S is a D-module. So we can study algebraic, analytic, or smooth solutions to our equations.

Example 1.2.3. *A single linear operator represented by a cyclic ideal $I_1 = \langle \partial^3 - \partial \rangle$ correspond to the space of solutions spanned by functions 1, e^x and e^{-x}. The solutions of ideal $I_2 = \langle (\partial - 1)^2 \rangle$ are spanned by e^x and xe^x.*

Remark 1.2.4. *Not all the ideals in the first Weyl algebra are cyclic. For example, the ideal* $I = \langle \partial^2, x\partial - 1 \rangle$ *cannot be generated by one element; In fact, according to Stafford: Every left ideal of the Weyl algebra can be generated by two elements.*

Clearly, K-linear space M is called a (left) D-module if there is a left action of A_n defined on M. When D-modules are employed to study systems of linear PDE, the first question one may ask is: what are the solutions? The concept of the solution space is readily available in the language of D-modules.

That is, to solve the system of equations represented by a cyclic D-module $M_1 = D/I$ in terms of the functional space represented by D-module M_2 means to describe the set of homomorphisms $Hom_D(M_1, M_2)$. This set has a structure of a K-linear space. Every element $\phi \in Hom_D(M_1, M_2)$ is defined by its value at the coset $\bar{1} \in D/I$; since $\bar{1}$ has to be annihilated by I, it is a solution to the corresponding system of PDEs.

Example 1.2.5. *Consider cyclic ideals* $I_1 = \langle \partial^3 - \partial \rangle$. *and ideal* $I_2 = \langle (\partial - 1)^2 \rangle$. *If* S *is a space of smooth functions, then*

$$\begin{aligned} Hom_D(D/I_1, S) &= Span_K \left\{1, e^x, e^{-x}\right\}, \\ Hom_D(D/I_2, S) &= Span_K \left\{e^x, xe^x.\right\} \end{aligned}$$

Therefore, if we look for solutions in the ring of one variable polynomials $K[x]$, *then*

$$\begin{aligned} Hom_D(D/I_1, K[x]) &\cong K, \\ Hom_D(D/I_2, K[x]) &\cong 0. \end{aligned}$$

1.3 Some Function Theory on Complex Domains.

Let Ω denote an open subset, called a region of the complex number space $\mathbb{C}^n$, if Ω is connected then Ω is called a domain. Since we have a topology on $\mathbb{C}$, we can talk about open sets. Just as in real analysis, there will be domains of differentiable functions.

Example 1.3.1. *Let* Ω *be a non-empty open set in* $\mathbb{C}$. *For instance,* Ω *would be* $\mathbb{C}$ *itself or* $\mathbb{C}$ *with with say finite number of points removed such as* $\mathbb{C}^\times$, *or the interior of the disk, or half plane.*

Definition 1.3.2 (Holomorphic functions). *They are functions defined on an open subset of complex number plane $\mathbb{C}$ with values in $\mathbb{C}$ that are complex differentiable at every point. This is a stronger condition than the real differentiability and implies that the function is infinitely often differentiable and can be described by Taylor Series.*

Remark 1.3.3.

The term analytic function is often used interchangeably with holomorphic function, although the term analytic is also used in broader sense of any function(real , complex or of more general type). The fact that the class of analytic functions coincides with the class of holomorphic functions, is a major theorem in complex analysis.

Definition 1.3.4 (meromorphic functions). *A meromorphic function on an open subset Ω of the complex plane is a function that is holomorphic on the whole of Ω, except a set of isolated points.*

1.4 Analytic Continuation of a function

This is one of the most important topic in complex variable theory as used today in many physical applications. The concept of analytic continuation just means enlarging the domain without giving up the property of being differentiable, i.e. holomorphic or meromorphic. After Weierstrass this process is called **analytic continuation**. All the elementary functions of real variables may be extended into the complex plane-replacing the real variable x by the complex variable z. For example, we could ask; is it possible to extend the definition of the sine function from the real line to all complex numbers? This is an example of analytic continuation. Namely, the same power series that defines the Taylor series for $\sin(x)$ with real x, also works for complex values of x.

In general, it is a big problem in complex function theory to continue functions analytically.

Example 1.4.1. *The Riemann zeta-function*

$$\zeta(z) = \sum_{n=1}^{\infty} n^{-z}$$

is easy to define for complex values with real part greater than 1, but it is quite difficult to find a continuation to the whole complex plane.

Question 1.4.2. *Why analytic continuation ?*

Moving into the complex plane opens up new opportunities for analysis. It defines the analytic functions in terms of its original definition and all its continuations. The power series method described in many coffee-table books in complex analysis e.g. see Ruel V. Churchill et al [18] is a common technique for doing it. However, we can also employ the integral techniques.

Generally one knows of an analytic function, such as the solution in series of a differential equation, where only the expansion around a point is usually found. However, one wants the general function which is guaranteed to exist by analytic continuation.

Chapter 2

Distributions Theory

In this chapter, we introduce the basic notion of distributions and their properties. Some of the uses are highlighted.

2.1 Introduction

It is true that we are used to describe the world using functions (especially physical processes in it). Sometimes, this view gives us the connection between several different quantities: position of an object, its speed and power applied to it is the value of a function and its first and second derivatives. So, if we know other derivatives, one more benefit is the possibility to extrapolate the function through Taylor series solution of differential equation. Every time we know something quite general about the whole function, and something precise only about its behaviour on a small set. In this manner, we can describe the function precisely everywhere by the use of derivatives.

Why we are not happy with functions?

Sometimes functions are not differentiable. If one thinks that nondifferentiable functions exist only in theory, it is not true. In applications one gets these as limits of differentiable functions, for example solving differential equation with a parameter. In many cases it is easier to consider an nondifferentiable function(or even noncontinuous function) as a critical "the most bad " case for differentiable function.

Example 2.1.1 (Dirac delta function). *The Dirac delta function on $\mathbb{R}$ is defined by the* *properties*

$$\delta(x) = \begin{cases} 0 & \textit{for} \quad x \neq 0, \\ \infty & \textit{for} \quad x = 0, \end{cases} \qquad \textit{and} \quad \int_{-\infty}^{\infty} \delta(x)dx = 1. \tag{2.1}$$

Let f be a nice function which vanishes at infinity, then,

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = \int_{-\infty}^{\infty} [f(x) - f(0)]\delta(x)dx + f(0)\int_{-\infty}^{\infty} \delta(x)dx = f(0)$$

because $[f(x) - f(0)]\delta(x) = 0$ on $\mathbb{R}$.

There is no such function δ with these properties and also we cannot interpret $\int_{-\infty}^{\infty} f(x)\delta(x)dx$ as an integral, in the usual sense. However, what does make sense is the assignment $f \longmapsto f(0)$. The δ function is thought of as a distribution.

2.2 What good are distributions?

Theory of distributions *also known as theory of generalized functions*(we do not distinguish between these terms) was invented to give a solid theoretical foundation to the delta function δ, which had been introduced as a technical device in mathematical formulation of quantum mechanics.

We now discuss some more concrete problems that will be solved in the calculus of distributions.

1. Derivative of discontinuous functions. One of the main achievements of 19th century analysis was to carefully examine notions such as continuity and differentiability, and to show that there are many continuous functions that are not differentiable.
Recall that f is differentiable at x, and its derivative is $f'(x) = L$, if the limit

$$\lim_{h \longrightarrow 0} \frac{f(x+h) - f(x)}{h} \tag{2.2}$$

exists and is equal to L.

While it was always clear that not every continuous function is differentiable, for example the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = |x|$ is not differentiable at 0, it was not until the work of Bolzano and Weierstrass that the full extent of the problem became clear: there are nowhere differentiable continuous functions. See Theorem 4.50 in [19].

Example 2.2.1 (The Dirichlet function). *The Dirichlet function is defined on $\mathbb{R}$ by*

$$f(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

every point $x \in \mathbb{R}$ is a point of discontinuity of f.

Question 2.2.2. *What makes us avoid differentiating discontinuous functions? Is it because it is impossible, unwise or simply out of ignorance?*

The good news is that, what we can call as unfair restriction poised by equation 2.2, does not prevent one from extending differentiation to continuous functions if one is willing to generalize the class in which the result (i.e. the derivative) will lie. As we will explain, a suitable notion of generalized functions is that of distributions. With this generalization, every continuous function, indeed every distribution, can be differentiated as many times as desired, with the result being yet another distribution. We will indeed generalise our concept of a function in such away that this new specification of a function will hold for all the continuous and piecewise continuous functions.

Example 2.2.3 (**Heaviside step-function**). *The Heaviside step-function is defined on $\mathbb{R}$ as,*

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{2.3}$$

then we see $H' = \delta$ by Theorem 2.4.13 in the following sense. If f is a nice function which vanishes at infinity, then integration by parts gives

$$
\begin{aligned}
\int_{-\infty}^{\infty} f(x)H'(x)dx &= [f(x)H(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f'(x)H(x)dx \\
&= -\int_{-\infty}^{\infty} f'(x)H(x)dx \\
&= -\int_{0}^{\infty} f'(x)dx \\
&= -[f(x)]_0^{\infty} \\
&= f(0) \\
&= \int_{-\infty}^{\infty} f(x)\delta(x)dx.
\end{aligned}
\tag{2.4}
$$

2. Partial Differential Equations:

Partial differential equations have been developed as a mathematical toolbox for the description of natural and engineered phenomena, ranging from vibrating strings to weather prediction to the design of the Boeing 777. Partial differential equations provide a language for modern science and analysis provides a theoretical background for their study. Consequently, differential equations are used to construct models of reality. Sometimes the reality we are modeling suggests that some solutions of differential equation need not be differentiable.

Let us first fix some definition about a solution to a partial differential equation.

Definition 2.2.4. *A partial differential equation for a real-valued function u in $n-$variables $x_1, \cdots, x_n$ is a relation of the form*

$$f(x_1, x_2, \cdots, u_{x_1}, \cdots, u_{x_1x_2}, \cdots) = 0 \tag{2.5}$$

where

$$u_{x_1} = \frac{\partial u}{\partial x_1}, \quad u_{x_1x_2} = \frac{\partial^2 u}{\partial x_1 x_2}, \cdots \quad e.t.c.$$

It may happen that this equation is supplemented by constraints on u and its partial derivatives of the boundary $\partial\Omega$ of the region $\Omega \subset R^n$ throughout which the independent variables $x_1, \cdots, x_n$ vary. These constraints are called **boundary conditions**, and if one of the variables is identified as time, the constraint associated with that variable is called an **initial or final condition**.

Definition 2.2.5 (Classical solution). *A solution of a partial differential equation is defined to be a function $u(x_1, ..., x_n)$ such that all the derivatives which appear in the partial differential equation* *exist, are continuous, and such that the equation, boundary conditions and initial conditions are all satisfied.*

It may be thought that with this definition of a solution the main emphasis should be focused on deriving methods by which a solution can be obtained. However, the story is not quite so simple. If we consider the wave equation (vibrating string equation)

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}; \quad u(0,t) = u(1,t) = 0; \quad u(x,0) = f, \quad \frac{\partial u}{\partial t}(x,0) = 0 \tag{2.6}$$

with u a function on $\mathbb{R}_x \times \mathbb{R}_t$. Then the general solution of this PDE, obtained by d'Alembert in the 18th century, must be of the form

$$u(x,t) = f(x+ct) + g(x-ct),$$

where f and g are arbitrary functions on $\mathbb{R}$. It is easy check by the chain rule that u solves the PDE as long as we can make sense of the differentiation. So, in the **classical sense,** f, g twice continuously differentiable, that is $f, g \in C^2(\mathbb{R})$ suffices.

Question 2.2.6. *But shouldn't this also work for rougher f, g?*

This question will be answered with the help of the following example.

Example 2.2.7.

Consider $f(x)$ given by

$$f(x) = \begin{cases} x & 0 \le x \le \frac{1}{2}, \\ 1-x & \frac{1}{2} \le x \le 1, \end{cases}$$

which is equivalent to plunking a string at the mid-point. Now the solution cannot be classical since $f(x)$ is not differentiable at $x = \frac{1}{2}$.

Question 2.2.8. *So what do we mean by a solution to the problem ?*

It is obvious, that we must widen the concept of a solution if we are going to allow initial conditions of the given form like in example 2.2. This restriction is both irksome and natural in many instances. It can be overcome by introducing so called weak solutions. By definition, these are functions u such that

$$\int u\left(\frac{\partial^2\phi}{\partial t^2} - c^2\frac{\partial^2\phi}{\partial x^2}\right) dxdt = 0 \tag{2.7}$$

for all sufficiently 'good' functions ϕ.

In general, idealization of physical problems often results in distributions. For instance, the sharp front for the wave equation discussed above, or point charges (the electron is supposed to be such!) are good examples. But a later curious application of great interest will be perhaps the applications of Dirac's Delta function in statistics where the δ-function approach provides us with a unified approach in treating discrete and continuous distributions; e.g Use Delta function to obtain discrete probability distributions see Santanu [15].

2.3 Formal Definition of Distributions

From the foregoing, we realize that most of the time when working with ordinary functions, we are subjected to what one might consider undesirable restrictions. In the case of differentiation, continuous functions need not be differentiable, and a function which is once differentiable need not be twice differentiable; similarly for Fourier integrals, all kinds of nice functions, $\cos x$ for instance does not possess Fourier transforms.

Distributions are important in physics and engineering where many noncontinuous problems naturally lead to differential equations whose solutions are distributions, such as the Dirac delta distribution. For example, Heaviside's mathematical innovations arising in the physics of telegraph cables (1880-7) were disregarded for many years. Only in the 1930s Hadamard, Sobolev, and others made systematic use of non-classical generalized functions. In 1952, Laurent Schwartz won a Fields Medal for systematic treatment of these ideas. In the next few sections we shall present the basic theory of these concepts according to the latter investigator.

The suitability of a particular function of a real variable for describing physical phenomena depends on its analytic properties such as continuity, differentiability, integrability, etc. These properties in turn depend on the structure or topology of the underlying real number system. For instance, to define the concept of continuity of a function of a real variable we must first investigate the convergence of sequences of real numbers. Analogously, the analytic properties of functional must be defined with respect to the convergence properties of sequences of functions $f(x)$ of the underlying function space. In other words, they must be defined with respect to the structure or topology of that function space.

Thus, the proper approach to the theory is via topological vector spaces-see Rudin's book [16] for the development along these lines. Here, the underlying point set will be an n-variable space $\mathbb{R}^n$ (point $x = (x_1, \cdots, x_n)$) or even a subset $\Omega \subset \mathbb{R}^n$ that is open.

Roughly speaking, the idea is; to deal with very 'bad' objects, first we need very 'good' ones. Many other names such as smooth, rapidly decreasing, e.t.c for test functions are common in the literature of distribution theory. Therefore, in order to give a generalised definition of a function, we must first introduce test (or testing) function.

Definition 2.3.1 (Support of a function). *The support of a function f on $\mathbb{R}^n$, denoted by the supp f is the closure of the set where f does not vanish;*

$$supp f = \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}.$$

A function has a **compact support** *if its support is a compact set; that is, there are numbers a, b such that $supp(f) \subset [a, b]$.*

Definition 2.3.2 (Test functions). *Let Ω be an open subset of $\mathbb{R}^n$. A test function on Ω is a function $\phi : \Omega \longrightarrow \mathbb{R}$ for which*

1. *supp ϕ is compact, and*

2. *ϕ has derivative of all orders.*

That is, $\phi(x)$ is continuously differentiable infinitely often, and vanishes identically outside some finite interval.

Example 2.3.3. *Consider the function*

$$f(x) = \begin{cases} 0 & for \quad x \leq 0 \\ e^{-1/x} & for \quad x > 0, \end{cases}$$

we claim it is infinitely differentiable. We can easily check that

$$f^{(n)}(x) = \begin{cases} 0 & for \quad x \leq 0 \\ P_n e^{-1/x} & for \quad x > 0 \end{cases}$$

(Here P_n means a rational function in x).

We use Greek letters $\phi, \varphi, \cdots$ etc. for test functions, and we denote the set of all test functions by $C_0^\infty(\Omega)$ or $\mathcal{D}(\Omega)$ the linear subspace of $C^\infty(\mathbb{R}^n)$ of those functions with compact support. If the underlying set Ω or $\mathbb{R}^n$ does not play a role in a particular definition, we simply refer to the set of test functions as $\mathcal{D}$.

Proposition 2.3.4. *All sums and products of test functions are again test functions. In particular, $\mathcal{D}$ is a vector space.*

Proof. This is immediately follows from the definition. □

Proposition 2.3.5. *All derivatives of test functions are test functions.*

Proof. This is also immediately follows from the definition. □

Proposition 2.3.6. *All Fourier transforms of test functions are again test functions. Therefore, the set $\mathcal{D}$ of test functions is well behaved.*

Thus, we can make a test function in $\mathcal{D}$ do anything a smooth function can do. We can force it take on prescribed values at a finite set of points.

Remark 2.3.7. *In writing down a formula for a function in $\mathcal{D}$ is often difficult, in fact no analytic function (other than $\phi \equiv 0$) can be in $\mathcal{D}$ because of the vanishing requirement. Thus, any formula of ϕ must be given 'in pieces'.*

Remark 2.3.8. *The space of test functions is not complete, since the limit of a sequence of test functions need not be a test function.*

2.4 Topology and Convergence of Test Functions

We study the topology on the space $C_0^\infty(\Omega)$ of arbitrary scalar valued smooth functions on an open subset Ω of $\mathbb{R}^n$, together with the associated space of distributions. To topologize $C_0^\infty(\Omega)$, we use the family of semi norms indexed by pairs (K, P) with K a compact subset of Ω and with P a polynomial, the $(K, P)^{th}$ semi norm being $\|f\|_{K,P} = sup_{x\in K} |(P(D)f(x))|$. The resulting topology is Hausdorff, and $C_0^\infty(\Omega)$ becomes a topological vector space. This topology is given by a countable subfamily of these semi norms and is therefore implemented by a metric. It is certainly sufficient to consider only the monomials D^α instead of all polynomials $P(D)$, and thus the P index of (K, P) can be assumed to run through a countable set.

Precisely, the collection of infinitely differentiable functions with compact support, $\mathcal{D}(\Omega)$, can be made a topological space, if we define convergence of functions in $\mathcal{D}(\Omega)$ as follows:

Definition 2.4.1. *A sequence of test functions $\{\phi_m\}_{m\in\mathbb{N}} \subseteq \mathcal{D}(\Omega)$ converges in ϕ in $\mathcal{D}(\Omega)$ if:*

1. *there is a compact $K \subset \Omega$ with supp $\phi_m \subset K$ for all n, and*
2. *all derivative $D^\alpha \phi_n$ converges uniformly to $D^\alpha \phi$.*

We have used the multi-index notation for partial derivatives, where

$$D^\alpha \phi(x) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \phi = \frac{\partial^{\alpha_1 + \cdots + \alpha_n}}{(\partial x_1)^{\alpha_1} \cdots (\partial x_n)^{\alpha_n}} \phi$$

the so called multi-index $\alpha = (\alpha_1, \cdots \alpha_n)$ is a n-tuple of non-negative integers. Of course, one defines $\alpha = (\alpha_1 + \cdots + \alpha_n)$.
This is a type of convergence of 'infinite order' since it implies uniform convergence for every derivative. Note that we do not demand that the derivatives of all orders shall simultaneously converge uniformly, but only that the derivatives of each order taken separately shall converge uniformly.

Proposition 2.4.2. *The space $\mathcal{D}(\mathbb{R}^n)$ is dense in $C_0(\mathbb{R}^n)$, the space of continuous real function of compact support with topology given by uniform convergence, i.e. for any $f \in C_0(\mathbb{R}^n)$ with supp$(f) \subset \Omega$ for some open $\Omega \subset \mathbb{R}^n$ and for any $\epsilon > 0$, there exists a $\varphi \in \mathcal{D}(\mathbb{R}^n)$ with supp $(f) \subset \Omega$, and such that for all $x \in \mathbb{R}^n$*

$$|f(x) - \varphi(x)| < \epsilon.$$

Recall that the space $\mathcal{D}(\Omega)$ of test functions is endowed with the norm, therefore we can define a functional;

Definition 2.4.3 (Functional). *A mapping $f : \mathcal{D} \longrightarrow \mathbb{R}$ is called a functional.*

A map $\mathcal{D} \longrightarrow \mathbb{R}$ means a rule which, given any $\phi \in \mathcal{D}$, produces a corresponding $z \in \mathbb{R}$; we write $z = f(\phi)$. Thus we can also speak of the action of the functional f on ϕ producing a number $f(\phi)$. So we will write $\langle f, \phi \rangle$ for this functional.

Remark 2.4.4. *This notational representation of f, which looks exactly like an inner product between f and ϕ and this similarity is intentional although misleading, since f need not always be representable as a true inner product.*

Definition 2.4.5 (Linear functional). *A functional f on $\mathcal{D}$ is linear if for all $\alpha, \beta \in \mathbb{R}$ and $\phi_1, \phi_2 \in \mathcal{D}$ then*

$$\langle f, \alpha\phi_1 + \beta\phi_2 \rangle = \alpha\langle f, \phi_1 \rangle + \beta\langle f, \phi_2 \rangle.$$

Definition 2.4.6 (Distribution). *A continuous linear functional on the vector space $\mathcal{D}$ is called a distribution T. i.e $\{\phi_i\} \longrightarrow \phi$ in $\mathcal{D} \Rightarrow \{T(\phi_i)\} \longrightarrow T(\phi)$ in $\mathbb{R}$.*

The set of all distributions on $\Omega \subset \mathbb{R}^n$ will be denoted $\mathcal{D}'(\Omega) \subset \mathcal{D}'(\mathbb{R}^n)$. We can see that $\mathcal{D}'(\mathbb{R}^n) \subset \mathcal{D}^*(\mathbb{R}^n)$; the set of all functionals on $\mathbb{R}^n$.

Remark 2.4.7. *One could ask if there exist functionals which are not continuous? Assuming the axiom of choice one can show that there exist linear maps $T : \mathcal{D}(\mathbb{R}^n) \longrightarrow \mathbb{R}$ which are not continuous in the ϕ of $\mathcal{D}(\mathbb{R}^n)$.*

We now look at various examples of distributions

Example 2.4.8. *An all important example is the so called* ***Heaviside Distribution***

$$\langle H, \phi \rangle = \int_0^\infty \phi(x)dx,$$

which is equivalent to the inner product of ϕ with well known Heaviside function

$$H(x) = \begin{cases} 1 & x > 0, \\ 0 & x \leq 0. \end{cases}$$

Example 2.4.9 (Functions as distributions).

First we need the following definition:

Definition 2.4.10. *Suppose $f(x)$ is locally integrable, that is, the Lebesgue integral $\int_I |f(x)dx|$ is defined and bounded for every finite interval I.*

Then the inner product

$$\langle f, \phi \rangle = \int_{-\infty}^{\infty} f(x)\phi(x)dx$$

is a linear functional if $\phi \in \mathcal{D}$. Thus, we have the result:

Theorem 2.4.11. *To every locally integrable function f there corresponds a distribution defined by*

$$\langle f, \phi \rangle = \int_{\infty}^{-\infty} f(x)\phi(x)dx,$$

where the integral is the Lebesgue integral.

Proof. See Theorem 1.14 [3] □

Question 2.4.12. *Do different functions define the same distribution?*

Any continuous function(or in general every locally integrable function f) induces a distribution through the usual inner product. If $\langle f, \phi \rangle = \int_{\infty}^{-\infty} f(x)\phi(x)dx$, then

$$|\langle f, \phi_n \rangle| \leq \int_{\infty}^{-\infty} |f||\phi_n|dx \leq M_n \int_I |f(x)|dx \longrightarrow 0,$$

where $M_n = max_x|\phi(x)|$ and $I = supp(\phi_n)$. It follows that two locally integrable functions which are the same almost everywhere, induce the same distribution. Thus, the following Result:

Theorem 2.4.13. *Two locally integrable functions, f and g have the same distribution if and only if $f = g$ almost everywhere.*

We earlier said that for any function f that is locally integrable we can interchangeably refer to its function values $f(x)$ and to its distributional values. There are numerous distributions which are not representable as an inner product.

Remark 2.4.14. *The fact that some continuous linear functionals on $\mathcal{D}$ exist which do not correspond to inner products should be compared to Riesz Representation Theorem: which states that any linear bounded linear functional on $\mathcal{L}^2$ can be represented as an inner product with some $\mathcal{L}^2$ function. Notice that the functional $f(u) = u(0)$ defined for all $u \in \mathcal{L}^2$ is not bounded linear functional. In fact, it is not even well defined in $\mathcal{L}^2$ since $u(0)$ can be changed arbitrarily without changing the function $u(x)$ in $\mathcal{L}^2$, and so the Riesz Representation Theorem does not apply here. This leads us to the following definition*

Definition 2.4.15. *A distribution is **Regular** if it can be represented as an inner product with some locally integrable function f, and is **Singular** if no such function exists.*

Example 2.4.16. *The delta distribution δ is defined by*

$$\langle \delta, \phi \rangle = \phi(0)$$

for all ϕ in $\mathcal{D}$ is a singular distribution. This is clearly a distribution to see that this cannot be represented as an inner product let

$$\phi_a(x) = \begin{cases} exp(\frac{a^2}{x^2-a^2}) & |x| < 0, \\ 0 & elsewhere. \end{cases}$$

Now, $\phi_a = max_x|\phi_a(x)| = \frac{1}{e}$ which is independent of a. For any locally integrable function $f(x)$

$$\left| \int_{\infty}^{-\infty} f(x)\phi_a(x)dx \right| \leq \frac{1}{e} \int_{a}^{-a} |f(x|dx \longrightarrow 0 \quad as \quad a \longrightarrow 0$$

which is not $\phi_a(0)$.
Thus there is no locally integrable function $f(x)$ for which $\int_{\infty}^{-\infty} f(x)\phi(x)dx = \phi(0)$ for every test function $\phi(x)$.

2.5 The Derivative of a Distribution

For a workable definition of the derivative of a distribution, i.e. a definition which yields consistent results when we differentiate a distribution that is also a function in the ordinary sense, we first develop the notion of the derivative of a distribution. We motivate the differentiation by first looking at a continuously function in Ω then for each i, $\partial f/\partial x_i$ is also a distribution in $\mathcal{D}'(\Omega)$ and we have for $\phi \in \mathcal{D}'(\Omega)$

$$\left\langle \frac{\partial f}{\partial x_i}, \phi \right\rangle = \int_{\Omega} \frac{\partial f}{\partial x_i} \phi \ dx = -\int_{\Omega} f \frac{\partial \phi}{\partial x_i} \ dx = -\left\langle f, \frac{\partial \phi}{\partial x_i} \right\rangle \tag{2.8}$$

where we have used the very important fact that ϕ vanishes outside a compact subset of Ω and thus the boundary terms in the integration by parts vanish. Motivated by the situation we give the following definition.

Definition 2.5.1. *Let $f \in \mathcal{D}'(\Omega)$. The distribution $f' = \partial f/\partial x_i$ is defined by*

$$\langle f', \phi \rangle = -\langle f, \phi' \rangle \tag{2.9}$$

for all $\phi \in \mathcal{D}$.

The operation of partial derivation is well defined since if $\phi \in \mathcal{D}(\Omega)$ then so is $\partial\phi/\partial x_i$. More than that, the operator

$$D_i : \mathcal{D}(\Omega) \longrightarrow \mathcal{D}(\Omega) \quad \text{given by}$$

$$D_i(\phi) = \frac{\partial \phi}{\partial x_i}$$

is clearly continuous. From equation [2.9] we see that the operator $D_i : \mathcal{D}'(\Omega) \longrightarrow \mathcal{D}'(\Omega)$ is also weakly continuous.

In general if $k \in \mathbb{N}^n$ then define $D^k f$ by

$$\left\langle D^k f, \phi \right\rangle = (-)^k \left\langle f, D^k \phi \right\rangle \tag{2.10}$$

Notice that the partial derivative of a distribution is always defined. Therefore in distributional sense any distribution is infinitely differentiable.

Example 2.5.2. *The derivative of the delta function is defined by*

$$\left\langle \delta', \phi \right\rangle = -\left\langle \delta, \phi' \right\rangle = -\phi(0) \tag{2.11}$$

Similarly the $n-th$ derivative of δ is defined by

$$\left\langle \delta^{(n)}, \phi \right\rangle = (-)^n \left\langle \delta, \phi^{(n)} \right\rangle = (-)^n \phi(0). \tag{2.12}$$

In general we have

Proposition 2.5.3. *Every distribution T is infinitely differentiable with*

$$D^k T(\phi) = (-)^k T(\phi^{(k)}).$$

Proof. Repeated use of Definition □

2.5.1 Approximation of distributions using test functions

Definition 2.5.4. *Consider an arbitrary distribution T, then a sequence of distributions $T_1, T_2, T_3, \cdots$ converges to a distribution T if the sequence $\langle T_n . \phi \rangle$ converges to the number $\langle T, \phi \rangle$ for all test functions.*

We simply write this as $T_n \longrightarrow T$ or equivalently say that T_n approximates T. Luckily if $\{T_n\}$ is a sequence of distributions for which $lim_{n \longrightarrow \infty} \langle T_n, \phi \rangle < \infty$ we can construct a distribution by putting

$$\langle T_n, \phi \rangle = lim_{n \longrightarrow \infty} \langle T_n, \phi \rangle < \infty$$

and in fact, $T_n \longrightarrow T$, hence the following results

Theorem 2.5.5. *For given distribution $T \in \mathcal{D}'$, there exists a sequence $\langle \phi_n \rangle$ of test functions such that $\phi_n \longrightarrow T$*

Proof. See Schrichartz [5] □

Chapter 3

The Analytic Continuation of Distributions

In this chapter, we show that distributions are analytically continuable. In particular, we consider analytic continuation of distribution with respect to a parameter: Gelfand's problem. Finally, we reduce the problem of analytical continuation of distribution to that of finding the Berstein-Sato polynomial.

3.1 Introduction

Suppose that $f_\lambda(x)$ are locally integrable functions that depend analytically on the parameter $\lambda \in U_0$. It is often the case that there is a larger region U_1 such that $f_\lambda(x)$ is analytic there. If for each $\phi \in \mathcal{D}(\Omega)$ the function $\Gamma(\lambda) = \langle f_\lambda(x), \phi(x) \rangle$, initially defined for $\lambda \in U_0$, can be continued to U_1.

Example 3.1.1.

The chief example is provided by the function $f_\lambda(x) := H(x)x^\lambda$,where $H(x)$ denotes the Heaviside step function. The function $f_\lambda(x)$ is locally integrable in $\mathbb{R}$ for all values of λ in the right half plane $Re\lambda > -1$ and in this case it defines a regular distribution, customarily denoted as x_+^λ. Let now $\phi \in \mathcal{D}(\mathbb{R})$ and set

$$\Gamma(\lambda) = \left\langle x_+^\lambda, \phi(x) \right\rangle = \int_0^\infty x^\lambda \phi(x) dx, \quad \mathcal{R}e\lambda > -1 \tag{3.1}$$

As we shall presently show, $\Gamma(\lambda)$ can be continued analytically to $\mathbb{C}\backslash\{-1, -2, -3, \cdots\}$. Indeed, the formula

$$\Gamma(\lambda) = \int_0^1 \left[x^\lambda \phi(x) - \phi(0) \right] dx + \int_1^\infty x^\lambda \phi(x) dx + \frac{\phi(0)}{\lambda + 1} \tag{3.2}$$

gives the continuation to the region $\mathcal{R}e\lambda > -2$, $\lambda \neq -1$. More generally the formula

$$\Gamma(\lambda) = \int_0^1 x^\lambda \left[\phi(x) - \sum_{j=0}^{n} \frac{\phi^{(j)}(0)x^j}{j!}\right] dx + \int_1^\infty x^\lambda \phi(x) dx + \sum_{j=0}^{n} \frac{\phi^{(j)}}{j!(\lambda+j+1)}$$

gives the analytic continuation of $\Gamma(\lambda)$ to the region $\mathcal{R}e\lambda > -(n+2)$, $\lambda \neq -1, \cdots, -(n+1)$ for $n \in \mathbb{Z}_+$.

Therefore, the method of analytic continuation allows us to define the generalized function x_+^λ for $\lambda \neq -1, -2, -3, \cdots$. As can be seen from 3.1 the analytic function $\Gamma(\lambda)$ has simple poles at $\lambda = -1, -2, -3, \cdots$ with residues

$$\text{Res}_\lambda = -k\Gamma(\lambda) = \frac{\phi^{(k-1)}(0)}{(k-1)!} \tag{3.3}$$

This means that the generalized function x_+^λ has poles at $\lambda = -k$ for $k = -1, -2, -3, \cdots$ with residues

$$\text{Res}_\lambda = -kx_+^\lambda = \frac{(-1)^{k-1}\delta^{(k-1)}(x)}{(k-1)!} \tag{3.4}$$

3.2 Further Development

We are now ready to describe I. M. Gel'fand problem . We shall work over the base field K of characteristic 0. As a motivation we look at the following problem.

3.3 The Analytic Continuation Problem.

Suppose $p(X) \in \mathbb{K}[x_1, \cdots, x_n]$ be a polynomial in n variables ($p : \mathbb{R}^n \longrightarrow K$) and let Ω be a connected component of $\mathbb{R}^n \backslash \{X : p(X) = 0\}$. We now define

$$p_\Omega(x) = \begin{cases} p(x) & \text{if } x \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

Let us take any $\lambda \in \mathbb{C}$ and consider the function $|p_\Omega|^\lambda$. It takes a minute to see that if Re $\lambda \geq 0$ then $|p_\Omega|^\lambda$ makes sense as a distribution on $\mathbb{R}^n$. Indeed,

$$\begin{aligned} |p^\lambda| &= |e^{\lambda \log p(x)}| \\ &= |e^{(\alpha+i\beta)\log p(x)}| \quad \text{for} \quad \lambda = \alpha + i\beta \\ &= |e^{\alpha \log p(x)}| \cdot |e^{i\beta \log p(x)}| \\ &= e^{\alpha \log p(x)} \quad \text{since} \quad |e^{ix}| = 1 \end{aligned} \tag{3.6}$$

It follows that the complex power p^λ exist as a continuous function in Ω when $Re(\lambda) \geq 0$. Therefore,

$$\Gamma_\phi(\lambda) = \left\langle |p_\Omega|^\lambda, \phi(x) \right\rangle = \int_\Omega |p_\Omega|^\lambda \phi(x) dx \tag{3.7}$$

is is convergent for any $\phi \in C_0^\infty(\mathbb{R}^n)$ - a smooth function on $\mathbb{R}^n$ with compact support. Moreover, the complex derivative is

$$\frac{d\Gamma_\phi}{d\lambda} = \int log|p_\Omega(x)||p_\Omega|^\lambda(x)\phi(x)dx.$$

So, Γ is an analytic function there. In particular $\Gamma_\phi(\lambda)$ is a function of the complex variable λ, defined in the half-space $Re(\lambda) > 0$.

It is easy to see that for Re$\lambda \geq 0$ we have a holomorphic family of distributions $\lambda \longrightarrow |p_\Omega(x)|^\lambda$

Question 3.3.1 (Gel'fand). *Can we analytically continue this family meromorphically in λ to the whole $\mathbb{C}$? Analytically continuing is the same thing as saying you can extend the integral $\int_\Omega |P_\Omega|^\lambda \phi(x)dx$ for other values of λ.*

The answer is yes!: This is a theorem of Bernstein and Sato (at about the same time). We can write

$$b(\lambda)p(x)^\lambda = \text{(differential operator with polynomial coefficients)}p(x)^{\lambda+1}. \tag{3.8}$$

This monic $b(\lambda)$ is called the *Bernstein-Sato polynomial of* $p(x)$.

Claim 3.3.2. *If we can find such a relation in 3.8, then we can continue $p(x)$ to all complex λ.*

To see how Berstein-Sato idea works, we look at the following toy example;

Example 3.3.3.

Let $n = 1$, $p(x) = x^2$ since p is always positive we need not consider the absolute value of p and $\Omega = \mathbb{C}_+$: the positive part of the complex plane. As shown in figure 3.1.

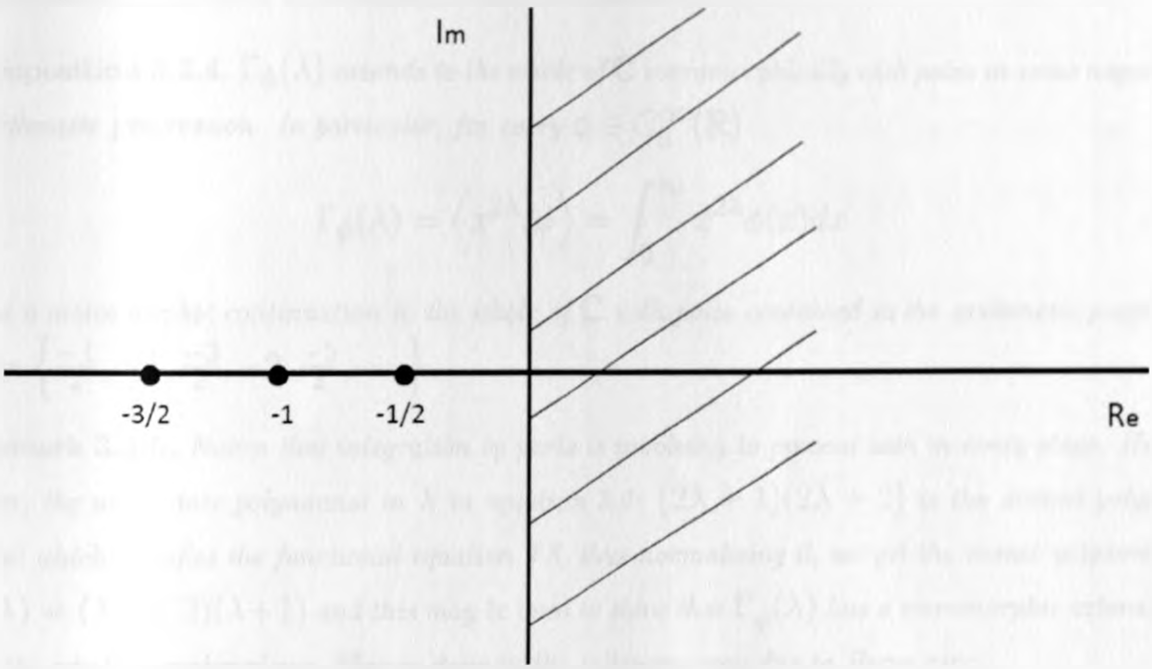

Figure 3.1: **Continuation to the entire complex plane.**

then for any test function $\phi(x)$ it follows by partial integration that

$$\begin{aligned}\Gamma_\phi(\lambda) = \left\langle x^{2\lambda}, \phi\right\rangle &= \int_0^\infty x^{2\lambda}\phi(x)dx \\ &= \left[\frac{x^{2\lambda+1}}{2\lambda+1}\phi(x)\right]_0^\infty - \frac{1}{2\lambda+1}\int_0^\infty x^{2\lambda+1}\phi'(x)dx \\ &= -\frac{1}{2\lambda+1}\int_0^\infty x^{2\lambda+1}\phi'(x)dx\end{aligned}$$

integrating again we get,

$$\begin{aligned}\int_0^\infty x^{2\lambda}\phi(x)dx &= -\left[\frac{x^{2\lambda+2}}{(2\lambda+1)(2\lambda+2)}\phi'(x)\right]_0^\infty + \frac{1}{(2\lambda+1)(2\lambda+2)}\int_0^\infty x^{2\lambda+2}\phi''(x)dx \\ &= \frac{1}{(2\lambda+1)(2\lambda+2)}\int_0^\infty x^{(2\lambda+2}\phi''(x)dx\end{aligned}$$

or more conveniently

$$(2\lambda+1)(2\lambda+2)\int_0^\infty x^{2\lambda}\phi(x)dx = \int_0^\infty x^{2\lambda+2}\phi''(x)dx \tag{3.9}$$

Continuing this process by induction we get the following.

Proposition 3.3.4. $\Gamma_\phi(\lambda)$ *extends to the whole of* $\mathbb{C}$ *meromorphically with poles in some negative arithmetic progression. In particular, for every* $\phi \in C_0^\infty(\mathbb{R})$

$$\Gamma_\phi(\lambda) = \left\langle x^{2\lambda}, \phi \right\rangle = \int_0^\infty x^{2\lambda}\phi(x)dx$$

has a meromorphic continuation to the whole of $\mathbb{C}$ *with poles contained in the arithmetic progression* $\left\{\frac{-1}{2}, -1, \frac{-3}{2}, -2, \frac{-5}{2}, \cdots\right\}$.

Remark 3.3.5. *Notice that integration by parts is involving to proceed with in every stage. However, the univariate polynomial in* λ *in equation 3.9:* $(2\lambda+1)(2\lambda+2)$ *is the desired polynomial which satisfies the functional equation 3.8, thus normalising it, we get the monic polynomial* $b(\lambda) = (\lambda+1/2)(\lambda+1)$ *and this may be used to show that* $\Gamma_\phi(\lambda)$ *has a meromorphic extension to the whole complex plane. This is done in the following way due to Bernstein:*

We observe that since $b(\lambda) = (\lambda+1/2)(\lambda+1)$ *then,*

$$\begin{aligned}
b(\lambda+1) =& (\lambda+2)(\lambda+\frac{3}{2})\\
b(\lambda+2) =& (\lambda+3)(\lambda+\frac{5}{2}),\\
b(\lambda+3) =& (\lambda+4)(\lambda+\frac{7}{2}),\\
&\vdots\\
&e.t.c
\end{aligned}$$

Remark 3.3.6. *The proposition works out not only for functions with compact support but also for functions which are rapidly decreasing at* $+\infty$ *together with all derivatives.*

Example 3.3.7. *We can take* $\phi(x) = e^{-x}$. *In this case we have* $\int_0^\infty e^{-x}x^\lambda dx = \Gamma(\lambda+1)$. *The proposition implies that* $\Gamma(\lambda)$ *has a meromorphic continuation with poles at* $0, -1, -2, \cdots$. *In fact, this is the way we analytically continue Gamma functions.*

To sum up, to analytically continue $|p(x)|^\lambda$ to all values of λ in the complex plane all what we need is to obtain the Bernstein-Sato polynomial for $|p(x)|$ and we are done. That is, given a polynomial $p = p[x_1, \cdots, x_n] \in \mathbb{K}[x_1, \cdots, x_n]$, we compute the the corresponding Bernstein-Sato polynomial and we find the corresponding integer shift to get the poles. However, thing are note that easy !?

Example 3.3.8. $p(x) = x^2 + y^3$. *If you manage to find the Bernstein polynomial of this, we shall be impressed. It takes a huge amount of calculation. The answer* $b(\lambda) = (\lambda + 1)(\lambda + 5/6)(\lambda + 7/6)$. *Direct calculation is very hard.*

The point is that finding Bernstein polynomials in general is very difficult. Fortunately, we don't always need to know $b(\lambda)$ explicitly. Bernstein showed how to prove that it exists without actually producing it. The existence result is actually very powerful.

Chapter 4

A Quick Introduction to D-module Theory

In this chapter we develop the necessary D-module theory to enable us prove the Bernstein Theorem in the next chapter. Therefore will define the Weyl Algebra and present its basic properties: it is a domain, simple and noetherian. Finally, we will consider the modules over the Weyl algebra to define the dimension which will eventually enable us to consider some special D-modules of dimension n called holonomic modules. The results in this chapter can be found in [1].

4.1 The Weyl Algebra

Let K be a field of characteristic 0. Lots of geometrically interesting concepts exist for the case $K = \mathbb{C}$; We may limit ourselves to $\mathbb{C}$ for simplicity when looking for examples. For a given positive integer n, we let $x_1, \cdots, x_n$ be n indeterminate variables. Then $K[X] = K[x_1, \cdots, x_n]$ is the ring of polynomials in n commuting indeterminates over K. We also introduce n new symbols denoting the partial derivatives with respect to $x_1, \cdots x_n$, and we call these $\partial_1, \cdots \partial_n$. Then we define the Weyl Algebra as follows

Definition 4.1.1 (The Weyl Algebra). *The Weyl algebra $A_n(K) = K[x_1, \cdots x_n, \partial_1, \cdots \partial_n]$ is the algebra with the usual laws of polynomial addition endowed with the following axioms of multiplication (commutation relations) due to Heisenberg:*

$$[x_i, x_j] = 0, \quad [\partial_i, \partial_j] = 0, \quad [\partial_i, x_j] = \delta_{ij},$$

It is worth noting that the last law comes from thinking of these as derivatives applied to a module of functions $f(x_1, \cdots, x_n)$. The product rule from calculus says that

$$\begin{aligned}\partial_i(x_i f) &= \partial_i(x_i) f + x_i \partial_i f \\ &= f + x_i \partial_i f \\ &= (1 + x_i \partial_i) f\end{aligned}$$

Thus, as operators we see that $\partial_i x_i = 1 + x_i \partial_i$. These rules together with the distributive laws define the noncommutative ring A_n in $2n$ variables.

Remark 4.1.2. *Most of the time we do not need to mention the underlying field K, so we will refer to $A_n(K)$ as simply A_n.*

The monomials of A_n form the so-called PBW (Poincare-Birkhoff- Witt) basis meaning that an arbitrary element $P \in A_n$ can be uniquely written in the form

$$P = \sum a_{\alpha\beta} x^\alpha \partial^\beta; \tag{4.1}$$

where $\alpha = (\alpha_1, \cdots, \alpha_n), \beta = (\beta_1, \cdots, \beta_n)$ are multi-indices, $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and $\partial^\beta = \partial_1^{\beta_1} \cdots \partial_n^{\beta_n}$, $a_{\alpha\beta} \in K$, the sum in (4.1) is finite. From this we can see that A_n has as a basis (as a vector space over K), terms of the form $x^\alpha \partial^\beta$. Hence, A_n can be thought of as the algebra of differential operators with polynomial coefficients acting on $K[x_1, \cdots, x_n]$ by formal differentiation.

Definition 4.1.3. *For $D \in A_n$, the degree of D is the largest $|\alpha| + |\beta|$ occuring in the summands of D.*

Example 4.1.4. *The degree of $3x_1x_2\partial_2 + x_1x_2^2\partial_1\partial_3$ is 5.*

The degree has the following properties; for $D, D' \in A_n$

$$\begin{aligned}\deg(D + D') &\leq max(\deg(D), \deg(D')), \\ \deg(DD') &= \deg(D) + \deg(D'), \\ \deg[D, D'] &\leq \deg(D) + \deg(D') - 2.\end{aligned}$$

Proposition 4.1.5. *The algebra A_n is a domain.*

Proposition 4.1.6. *$A_n(K)$ is a simple algebra, i.e any two sided ideal in A_n is $\{0\}$ or A_n*

Proof. Assume $I \subset A_n$ be a two sided ideal and $0 \neq P = \sum a_{\alpha\beta} x^\alpha \partial^\beta \in I$ If x_j occurs in P with the highest order $m > 0$, then $0 \neq [\partial_j, P] \in I$ and x_j occurs in $[\partial_j, P]$ with the highest order $m - 1$. By applying $[\partial_j,]$ m times to P one obtains a nonzero element in I without the variable x_j. Using all the $[\partial_j,]$, $[x_j,]$ gives $1 \in I$. So $I = A_n$ □

Remark 4.1.7. *An A_n-module is a D-module. By an A_n-module we mean a left unitary A_n-module with a finite number of generators, unless otherwise specified.*

4.2 Grading and Filtrations

Definition 4.2.1. *A K-algebra R is $\mathbb{N}$-graded or simply, graded, if there exists a set of K-subspaces G_i such that:*

$$G_i . G_j \subseteq G_{i+j}$$
$$R = \bigoplus_{i \in \mathbb{N}} G_i$$

The nonzero elements of the subspace G_n are called homogeneous of degree n. The algebra $P = K[x_1, \cdots, x_n]$ is a good example of a graded K-algebra. We let P_i be a vector subspace of P over K generated by monomials of degree i for $i \geq 0$ i.e. the usual polynomials of degree i. Then

$$P_i P_j = P_{i+j} \quad \text{for} \quad i, j \geq 0 \quad \text{and} \quad P = \bigoplus_i P_i.$$

This is called the grading by degree.

Remark 4.2.2. *It is unfortunate that due to noncommutativity, the degree of an operator in the Weyl algebra-A_n cannot be used to make the A_n graded ring. This is because elements like $\partial_1 x_1$ ought to be homogeneous of degree 2, but it is also equal to $x_1 \partial_1 + 1$ which is not homogeneous. To use this degree effectively we must generalise graded rings, to get filtered rings.*

Definition 4.2.3. *A filtration of an algebra R is an increasing sequence of K-subspaces*

$$\{0\} = F_{-1} \subset F_0 \subset F_1 \subset F_2 \subset \cdots \subset R$$

with $\bigcup_j F_j = R$ and $F_i F_j \subseteq F_{i+j}$ for all i, j. F_0 is a subalgebra of R

By putting $F_i = \bigoplus_{i=0}^{i} G_k$ in a graded algebra $R = \bigoplus G_j$ we see that every graded algebra is filtered. However, there are filtered algebras which do not have a natural grading.

The Weyl algebra which has many different filtrations. We shall endow A_n with the following two filtrations.

1. **Bernstein filtration**:
   Set $B_j = \{a_{\alpha\beta}x^\alpha\partial^\beta : |\alpha| + |\beta| \leq j\}$ then $\bigcup B_j = A_n$ is a filtered ring. The advantage of this is that B_j are finite dimensional over K.

2. **Standard filtration (by the order of a differential operator)**:
   We put $A_n = \bigcup C_j$ $C_j = \{a_{\alpha\beta}x^\alpha\partial^\beta : |\beta| \leq j\}$. Note that $C_0 = K[X]$ is an infinite dimensional K-vector space. However, the order filtration has the advantage that unlike the Bernstein filtration, it is well defined for other rings of differential operators. That is, this filtration can be generalized to algebras of differential operators on varieties; the C_j are finite $K[x]$-modules.

We now denote by $\{\mathcal{F}_j\}$ any of these filtrations. Thus,

$$\mathcal{F}_i\mathcal{F}_j \subset \mathcal{F}_{i+j}, \text{and} \cup \mathcal{F}_j = A_n.$$

In general for any ring R with filtration $F : F_0 \subset F_1 \subset F_2 \subset F_3 \subset \cdots$ one constructs the associated graded ring

$$gr^F R = \bigoplus_{i=1}^{\infty} F_i/F_{i-1}$$

Therefore, for our ring with filtration $\mathcal{F}$, we construct the associated graded ring and denote it with

$$gr^{\mathcal{F}} A_n = \bigoplus_{i=0}^{\infty} (\mathcal{F}_i/\mathcal{F}_{i-1}).$$

Proposition 4.2.4. *$gr^{\mathcal{F}} A_n$ is canonically isomorphic to a polynomial ring $S_n = K[\bar{x}_1, \cdots, \bar{x}_n, \bar{\partial}_1, \cdots, \bar{\partial}_n]$ in $2n$ variables, $\bar{x}_j, \bar{\partial}_j \in \mathcal{F}_1/\mathcal{F}_0$*

Proof. The proof uses $[\partial_i, x_j] = \delta_{ij}$ □

For example, in the case of $A_2(K)$ the corresponding polynomial is S_2 generated by

$$\bar{x_1} = x_1 + K, \quad \bar{x_2} = x_2 + K, \quad \bar{\partial_1} = \partial_1 + K, \quad \bar{\partial_2} = \partial_2 + K,$$

S_n is an algebra of polynomials in 4 variables and it is commutative because $[\bar{x_i}, \bar{\partial_j}] = \bar{0} \quad \forall i, j$

4.2.1 Filtration of A_n modules.

Let M be an A_n-modules. A filtration of M is an increasing sequence of K-subspaces

$$\{0\} = \Gamma_{-1} \subset \Gamma_0 \subset \Gamma_1 \subset \Gamma_2 \cdots \subset M$$

such that

1. $\bigcup \Gamma_j = M$
2. $\mathcal{F}_i \Gamma_j \subset \Gamma_{i+j}$ for all i, j.
3. $\mathcal{F}_0$-modules Γ_j is a K-vector space of finite dimension.

There is an associated graded module

$$gr^{\Gamma} M = \bigoplus_{j=0}^{\infty} \Gamma_j / \Gamma_{j-1} \quad \text{over} \quad gr^{\mathcal{F}} A_n$$

Remark 4.2.5. *Let M be an A_n-module and Γ a filtration on M. If $gr^{\Gamma} M$ is finitely generated then it is noetherian, hence M is finitely generated . However , it is not always true that if M is finitely generated over A_n then $gr^{\Gamma} M$ is finitely generated over $gr^{\Gamma} A_n$. When $gr^{\Gamma} M$ is finitely generated we say that Γ is good filtration.*

Definition 4.2.6. *Γ is called a good filtration of M if $gr^{\Gamma} M$ is finitely generated over $gr A_n$.*

Proposition 4.2.7.

1. If $gr^{\Gamma} M$ is finitely generated over $gr A_n$ then M is finitely generated over A_n.
2. Γ is good, if and only if there exists some j_0 such that $\mathcal{F}_i \Gamma_j = \Gamma_{i+j}$ for all $i \geq 0, j \geq j_0$.
3. If Γ and Ω are filtrations of M and Γ is good, then there exists a j_1, such that $\Gamma \subset \Omega_{j_1+j}$ for all j. If Γ and Ω are both good, then there exist a j_2, such that $\Omega_{j-j_2} \subset \Gamma \subset \Omega_{j+j_2}$ for all j.

Proposition 4.2.8. *Finitely generated A_n-modules admit good filtrations.*

Let M be a left A_n module with filtration Γ. Let N be a submodule, and set Γ', Γ'' to be the induced filtrations on N, M/N. Then we have an exact sequence

$$0 \longrightarrow gr^{\Gamma'}N \longrightarrow gr^{\Gamma}M \longrightarrow gr^{\Gamma''}M/N \longrightarrow 0$$

Proposition 4.2.9. *Let M be a left A_n module, if M is finitely generated over A_n then M is Noetherian.*

4.3 Dimensions and Multiplicities

Remark 4.3.1. *From now onwards we restrict ourselves A_n endowed with the Berstein filtration.*

Consider the Bernstein filtration B_j on A_n. For this filtration we have

$$dim_K B_j = \frac{j^{2n}}{2n!} + \text{terms of lower order in } j \tag{4.2}$$

To see this we take the simplest case $A_1(\mathbb{C})$

$$B_{-1} \subset B_0 \subset B_1 \subset B_2 \subset B_3 \subset \cdots$$

where B_j are given as,

$$\begin{aligned} B_{-1} =& \{0\}, \\ B_0 =& \mathbb{C}1, \\ B_1 =& \mathbb{C}1 + \mathbb{C}x + \mathbb{C}\partial, \\ B_2 =& \mathbb{C}1 + \mathbb{C}x + \mathbb{C}\partial + \mathbb{C}x^2 + \mathbb{C}x\partial + \mathbb{C}\partial^2 \\ B_3 =& \mathbb{C}1 + \mathbb{C}x + \mathbb{C}\partial + \mathbb{C}x^2 + \mathbb{C}x\partial + \mathbb{C}\partial^2 + \mathbb{C}x^3 + \mathbb{C}x^2\partial + \mathbb{C}x\partial^2 + \mathbb{C}\partial^3 \end{aligned}$$

e.t.c

Now we consider the map $j \longmapsto dim_{\mathbb{C}}(B_j)$; that is $1 \longmapsto 3, 2 \longmapsto 6, 3 \longmapsto 10, \cdots, j \longmapsto \frac{(j+1)(j+2)}{2} = \frac{1}{2}j^2$ + lower order terms in j as claimed in equation (4.2)

Definition 4.3.2 (Hilbert polynomial). *Let M be a finitely generated A_n-module with a good filtration $\{\Gamma_j\}$, so that $gr^{\Gamma} M$ is a finitely generated module over the graded ring $S = gr^B A_n = K[\bar{x}_1, ..., \bar{x}_n, \bar{\partial}_1, ..., \bar{\partial}_n]$. Then, there exists a polynomial*

$$\chi(M,\Gamma,t) = \frac{m}{d!}t^d + O(t^{d-1}), \quad m, d \in \mathbb{Z}_+$$

such that

$$dim_K \Gamma_j = \sum_{i=0}^{j} dim_K \Gamma_j/\Gamma_{j-1} = \chi(M,\Gamma,j) \quad \text{for} \quad j >> 0$$

This is the Hilbert polynomial of $gr^{\Gamma} M$. Since $gr^B A_n$ is a polynomial ring in $2n$ variables, $d \leq 2n$.

If Γ, Ω are good filtration on M, by definition there exist a k such that:

$$\chi(M,\Omega,j-k) \leq \chi(M,\Gamma,j) \leq \chi(M,\Omega,j+k)$$

since the leading term of χ is not affected by a shift in j. The numbers $d = d(M)$ and $m = m(M)$ do not depend on the filtration Γ; d is called the **dimension** of the module M while m is called the **multiplicity** of M.

Example 4.3.3.

Let $M = A_n$ and B be the Bernstein filtration. We want to find $\chi(A_n, B, t)$, so we need to calculate the $\dim_K(B_j)$. Recall that B_i is spanned by the monomials $x^\alpha \partial^\beta$ with $|\alpha| + |\beta| \leq i$; so we want to find nonnegative solution to

$$\alpha_1 + \cdots + \alpha_n + \beta_1 + \cdots + \beta_n \leq i.$$

The number of solutions is

$$\binom{2n-1}{2n-1} + \binom{2n}{2n-1} + \binom{2n+1}{2n-1} + \cdots + \binom{2n+i-1}{2n-1} = \binom{2n+i}{2n}$$
$$= \frac{(i+2n)(i+2n-1)\cdots(i+1)}{2n!}$$

$$\text{So} \quad \chi(A_n, B, t) = \frac{(t+2n)(t+2n-1)\cdots(t+1)}{2n!}$$

This is a polynomial of degree $2n$ with leading coefficient $1/(2n!)$. Thus $d(A_n) = 2n$ and $m(A_n) = 1$.

Example 4.3.4.

Let

$$M = \frac{A_n}{\Sigma_{j=1}^{n} A_n \partial_j} \cong K[x_1, \cdots, x_n]$$

and Γ be a standard filtration then $\dim(\Gamma_j) = \binom{n+j}{n}$, $d = n$ and$m = 1$.

Proposition 4.3.5. *Let $0 \longrightarrow N \longrightarrow M \longrightarrow M/N \longrightarrow 0$ be an exact sequence of A_n-modules with M a finitely generated A_n-module and Γ a good filtration of M. Let N be a submodule of M. Denote by Γ' and Γ'' the filtration induced by Γ on N and M/N respectively. Then*

$$0 \longrightarrow gr^{\Gamma'} N \longrightarrow gr^{\Gamma} M \longrightarrow gr^{\Gamma''} M/N \longrightarrow 0$$

is an exact sequence of $gr^B A_n$ modules, Γ' and Γ'' are good filtrations,and we have

1. $\chi(M, \Gamma, t) = \chi(M, \Gamma', t) + \chi(M, \Gamma'', t)$
2. $d(M) = max(d(N), d(M/N))$
3. *If* $d(N) = d(M/N)$ *then* $m(M) = m(N) + m(M/N)$

Corollary 4.3.6. *Let $M = \bigoplus_{i=1}^{k} M_i$ be a direct sum of finitely generated A_n modules. Then $d(M) = max(d(M_i))$. and if $d(M) = d(M_i)$ for $1 \leq i \leq k$ then $m(M) = \sum_{i=1}^{k} m(M_i)$.*

Corollary 4.3.7. *If M is finitely A_n-module then $d(M) \leq 2n$.*

Proof. Since M is finitely generated implies there exists l elements which generates M and there exists a surjection

$$A_n^l \xrightarrow{p} M \longrightarrow 0$$

The theorem gives

$$d(A_n^l) = max((d(M), d(\ker p)).$$

On the other hand the previous corollary gives

$$d(A_n^l) = d(A_n) = 2n$$

So $d(M) \leq 2n$. □

We use the following technical lemma due to A. Joseph in the proof of Bernstein Inequality .

Lemma 4.3.8. *Let M be finitely generated A_n module with filtration Γ with respect to B and $\Gamma_0 \neq 0$. Then the K-linear map*

$$B_i \longrightarrow Hom_K(\Gamma_i, \Gamma_{2i}),$$

defined by sending $a \in B_i$ to the map $u \longrightarrow au$, is surjective

Proof. It is enough to show that $a\Gamma_i \neq \{0\}$ whenever $a \in B_i$ is not 0. We proceed by induction on i. For $i = 0, B_0 = K$, so this follows from the fact that that $\Gamma_0 \neq 0$. Assume on the contrary, $a\Gamma_i = \{0\}$. We must have $a \notin K$, so a contains a term of the form $cx^\alpha\partial^\beta$,with $|\alpha| + |\beta| > 0$. Suppose that some $\alpha_j \neq 0$. Then $[a, \partial_j]$ contains a terms $\alpha_j c x^{\alpha - e_j}\partial^\beta$.In particular, it will be nonzero element of B_{i-1}.We have

$$[a, \partial_j]\Gamma_{i-1} = a\partial_j\Gamma_{i-1} + \partial_j a\Gamma_{i-1}.$$

Since $\partial_j\Gamma_{i-1} \subset \Gamma_i$ and $a\Gamma_i = 0$, we have

$$[a, \partial_j]\Gamma_{i-1} = 0,$$

which is a contradiction.

Similarly, suppose that some $\beta_j \neq 0$. Then $[a, x_j]$ contains a term $\beta_j c x^\alpha\partial^{\beta - e_j}$. In particular it will be a nonzero element of B_{i-1}. We have

$$[a, x_j]\Gamma_{i-1} = ax_j\Gamma_{i-1} + x_j a\Gamma_{i-1}.$$

Since $x_j\Gamma_{i-1} \subset \Gamma_i$ and $a\Gamma_i = 0$, we have

$$[a, x_j]\Gamma_{i-1} = 0,$$

which is also a contradiction. □

Theorem 4.3.9 (Bernstein Inequality 1971 [8]). *Let $M \neq 0$ be a finitely generated A_n-module. Then $d(M) \geq n$.*

Proof. We create a good filtration Γ of M by having a set of generators be in Γ_0. So $\Gamma_0 \neq 0$. The previous lemma then applies to give an injection

$$B_i \longrightarrow Hom_K(\Gamma_i, \Gamma_{2i})$$

So $\dim_K(B_i) \leq \dim_K(Hom_K(\Gamma_i, \Gamma_{2i})) = \dim_K(\Gamma_i)\dim_K(\Gamma_{2i})$.

Writing $\chi(t)$ for $\chi(M, \Gamma, t)$, for i sufficiently large we have that

$$\dim_K(B_i) \leq \chi(i)\chi(2i).$$

But since $d(M) = 2n$, we know that $\dim_K(B_i)$ is a polynomial in i of degree $2n$. So

$$2n \leq \deg(\chi(i)\chi(2i)).$$

But $\chi(i)$ is a polynomial of degree $d(M)$, so we have

$$2n \leq 2d(M).$$

Thus the result follows. □

4.4 Holonomic modules

Definition 4.4.1. *A finitely generated A_n module M is called a holonomic module if it is zero or if $d(M) = n$.*

Example 4.4.2. *$K[x_1, \cdots, x_n]$ is holonomic A_n- module.*

Proposition 4.4.3. *Submodules, quotients, and finite direct sums of holonomic modules are holonomic.*

Theorem 4.4.4. *Holonomic modules are Artinian.*

Proof. Let M be holonomic. Suppose it has a descending chain

$$M = N_0 \supset N_1 \supset \cdots \supset N_l$$

Since submodules and quotients of holonomic modules are holonomic, $n = d(N_i) = d(N_i/N_{i+1})$. Then

$$m(M) = \sum_{i=0}^{l-1} m(N_i/N_{i+1}) + m(N_l)$$

Since multiplicity is a nonnegative integer, this sum must be at least l. Hence M cannot have an infinite descending chain. □

Chapter 5

The Bernstein-Sato polynomials

In this section we show the construction of a particular family of holonomic A_n modules which are necessarily for the proof of the existence of the Bernstein-Sato polynomials. First, we need the following result.

Lemma 5.0.5. *Let M be a left A_n module with filtration Γ with respect to Berstein filtration of A_n. Suppose there exists c, c_1 such that for all $j >> 0$,*

$$\dim_K \Gamma_j \le c\frac{j^n}{n!} + c_1(j+1)^{n-1}.$$

Then M is holonomic with $m(M) \le c$. In particular, M is finitely generated.

Proof. We prove the lemma in two parts. First we show that the conclusion holds for every finitely generated submodule of M, and then prove that M itself is finitely generated.

We now show that every finitely generated submodule of M is holonomic with $m(M) \le c$. Let N be a finitely generated submodule. Then it has a good filtration Ω. Thus there exists an integer r with $\Omega_j \subset \Gamma_{j+r} \cap N$, so that $\dim_K(\Omega_j) \le \dim_K(\Gamma_{j+r})$.By our hypothesis, for $j >> 0$

$$\chi(j) \le c\frac{(j+r)^n}{n!} + c_1(j+r)^{n-1}.$$

Hence $d(N) \le n$ and $m(N) \le c$. By the Bernstein Inequality, $d(N) = n$. We are done with the first part.

Now we want to prove that M is finitely generated. Choose an ascending chain of finitely generated submodules of M

$$N_1 \subset N_2 \subset \cdots \subset N_l$$

Each is holonomic with $m(N_i) \leq c$. And since the N_i have the same dimension, their multiplicities satisfy

$$m(N_i) = M(N_{i-1}) + m(N_i/N_{i-1}).$$

So we have

$$m(M) = m(M) + \sum_{i=2}^{l} m(N_i/N_{i-1})$$

Since $m(N_l) \leq c$ and multiplicities are nonnegative integers, all ascending chains of finitely generated submodules must have fewer than c steps. This implies that M must be finitely generated.

We can thus apply the first part to M to see that it is holonomic with $m(M) \leq c$. □

Remark 5.0.6. *If M is holonomic, then its length is bounded by $m(M)$ and holonomic modules are cyclic.*

Theorem 5.0.7. *Let $p \in K[X]$. Then $K[X, p^{-1}]$ is a holonomic A_n module.*

Proof. Let m be the degree of p. We define a filtration Γ for $K[X, p^{-1}]$, and show that this filtration is amenable to an estimate as in the lemma. Set

$$\Gamma_l = \{f/p^l : \deg(f) \leq (m+1)l\}.$$

It is easy to check that all elements of $K[X, p^{-1}]$ are in some Γ_l. It is similarly easy to see that

$$B_i \Gamma_l \subset \Gamma_{l+i}$$

and that the dimension of Γ_l is bounded by the dimension of the space of polynomials of degree $\leq (m+1)l$, so is finite dimensional. In particular, a similar counting trick as employed in the proof that $d(A_n) = 2n$ shows that

$$\dim_K((m+1)l + n) \cdots ((m+1)l + 1)/n!.$$

Since the highest order term in l on the left hand side is $(m+1)^n l^n/n!$, there is a c such that

$$\dim_K \Gamma_l \leq cl^n/n!.$$

We apply the lemma and are done. □

The following example will illustrate:

Example 5.0.8. $M = \mathbb{C}[x, x^{-1}]$ *is holonomic. A filtration* Γ *on* M *is given by* $\Gamma : \Gamma_0 \subset \Gamma_1 \subset \Gamma_2 \subset \Gamma_3 \subset \cdots$ *where*

$$\begin{aligned}\Gamma_0 &= \mathbb{C}1 \\ \Gamma_1 &= \mathbb{C}1 + \mathbb{C}x + \mathbb{C}x^{-1} \\ \Gamma_2 &= \mathbb{C}1 + \mathbb{C}x + \mathbb{C}x^{-1} + \mathbb{C}x^2 + \mathbb{C}x^{-2} \\ \Gamma_3 &= \mathbb{C}1 + \mathbb{C}x + \mathbb{C}x^{-1} + \mathbb{C}x^2 + \mathbb{C}x^{-2} + \mathbb{C}x^3 + \mathbb{C}x^{-3}\end{aligned}$$

Now the map $j \longrightarrow \dim_{\mathbb{C}} \Gamma_j; \quad 1 \longmapsto 3, 2 \longmapsto 3+2, 3 \longmapsto 3+2+2$
thus $j \longrightarrow 3 + 2(j-1)$.

Let λ be transcendental over $K[X] = K[x_1, \cdots, x_n]$. Consider the Weyl algebra $A_n(K(\lambda))$ over $K(\lambda)$; the field of rational functions in λ. Let $p \in K[x_1, \cdots, x_n]$ with $\deg(p) = m$. We define a symbol p^λ on which the element ∂_j acts by formal differentiation:

$$\partial_j p^\lambda = \lambda p^{-1} \frac{\partial p}{\partial x_j} p^\lambda$$

So we define an $A_n(K(\lambda))$ module $M = K(\lambda)[X, p^{-1}]p^\lambda$ generated over $K(\lambda)[X, p^{-1}]$ by the generator p^λ. Further let $N = A_n(K(\lambda)).p^\lambda$ Introducing in M a filtration similar to the one in Theorem 5.0.7 that is $\Gamma_j := \{f.p^{-1}.p^\lambda | \deg f \leq (m+1)j\}$. Then M and N are holonomic.

Theorem 5.0.9 (Generalisation of theorem 5.0.7). *Let* $p \in K[X]$, *then* $K(\lambda)[X, p^{-1}]p^\lambda$ *is a holonomic* $A_n(K(\lambda))$ *module. Hence its submodule* $A_n(K(\lambda))p^\lambda$ *is also holonomic.*

We need also the following result:

Lemma 5.0.10. *Now we have automorphism* t *of* $K(\lambda)[X, p^{-1}]p^\lambda$ *defined by*

$$t(\lambda^i p^\lambda) = (\lambda + 1)^i p p^\lambda$$

This is $A_n(K)$ *linear, but not* $A_n K(\lambda)$ *linear*

5.1 Existence of Bernstein-Sato polynomial

We are now ready to prove that for every polynomial $p \in K[X]$ there exists a corresponding a Bernstein-Sato polynomial.

Theorem 5.1.1. *Let $p \in K[x_1, \cdots, x_n]$, then there exist $B(\lambda) \in K[\lambda]$ and $d(\lambda) \in A_n(K)[\lambda]$ such that*

$$B(\lambda)p^\lambda = d(\lambda)pp^\lambda \tag{5.1}$$

Proof. Since $A_n(K(\lambda))p^\lambda$ is holonomic, it is Artinian. Hence the descending chain

$$A_n(K(\lambda))p^\lambda \supset A_n(K(\lambda))pp^\lambda \supset A_n(K(\lambda))p^2p^\lambda \supset \cdots$$

must terminate. So there exist k such that

$$p^kp^\lambda \in A_n(K(\lambda))p^{k+1}p^\lambda$$

We apply t^{-k} to both sides, Noting that $t^{-k}(p^kp^\lambda)$ is p^λ times a rational function in λ we get

$$p^\lambda \in A_n(K(\lambda))pp^\lambda$$

Clearing the denominators of λ, we get that for some $B(\lambda) \in K[\lambda]$,

$$B(\lambda) \in A_n(K)[\lambda]pp^\lambda$$

So there is a $d(\lambda) \in A_n(K)[\lambda]$ with

$$B(\lambda)p^\lambda = d(\lambda)pp^\lambda$$

□

Remark 5.1.2. *This equation [5.1] is called the* **Bernstein's functional equation**. *$B(\lambda)$ and $d(\lambda)$ are not uniquely determined by this functional equation. However, by the fact that the polynomial ring $K[\lambda]$ is a principal ideal domain, there exists a unique polynomial $b(\lambda)$ of smallest possible degree which enters in a functional equation satisfied by a given polynomial p. Here $b(\lambda)$ is normalised so that its highest coefficient is 1 and $b(\lambda)$ is called the Bernstein-Sato polynomial of p.*

Definition 5.1.3. *The* ***Bernstein-Sato polynomial*** *or* ***b-function*** *of p, written $b(\lambda)$, is defined to be the monic generator of the ideal of all possible $B(\lambda)$ satisfying the Bernstein's functional equation.*

Example 5.1.4. *Let $p(x) = x$. Then*

$$\frac{d}{dx}x^{\lambda+1} = (\lambda+1)x^{\lambda}$$

Hence $b(\lambda) = \lambda + 1$ is the b-function of p and $d(\lambda) = \frac{d}{dx}$.

Example 5.1.5. *Let $p(x) = x_1^2 + \cdots + x_n^2$*
Then

$$\partial_i p(x)^{\lambda+1} = 2(\lambda+1)x_i p(x)^{\lambda}$$
$$\partial_i^2 p(x)^{\lambda+1} = 4\lambda(\lambda+1)x_i^2 p(x)^{\lambda-1} + 2(\lambda+1)p(x)^{\lambda}$$

Hence $d(\lambda) = \sum_{i=1}^{n} \partial_i^2$ satisfies

$$\begin{aligned} d(\lambda)p(x)^{\lambda+1} &= 4\lambda(\lambda+1)p(x)p(x)^{\lambda-1} + 2n(\lambda+1)p(x)^{\lambda} \\ &= 4(\lambda+1)(\lambda+\frac{n}{2})p(x)^{\lambda} \end{aligned}$$

Thus

$$b(\lambda) = (\lambda+1)(\lambda+\frac{n}{2})$$

is a b-function.

5.2 Solution of the Analytic Continuation Problem

Let $\phi(x) \in C_0^{\infty}(R^n)$ - a test function on $\mathbb{R}^n$. Consider a polynomial $p = p[x_1, \cdots, x_n] \in \mathbb{R}[x_1, \cdots, x_n]$ with nonnegative values on $\mathbb{R}^n$ and let $\lambda \in \mathbb{C}$, $Re\lambda \geq 0$ then p^{λ} as a continuous function in $\mathbb{R}^n$ thus it makes sense as a distribution on $\mathbb{R}^n$.

In fact, we define

$$\Gamma_\phi(\lambda) = \int_{\mathbb{R}^n} p(x)^\lambda \phi(x)dx \tag{5.2}$$

Then $\Gamma_\phi(\lambda) \in \mathcal{D}'$ and $\lambda \longrightarrow \Gamma_\phi(\lambda)$ is holomorphic family on $\{\lambda \in \mathbb{C} | \mathcal{R}e\lambda \geq 0\}$

Corollary 5.2.1. *Let $\Gamma_\phi(\lambda) \in \mathcal{D}'$ has a meromorphic extension to all of $\mathbb{C}$ with poles in $\{\lambda | b(\lambda + j) = 0\}$ for some $j \in \mathbb{N}$.*

Proof. From

$$B(\lambda)p^\lambda = d(\lambda)p^{\lambda+1}$$

By partial integration

$$\begin{aligned}
\Gamma_\phi(\lambda) =& \frac{1}{b(\lambda)} \int d(\lambda)p^{\lambda+1}\phi(x)dx \\
=& \frac{1}{b(\lambda)} \int p^{\lambda+1} d(\lambda)\phi(x)dx \\
=& \frac{1}{b(\lambda)} \frac{1}{b(\lambda+1)} \int p^{\lambda+2} d(\lambda+1)d(\lambda)\phi(x)dx \\
=& \cdots \\
=& \frac{1}{b(\lambda)} \frac{1}{b(\lambda+1)} \cdots \frac{1}{b(\lambda+N-1)} \int p^{\lambda+N} d(\lambda+N-1)\cdots d(\lambda+1)d(\lambda)\phi(x)dx
\end{aligned}$$

By taking N very large, we see that $\Gamma_\phi(\lambda)$ is meromorphically continued to the whole space $\lambda \in \mathbb{C}$ and that its poles are contained in $\{\alpha_j - k, j = 1, \cdots, m, k = 0, 1, 2, \cdots\}$ where α_j is any root of $b(\lambda)$. □

5.3 Conclusion

In this report, we have seen the treatment of purely analytic problem by algebraic means. This gives a beautiful and easy to follow solution compared to its analytic flavoured solution. This is exactly one of the reasons why we ought to study D-module theory. Generally, numerous results can be reformulated and proved in this manner.

The theory of algebraic D-modules was developed to provide an algebraic study of partial differential equation. In fact, as a motivation for the theory, we explained in chapter 1 how to make the switch from the usual point of view of systems of partial differential equations to the point of view of D-modules.

Most importantly, the theory of algebraic D-modules, becomes very interesting if one is familiar with basic knowledge of linear algebra, homological algebra, sheaf theory and analytic geometry; that is algebraic geometry and linear symplectic geometry; for instance, it is not easy to give a geometrical interpretation of the dimension of A_n-module, if one is not conversant with the language of algebraic geometry and linear symplectic geometry. Accordingly, using D-modules one could reformulate the famous Cauchy-Kovalevskaya theorem;on the existence and uniqueness of a local solution and prove it in a purely algebraic form after introducing the characteristic variety of a system. See the proof in [20].

Finally, we will wish to point out that the theory of D-modules, which is now often referred as "Algebraic Analysis", has grown steadily during the past years and important applications to other fields of mathematics and physics (e.g. singularities theory, group representations, Feynman integrals, e.t.c) were developed. Therefore, the task facing a student who would like to work in these directions is quite impressive: the research literature is huge and many research areas are still open.

Bibliography

[1] S.C.Coutinho: *A primer of algebraic D-modules*, Cambridge University Press. **42**, (1995)

[2] Gunter R. Krause and Thomas H. Lenagan, *Growth of algebras and Gelfand- Kirillov dimension, revised ed.*,Graduate Studies in Mathematics, vol. 22, American Mathematical Society, Providence, RI, 2000.

[3] D.H. Griffel: *Applied Functional Analysis*, Mathematics and Application Ellis Horwood Series.**42**, (1981)

[4] J.E Björk: *Rings of Differential Operators*,Mathematical Library vol. 21 Amsterdam, (1979)

[5] Robert S. Schrichartz: *A guide to Distribution Theory and Fourier Transforms*, Studies in Advanced Mathematics, CRC Press.**42**, (1994)

[6] I. N. Bernstein:*Modules over a ring of differential operators. An investigation of the fundamental solutions of equations with constant coefficients,*,Funkcional. Anal. i Prilozen. 5 (1971), no.2, 1-16.

[7] I. N. Bernstein and S. I. Gelfand: *Meromorphy of the function* P^{λ}, Funtioonal. Anal. i Prilozen. 3 (1969), no. 1,84-85.

[8] I. N. Bernstein :*The analytic continuation of generalised functions with parameter*, Funtional.Anal. and Appl. vol.6 (1972), no.4,26-40

[9] Masaki Kashiwara:*B-functions and holonomic systems. Rationality of roots of b-functions*, Invent. Math. 38 (1976/77), no. 1, 33Ů53. MR MR0430304 (55 #3309)

[10] I.M. Gelfand and G.E. Shilov:*Generalised functions*. Vol.1, Academic Press, New York 1964.

[11] M.F. Atiyah:*Resolution of singularities and division of distributions* Comm. Pure and App. Mathematics 23,1970.

[12] Anton Leykin:*Introduction to the algorithmic D-module theory* Institute for Mathematics and its Applications,August 2006.

[13] M. Kashiwara:*D-modules and microlocal Calculus translated from 2000 Japanesse original by Mutsumi Saito.* Translation of Mathematic Monographs, 217, Iwanami Series of Modern Mathematics, AMS, Providence, RI,2003.

[14] Malcolm A. H. Maccallum and Alexander V. Mikhaillov:*Algebraic Theory of Differential Equations.*, London Mathematical Society Lecture Note Series: 357, Cambridge University Press, 2009

[15] Santanu Chakraborty:*Some Applications of Dirac's Delta Function in Statistics for More Than One Random Variable.*, Applications and Applied Mathematics: Vol. 3, No. 1 (June 2008) pp. 42-54

[16] Walter Rudin:*Functional Analysis.*, McGraw-Hill, Inc,. 1973.

[17] Ruel V. Churchill,James W. Brown and Roger F. Verhey:*Complex Variable and Applications.*, McGraw-Hill, Inc,. 1976.

[18] Ricardo Estrada and Ram P. Kanwal :*A Distributional Theory for Asymptotics Theory and Applications.*, Birkhauser Boston, 2002.

[19] William R. Wade :*An Introduction to Analysis.* Prentice Hall, New Jersey 1995.

[20] Masaki Kashiwara :*Algebraic study of systems of partial differential equations.* Master Thesis, Tokyo University, December 1970. Translation by Andrea D'Agnolo and Jean-Pierre Schneiders, Memoires 65 de la S.M.F, 1995.

Index