ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea Magistrale in Matematica

Kernel approximations in Lie groups and application to group-invariant CNN

Tesi di Laurea in Analisi Matematica

Relatore:
Chiar.ma Prof.ssa
Giovanna Citti

Correlatori:
Chiar.mi Prof.
Remco Duits
Mattia Francesco Galeotti

Presentata da:
Giulia Lupi

Anno Accademico 2020-2021

Contents

Chapter 1

Introduction

The aim of this thesis is to study the following convection diffusion dilation/erosion equation in a Lie group:

$$\begin{cases} \frac{\partial W}{\partial t}(p,t) = -cW(p,t) - (-\Delta_{\mathcal{G}_1})^{\alpha} W(p,t) \pm \parallel \nabla_{\mathcal{G}_2} W(p,t) \parallel_{\mathcal{G}_2}^{2\alpha} & for\ p \in G/H, t \geq 0 \\ W(p,0) = U(p) & for\ p \in G/H. \end{cases} \tag{1.1}$$

This equation has been introduced by R.Duits,B.Smets, E.Bekkers and J.Portegies in [1] in order to solve image processing problems through PDE-based Group Equivariant Convolutional Neural Networks.

In chapter 2 we recall notions of sub-Riemannian theory. In particular, we prove the connectivity property which allow us to define a distance if the Hörmander condition is satisfied. We then suppose the Hörmander condition is satisfied and we define the notion of sub-Riemannian Laplacian operator and sub-Riemannian heat operator; with this hypothesis these operators turn out to be subelliptic.

In chapter 3 we consider equation (1.1) introduced by R.Duits,B.Smets, E.Bekkers and J.Portegies in [1] in order to solve image processing problems. In particular, since this equation is defined on a homogeneous space we recall notions on groups and homogeneous spaces. We then analyze each part of the equation separately; the convection part of the PDE is solved by a translation of the initial condition. Instead, the solution of the fractional diffusion PDE is written as a linear convolution of the fractional diffusion kernel with the initial condition. We then analyze the dilation/erosion part of the PDE: the solution is written as a morphological convolution of the dilation/erosion kernel with the initial condition. In order to implement these equations and apply them in image processing, it will be necessary to provide explicitly expressions of these kernels. First we propose to make the approximation of the fractional diffusion kernel through the parametrix method.

In chapter (4) we analyze the parametrix method for operators defined on $\mathbb{R}^n$. First of all we recall the notion and the properties of the foundamental solution of a parabolic operator with constant coefficients. We then consider a parabolic operator with Hölder-continuous coefficients, whose foundamental solution is found using the parametrix method. In particular, we write the foundamental solution as a convolution of a convergent series with the foundamental solution of the operator with constant coefficients.

In chapter (5) we extend the parametrix method for operators defined on a Lie group where equation (1.1) is defined. We will approximate the foundamental solution of a parabolic operator on this group with the foundamental solution defined on a homogeneous Carnot group, always with the parametrix method. As a result the foundamental solution will be expressed as a convolution of a convergent series whose terms are iteration of the foundamental solution of the parabolic operator defined on a Carnot group.

In chapter (6) we consider equation (1.1) in order to do approximations of the kernels found in chapter (3). The approximate fractional diffusion kernel for $\alpha = 1$ is found through the parametrix method. For a generic value of α we then consider a key relation between the $\alpha-$kernels and the Laplacian kernel. For the dilation/erosion equation we first prove a foundamental relation between the dilation/erosion kernel and the fractional diffusion kernel: in order to do that we need to introduce the notion of Cramér-Fourier transform. We then find the approximate dilation/erosion kernel starting from the fractional diffusion kernel.

In chapter (7) we consider a possible application to image processing problems of the equation discussed within the thesis. In particular we describe the PDE-based Group Equivariant Convolutional Neural Network defined by R.Duits,B.Smets, E.Bekkers and J.Portegies in [1].

Chapter 2

Sub-riemannian manifolds

The aim of this chapter is to describe the roto-translation group $SE(2)$ with a sub-Riemannian metric on it and to clarify the role in image processing and in modelling the visual cortex.

The chapter is organized as follows: in the first section we give the main definitions on Lie groups. These notions will serve us in the second section, where we will give the definition of sub-Riemannian manifold and we will prove the Hörmander theorem. Finally, in the third section we will explain the model of the visual cortex as the $SE(2)$ group.

2.1 The notion of Lie Group

In this first section we give some basic definitions on Lie groups. These are well known notions of Lie group theory, that can be found for example in [12].

Definition 2.1.1. An internal operation $\cdot$ on a set G is a function which associates to each couple of elements (a, b) of G another element denoted as $a \cdot b$ of G. This condition is also called closure: for all $a, b \in G$, $a \cdot b \in G$. To qualify the set as a group, the operation must satisfy three requirements known as the group axioms:

- Associativity. For all $a, b, c \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- Identity element. There exists an element $e \in G$, such that for all $a \in G$, $e \cdot a = a \cdot e = a$.
- Inverse element. For all $a \in G$, there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$, where e is the identity element.

If G is a group and $\forall a, b \in G$, $a \cdot b = b \cdot a$, then G is called an Abelian or commutative group. A subgroup is a subset G' of G satisfying the group axioms.

The distinguishing feature of a Lie group is that it also carries the structure of a smooth manifold, so that the group operation is continuous.

Definition 2.1.2. A Lie group is a group which also carries the structure of a manifold in such a way that both the group operation

$$m : G \times G \to G, \quad m(g,h) = g \cdot h, \quad g,h \in G \tag{2.1}$$

and the inversion

$$i : G \to G \quad i(g) = g^{-1} \quad g \in G \tag{2.2}$$

are smooth maps between manifolds.

Examples of Lie groups are:

- The Euclidean space $\mathbb{R}^n$, with the usual sum as group law
- The circle S^1 of angles mod 2π, with the standard sum of angles

Definition 2.1.3. Let C be a smooth curve on a manifold $\mathcal{M}$, parametrized by $\gamma : I \to \mathcal{M}$, where I is an interval of $\mathbb{R}$. In local coordinates $x = (x^1, ..., x^n)$, C is given by n smooth functions $\gamma(s) = (\gamma_1(s), ..., \gamma_n(s))$ of the real variable s. At each point $x = \gamma(s)$ of C, the curve has a tangent vector

$$X \mid_x = \gamma'(s) = \gamma_1'(s)\frac{\partial}{\partial x^1} + ... + \gamma_n'(s)\frac{\partial}{\partial x^n}. \tag{2.3}$$

By this definition it becomes clear that tangent vectors can be identified with directional derivatives in the direction:

$$\vec{X} \mid_x = (\gamma_1', ..., \gamma_n')$$

In particular if we apply the tangent vector X to a smooth function $f : \mathcal{M} \to \mathbb{R}$ we obtain:

$$Xf = \langle \vec{X}, \nabla f \rangle$$

Within the thesis we will then identify differential operators and vectors, we will see this fact in details in the next section.

Definition 2.1.4. The collection of all possible tangent vectors to all possible tangent curves passing through a point $x \in \mathcal{M}$ is called the tangent space to $\mathcal{M}$ at x, and is denoted as $T\mathcal{M} \mid_x$.

Definition 2.1.5. An integral curve of a vector field $\vec{X}$ is a smooth parametrized curve $x = \gamma(s)$ whose tangent vector at any point coincides with the value of $\vec{X}$ at the same points for all s:

$$\gamma'(s) = \vec{X} \mid_{\gamma(s)}$$

An integral curve is called maximal integral curve if it is not contained in a longer integral curve.

Definition 2.1.6. For two differential operators X and Y, their Lie bracket (or commutator) is defined by their actions on functions $f : \mathcal{M} \to \mathbb{R}$:

$$[X,Y]f = X(Yf) - Y(Xf). \tag{2.4}$$

Note that the Lie Bracket is a measurement of the non-commutativity of the operators: it is defined as the difference of applying first Y and then X and vice versa. In particular $[X,Y]$ is identically 0 if X and Y commute.

Definition 2.1.7. Let G be a Lie group. For any element $g \in G$, we define the left-multiplication (or left-translation) $L_g : G \to G$ by:

$$L_g(h) = g \cdot h, \quad \forall h \in G \tag{2.5}$$

where $\cdot$ denotes the group operation in G

Definition 2.1.8. An operator X on G is called left-invariant if:

$$X(f \circ L_g) = (Xf) \circ L_g, \quad \forall g \in G \tag{2.6}$$

Definition 2.1.9. The Lie algebra of a Lie group G is the vector space of all left-invariant vector fields on G

Intuitively the Lie algebra associated to a Lie group encodes its differential structure, and it is identified as the tangent space at the origin.

2.2 Hörmander vector fields and sub-Riemannian manifolds

In this section we will prove the connectivity property, which states that, if $X_1, ..., X_m$ are Hörmander vector fields defined on a manifold M, then every couple of points can be connected with an admissible integral curve. The proof follows the approach of [18]. This theorem will permit us to define a distance on a sub-Riemannian manifold. Before stating the theorem we need to give some basic definitions on sub-Riemannian theory.

Definition 2.2.1. A Riemannian metric on a differentiable manifold M is given by a scalar product on each tangent space T_pM which depends smoothly on the base point p. A Riemannian manifold is a differentiable manifold, equipped with a Riemannian metric.

Example 2.2.2. *In order to understand the concept of a Riemannian metric, we again need to study local coordinate representations and the transformation behavior of these expressions.*

Thus, let $x = (x^1, ..., x^d)$ be local coordinates. In these coordinates, a metric is represented by a positive definite, symmetric matrix $(g_{ij}(x))_{i,j=1,...,d}$ where the coefficients depend smoothly on x.
The product of two tangent vectors $v, w \in T_pM$ with coordinate representations $(v^1, ..., v^d)$ and $(w^1, ..., w^d)$ is

$$\langle v, w\rangle_g = g_{ij}(x(p))v^i w^j \tag{2.7}$$

and

$$\|v\|_{g(x)} = \sqrt{\langle v, v\rangle_g} \tag{2.8}$$

Definition 2.2.3. Let M be a Riemannian manifold equipped with a Riemannian metric g. Let $[a, b]$ be a closed interval in $\mathbb{R}$, $\gamma : [a, b] \to \mathbb{R}$ a curve of class C^1.The length of γ then is defined as

$$L(\gamma) = \int_a^b \|\gamma'(t)\|_{g(\gamma(t))} \tag{2.9}$$

Definition 2.2.4. On a connected and differentiable manifold the distance between two points $p, q \in M$ can be defined as:

$$d(p, q) = inf\{L(\gamma) : \gamma : [a, b] \to M \; piecewise \; smooth \; curve \; with \; \gamma(a) = p, \gamma(b) = q\} \tag{2.10}$$

Note that M is connected, hence $d(p, q) < \infty \; \forall p, q \in M$.

We call vector field a map $X : \Omega \to \cup_p T_p$ such that $X \mid_p \in T_pM$ for all $p \in M$. In the sequel we will always identify vector fields and differential operators as follows. Let $X = (a_1, ..., a_n)$ be a vector field of class $C^\infty(\Omega, \mathbb{R}^n)$, $\Omega \subseteq \mathbb{R}^n$.
If X is a smooth first order differential operator, $X = \sum_{i=1}^n a_i\partial_i$ and I is the identity map $I(\xi) = \xi$, then it is possible to represent the vector field with the same components as the differential operator X in the form

$$XI(\xi) = (a_1, ..., a_n) \tag{2.11}$$

Definition 2.2.5. Let M be a differentiable manifold. We call distribution Δ a set $(\Delta_p)_{p\in M}$ such that $\forall p \; \Delta_p$ is a subspace of T_pM.

Notice that the definition of distribution implies that Δ is subbundle of the tangent bundle. If $(X_1, ..., X_n)$ are differential operators then
$\forall p \in M \; \Delta_p = \langle X_1 I \mid_p, ..., X_n I \mid_p\rangle$ define the distribution $\Delta = (\Delta_p)_{p\in M}$.
Note that T_pM has the same dimension $\forall p \in M$ while Δ_p and Δ_q, with $p, q \in M \; p \neq q$ have not necessarily the same dimension.

Definition 2.2.6. Let M be a connected and differentiable manifold, Δ a distribution $\Delta \subseteq TM$ and g a metric defined on Δ. Then (M, g, Δ) is called sub-Riemannian manifold.

The problem now is how to define the length of a curve defined on a sub-Riemannian manifold. We then have to distinguish between admissible and non admissible curves.

Definition 2.2.7. Let $\gamma : [a, b] \to M$ be a smooth curve. Then γ is called an admissible curve if $\forall t \in [a, b]$ it holds $\gamma'(t) \in \Delta_{\gamma(t)}$.

Definition 2.2.8. Let $\gamma : [a, b] \to M$ be an admissible curve. We call

$$l(\gamma) = \int_a^b \parallel \gamma'(t) \parallel_{g(\gamma(t))}$$

the length of γ

Notice that $\gamma'(t) \in \Delta_{\gamma(t)}$ then it follows from the definition of sub-Riemannian manifold that $\parallel \gamma'(t) \parallel_{g(\gamma(t))}$ exists. The difference with Riemannian manifolds is that on a sub-Riemannian manifold it is not always true that any couple of points can be connected. Hence, the distance can be defined as follows

Definition 2.2.9. Let (M, Δ, g) be a sub-Riemannian manifold and let $x, y \in M$. If it exists at least one admissible curve that connects x and y we call

$$d(x, y) = \inf\{l(\gamma) : \gamma : [a, b] \to M \textit{ is admissible}, \gamma(a) = x, \gamma(b) = y\} \quad (2.12)$$

the distance between x and y. If it doesn't exist any admissible curve that connects x and y we set

$$d(x, y) = +\infty$$

Remark 2.2.10. *If we have* $X = \sum_{i=0}^n a_i \partial_{x_i}$, $Y = \sum_{j=0}^n b_j \partial_{x_j}$ *and* $f \in C_c^\infty(\mathbb{R}^n)$ *then the explicit computation of the commutator is:*

$$\begin{aligned}
[X, Y]f = (XY - YX)f &= \sum_{i,j=0}^n a_i \partial_{x_i}(b_j \partial_{x_j} f) - \sum_{i,j=0}^n b_j \partial_{x_j}(a_i \partial_{x_i} f) = \\
&= \sum_{i,j=0}^n a_i \partial_{x_i} b_j \partial_{x_j} f + \sum_{i,j=0}^n a_i b_j \partial^2_{x_i, x_j} f - \sum_{i,j=0}^n b_j \partial_{x_j} a_i \partial_{x_i} f - \sum_{i,j=0}^n b_j a_i \partial^2_{x_i, x_j} f = \\
&= \sum_{j=0}^n (\sum_{i=0}^n a_i \partial_{x_i} b_j) \partial_{x_j} f - \sum_{i=0}^n (\sum_{j=0}^n b_j \partial_{x_j} a_i) \partial_{x_i} f = \\
&= \sum_{j=0}^n (X b_j \partial_{x_j} f - Y a_j \partial_{x_j} f)
\end{aligned} \quad (2.13)$$

Note that $[X, Y]$ *is a first order differential operator.*

Definition 2.2.11. If $X_1, X_2, ..., X_m$ are differential operators we call Lie algebra generated by $X_1, ..., X_m$ and denote it as

$$\mathcal{L}(X_1, ..., X_m) \tag{2.14}$$

the linear span of the operators $X_1, ..., X_m$ and their commutators of any order.
We will say that the vectors

$$X_1, ..., X_m \textit{ have degree } 1, \ deg(X_i) = 1 \tag{2.15}$$

$$\forall i, j \leq m : \ [X_i, X_j] \textit{ have degree } 2, \ deg([X_i, X_j]) = 2 \tag{2.16}$$

and define in an analogous way higher order commutators.

Example 2.2.12. *In general the degree is not unique. Indeed, if we consider the vector fields in* $\mathbb{R}^2 \times S^1$

$$X_1 = cos(\theta)\partial_1 + sin(\theta)\partial_2 \quad X_2 = \partial_\theta \tag{2.17}$$

the vector X_1 *has degree* 1*, but it also has degree* 3*, since in this specific example* $X_1 = -[X_2, [X_2, X_1]]$.

If the degree is not unique, we call degree of a vector field the minimum degree.

In the sequel we will consider the Lie group $M = \mathbb{R}^n$. Since $m < n$, in general $\mathcal{L}(X_1, ..., X_m)$ will not coincide with the tangent bundle. If these two spaces coincide, we will say that the Hörmander condition is satisfied.

Definition 2.2.13. Let $\Omega \subset \mathbb{R}^n$ be an open set, and let $(X_j)_{j=1,...,m}$ be a family of smooth vector fields defined on Ω. If $\forall \xi \in \mathbb{R}^n$ the condition

$$rank(\mathcal{L}(X_1, ..., X_m))(\xi) = n, \tag{2.18}$$

is satisfied, we say that the vector fields $(X_j)_{j=1,...,m}$ satisfy the Hörmander rank condition

If this condition is satisfied, at every point ξ we can find a number s such that $(X_j)_{j=1,...,m}$ and their commutators of degree smaller or equal to s span the space at ξ. At every point we can select a basis $\{X_j : \ j = 1, ..., n\}$ of the space made out of commutators of the vector fields $\{X_j : \ j = 1, ..., m\}$. In general the choice of the basis will not be unique, but we will choice a basis such that for every point

$$Q = \sum_{j=1}^{n} deg(X_j) \tag{2.19}$$

is minima. The value of Q is called local homogeneous dimension of the space.

Example 2.2.14. *Consider $X_1, ..., X_m$ vector fields defined on $\mathbb{R}^n$. The simpliest example of family of vector fields is the Euclidean one. In this case, if $m = n$, then the Hörmander condition is satisfied while it is violated if $m < n$.*

Example 2.2.15. *We denote by $\xi = (x_1, x_2, \theta)$ a point in $\mathbb{R}^2 \times S^1$. Using the notation of (2.17) X_1, X_2 are the generators of the Lie algebra. In fact their commutator is*

$$X_3 = [X_2, X_1] = -sin(\theta)\partial_1 + cos(\theta)\partial_2 \tag{2.20}$$

and X_1, X_2, X_3 are linearly indipendent and generate $\mathbb{R}^2 \times S^1$.

The lemmas that follow will allow us to prove the connectivity property. First of all we need to define what is an integral curve.

Definition 2.2.16. If X is a smooth first order differential operator, we call integral curve of the vector field XI starting at ξ_0 a curve γ such that

$$\begin{cases} \gamma'(t) = (XI)(\gamma(t)) \\ \gamma(0) = \xi_0 \end{cases}$$

The curve will also be denote by $\gamma(t) = exp(tX)(\xi_0)$

Lemma 2.2.17. *Let X be of class C^∞ and f of class C^1. If $\gamma(t) = exp(tX)(\xi_0)$, then $\exists (f \circ \gamma)'(t) = (Xf)(\gamma(t))$*

Proof: If $X = \sum_{i=1}^n a_i \partial_i$ then

$$\frac{d}{dt}(f\circ\gamma)(t) = \langle \nabla f(\gamma(t)), \gamma'(t)\rangle = \langle \nabla f(\gamma(t)), XI(\gamma(t))\rangle = \sum_{i=0}^n a_i\partial_i f(\gamma(t)) = Xf(\gamma(t)). \tag{2.21}$$

□

Lemma 2.2.18. *Let X be of class C^∞, then*

$$exp(tX)(\xi) = \xi + t(XI)(\xi) + \frac{t^2}{2}X^2 I(\xi) + o(t^2) \quad t \to 0 \tag{2.22}$$

In general we have:

$$exp(tX)(\xi) = \xi + t(XI)(\xi) + \frac{t^2}{2}X^2 I(\xi) + ... + \frac{t^h}{h!}X^h I(\xi) + o(t^h) \tag{2.23}$$

Proof: We have $\gamma(t) = exp(tX)(\xi)$, then $\gamma(0) = \xi$.
$\gamma'(t) = (XI)(\gamma(t))$, then $\gamma'(0) = (XI)(\xi)$.
$\gamma''(t) = \frac{d}{dt}(XI \circ \gamma)(t) = X^2 I(\gamma(t))$, then $\gamma''(0) = X^2 I(\gamma(0)) = X^2 I(\xi)$.

$\gamma^h(t) = (X^h I)(\gamma(t))$, then $\gamma^h(0) = X^h I(\xi)$.
The Taylor expansion ensures that

$$\gamma(t) = \gamma(0) + t\gamma'(0) + \frac{t^2}{2}\gamma''(0) + o(t^2) \quad t \to 0 \tag{2.24}$$

Hence

$$exp(tX)(\xi) = \xi + t(XI)(\xi) + \frac{t^2}{2}(X^2 I)(\xi) + o(t^2) \quad t \to 0 \tag{2.25}$$

□

Definition 2.2.19. Let X be a first order differential operator. We call Lie derivative of f in the direction of the vector X on the tangent space to $\mathbb{R}^n$ at a point ξ the derivative with respect to t in $t = 0$ of the function $f \circ exp(tX)(\xi)$.

Lemma 2.2.20. *Let f be of class C^k and X of class C^∞, then*

$$f(exp(tX)(\xi)) = f(\xi) + t(Xf)(\xi) + ... + \frac{t^k}{k!}X^k f(\xi) + o(t^k) \tag{2.26}$$

Proof: We prove the thesis for $k = 2$, the proof for $k > 2$ is analogous.
$\gamma(t) = exp(tX)(\xi)$. We have: $(f \circ \gamma)(0) = f(\gamma(0)) = f(\xi)$, $(f \circ \gamma)'(t) = Xf(\gamma(t)) \Rightarrow (f \circ \gamma)'(0) = (Xf)(\xi)$.
$(f \circ \gamma)''(t) = (Xf \circ \gamma)'(t) = (X^2 f)(\gamma(t)) \Rightarrow (f \circ \gamma)''(0) = (X^2 f)(\xi)$.
If f is of class C^h then $(f \circ \gamma)^h(0) = (X^h f)(\xi)$.
If f is of class C^2 and we consider the second order Taylor expansion:

$$(f \circ \gamma)(t) = (f \circ \gamma)(0) + t(f \circ \gamma)'(0) + \frac{t^2}{2}(f \circ \gamma)''(0) + o(t^2) \tag{2.27}$$

Hence:

$$f(exp(tX)(\xi)) = f(\xi) + t(Xf)(\xi) + \frac{t^2}{2}X^2 f(\xi) + o(t^2) \tag{2.28}$$

□

Theorem 2.2.21. *Let X, Y be of class C^2,then:*

$$exp(-tY)exp(-tX)exp(tY)exp(tX)(\xi) = \xi + t^2[X,Y] + o(t^2) \tag{2.29}$$

Proof: $exp(tX)(\xi) = \xi + t(XI)(\xi) + \frac{t^2}{2}X^2 I(\xi) + o(t^2)$.
Pose $\eta = exp(tX)(\xi)$. First of all, we compute:

$$\begin{aligned} exp(tY)(exp(tX)(\xi)) &= exp(tY)(\eta) = \eta + tYI(\eta) + \frac{t^2}{2}Y^2 I(\eta) + o(t^2) = \\ =&\xi + tXI(\xi) + \frac{t^2}{2}X^2 I(\xi) + o(t^2) + tYI(exp(tX)(\xi)) + \frac{t^2}{2}Y^2 I(exp(tX)(\xi)) + o(t^2) = \\ =&\xi + tXI(\xi) + \frac{t^2}{2}X^2 I(\xi) + o(t^2) + t(YI(\xi) + tXYI(\xi) + o(t)) + \frac{t^2}{2}Y^2 I(exp(tX)(\xi)) + o(t^2) \end{aligned} \tag{2.30}$$

Note that

$$\frac{t^2}{2}Y^2I(exp(tX)(\xi)) = \frac{t^2}{2}Y^2I(\xi + o(1)) = \frac{t^2}{2}Y^2I(\xi)$$

Hence

$$exp(tY)(exp(tX)(\xi)) = \xi + t(X+Y)I(\xi) + t^2(\frac{X^2}{2}I(\xi) + XYI(\xi) + \frac{Y^2}{2}I(\xi)) + o(t^2) \tag{2.31}$$

In particular

$$exp(tY)(exp(tX)(\xi)) = \xi + t(X+Y)I(\xi) + o(t) = exp(t(X+Y))(\xi) + o(t)$$

Pose $\nu = epx(tY)exp(tX)(\xi)$.Now we compute:

$$\begin{aligned}
&exp(-tX)exp(tY)exp(tX)(\xi) = exp(-tX)(\nu) = \nu - tXI(\nu) + \frac{t^2}{2}X^2I(\nu) + o(t^2) = \\
&= \xi + tXI(\xi) + tYI(\xi) + \frac{t^2}{2}X^2I(\xi) + t^2XYI(\xi) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) - tXI(exp(t(X+Y)(\xi)) + \\
&o(t)) + \frac{t^2}{2}X^2I(\xi + o(1)) = \\
&= \xi + tXI(\xi) + tYI(\xi) + \frac{t^2}{2}X^2I(\xi) + t^2XYI(\xi) + \frac{t^2}{2}Y^2I(\xi) - t(XI(\xi) + t(X+Y)XI(\xi) + \\
&+ o(t)) + \frac{t^2}{2}X^2I(\xi) + o(t^2) = \\
&= \xi + tYI(\xi) + \frac{t^2}{2}X^2I(\xi) + t^2XYI(\xi) + \frac{t^2}{2}Y^2I(\xi) - t^2X^2I(\xi) - t^2YXI(\xi) + \frac{t^2}{2}X^2I(\xi) + o(t^2) = \\
&= \xi + tYI(\xi) + t^2[X,Y]I(\xi) + \frac{t^2}{2}Y^2I(\xi) + o(t^2)
\end{aligned} \tag{2.32}$$

In particular

$$exp(-tX)exp(tY)exp(tX) = \xi + tYI(\xi) + o(t)$$

Finally, we compute:

$$\begin{aligned}
&exp(-tY)exp(-tX)exp(tY)exp(tX) = exp(-tY)(\eta) = \eta - tYI(\eta) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) = \\
&= \xi + tYI(\xi) + t^2[X,Y]I(\xi) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) - tYI(exp(tY)(\xi) + o(t)) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) = \\
&= \xi + tYI(\xi) + t^2[X,Y]I(\xi) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) - tYI(\xi) - t^2Y^2I(\xi) + \frac{t^2}{2}Y^2I(\xi) + o(t^2) = \\
&= \xi + t^2[X,Y]I(\xi) + o(t^2)
\end{aligned} \tag{2.33}$$

□

We first observe that

$$exp(-tY)exp(-t^2[X,Y])exp(tY)exp(t^2[X,Y]) = \xi + t^3[[X,Y],Y] + o(t^3) \tag{2.34}$$

$exp(-t^2[X,Y])$ is obtained composing $exp(tX)$ and $exp(tY)$. Then, we can conclude that, composing together $exp(tX)$ and $exp(tY)$ we can find increments in the direction of the commutators of any degree.

Definition 2.2.22. If $deg(X) = 1$, we define

$$C_X(t) = exp(tX)(\xi) = \begin{cases} C'_X(0) = (XI)(\xi) \\ C_X(0) = \xi \end{cases}$$

Note that, if we define

$$C_{[X,Y]}(t) = exp(-tY)exp(-tX)exp(tY)exp(tX)(\xi) = \xi + t^2[X,Y](\xi) + o(t^2)$$

then $C'_{[X,Y]}(t) = 0$ when $t = 0$.

Hence, in order to mantain the homogeneity, we have to give a different definition of integral curves for differential operators of degree higher than 1. We start by giving this definition for vector fields of degree 2 and then generalize it.

Definition 2.2.23. If X is of class C^∞ and $deg(X) = 2$, $X = [X_1, X_2]$, we define:

$$C_X(t) = \begin{cases} C_{X_2}(-\sqrt{|t|})C_{X_1}(-\sqrt{|t|})C_{X_2}(\sqrt{|t|})C_{X_1}(\sqrt{|t|})(\xi) & if\ t > 0 \\ C_{X_2}(-\sqrt{|t|})C_{-X_1}(-\sqrt{|t|})C_{X_2}(\sqrt{|t|})C_{-X_1}(\sqrt{|t|})(\xi) & if\ t < 0 \end{cases}$$

Theorem 2.2.24. *$C_X(0) = \xi$ and $C'_X(0) = XI(\xi)$*

Proof:

$$C_X(t) = \begin{cases} C_{X_2}(-\sqrt{|t|})C_{X_1}(-\sqrt{|t|})C_{X_2}(\sqrt{|t|})C_{X_1}(\sqrt{|t|})(\xi) & if\ t > 0 \\ C_{X_2}(-\sqrt{|t|})C_{-X_1}(-\sqrt{|t|})C_{X_2}(\sqrt{|t|})C_{-X_1}(\sqrt{|t|})(\xi) & if\ t < 0 \end{cases} =$$

$$= \begin{cases} \xi + |t|\,[X_1, X_2]I(\xi) + o(t) & if\ t > 0 \\ \xi - |t|\,[X_1, X_2]I(\xi) + o(t) & if\ t < 0 \end{cases} = \xi + t[X_1, X_2]I(\xi) + o(t) \Rightarrow$$

$$\Rightarrow \begin{cases} C'_X(0) = [X_1, X_2]I(\xi) = XI(\xi) \\ C_X(0) = \xi \end{cases}$$

□

Definition 2.2.25. If $deg(X) = j$ and $X = [X_1, X_2]$ with $deg(X_1) = 1$ and $deg(X_2) = j - 1$, we define:

$$C_X(t)(\xi) = \begin{cases} C_{X_2}(-|t|^{1/j})C_{X_1}(-|t|^{1/j})C_{X_2}(|t|^{1/j})C_{X_1}(|t|^{1/j})(\xi) & if\ t > 0 \\ C_{X_2}(-|t|^{1/j})C_{-X_1}(-|t|^{1/j})C_{X_2}(|t|^{1/j})C_{-X_1}(|t|^{1/j})(\xi) & if\ t < 0 \end{cases}$$

Hence:

$$\begin{cases} C'_X(0) = (XI)(\xi) \\ C_X(0) = \xi \end{cases}$$

As we have said before, the span of a family of Hörmander vector fields generate the manifold M. Then it is a basis of the space: the following definition gives a particular way to construct this basis.

Definition 2.2.26. Let $X_1, ..., X_m$ be Hörmander vector fields. Then we can use them to define a basis. Consider $X_1, ..., X_m$ vector fields of degree 1 and complete them with a basis of $\Delta^{(2)} = span\{X_1, ..., X_m, [X_i, X_j] : i, j = 1, ..., m\}$ of degree 2. Repeat the same procedure with vector fields of increasing degree. At the end in a finite number of iterations we will obtain a basis:

$$\underbrace{X_1, ..., X_m}_{Degree\ 1}, \underbrace{X_{m+1}, ..., X_{dim(\Delta^{(2)})}}_{Degree\ 2}, ..., \underbrace{X_{dim(\Delta^{(k)})}}_{Degree\ k} \tag{2.35}$$

This is called an adapted basis.

Definition 2.2.27. Consider γ a curve defined on the subriemannian manifold (M, Δ, g). We call γ an admissible curve if and only if γ is an integral curve of a vector field of degree 1.

We now want to prove that, if $X_1, ...X_m$ are Hörmander vector fields, then it is possible to connect every two points in the subriemannian manifold with admissible integral curves. First of all we have to prove the following lemma

Lemma 2.2.28. *Let (M, Δ, g) be a sub-Riemannian manifold and let$X_1, ..., X_n$ be an adapted basis. If $s \in \mathbb{R}^n, \xi \in M$ consider*

$$\hat{C}(s)(\xi) = C_{X_1}(s_1)C_{X_2}(s_2)...C_{X_n}(s_n)$$

, where $s = (s_1, ..., s_m)$.
Then $\hat{C}$ is a local diffeomorphism around the point $s = 0$.

Proof: Consider $\frac{\partial \hat{C}}{\partial s_j}(0)$. Let j be fixed, then

$$\hat{C}(0, ..., s_j, ..0)(\xi) = C_{X_j}(s_j)(\xi)$$

If we now consider the derivative with respect to s_j we obtain:

$$\frac{\partial \hat{C}}{\partial s_j}(0)(\xi) = \frac{d}{ds_j} C_{X_j}(0)(\xi) = X_j I(\xi)$$

Then the Jacobian matrix of $\hat{C}$ in $s = 0$ is

$$J\hat{C}(0)(\xi) = (X_1 I(\xi), ..., X_n I(\xi))$$

where $X_1, ..., X_n$ is a basis. This implies that $det(J\hat{C}(0)(\xi)) \neq 0$, then $\hat{C}$ is a local diffeomorphism around 0. This completes the proof. □

The previous lemma implies that it exists a neighborhood U of 0 and a neighborhood V of ξ such that

$$\forall \eta \in V, \exists ! s \in U : \eta = \hat{C}(s) \tag{2.36}$$

This means that it exists an admissible integral curve which connects ξ and η.

We can now prove the following connectivity property

Theorem 2.2.29. *Let (M, Δ, g) be a connected sub-Riemannian manifold whose basis satisfies the Hörmander condition. Then $\forall \xi, \xi_0 \in M$ it exists an admissible integral curve on M which connects ξ and ξ_0.*

Proof: Let ξ_0 be fixed and consider:

$$A = \{\xi : \exists \gamma \textit{ admissible integral curve of extremes } \xi, \xi_0\} \tag{2.37}$$

We want to prove that A is an open and a closed subset of M. This fact will imply, thanks to tha fact that M is connected ,that $A = M$.
First of all we verify that A is an open subset of M. Consider $\xi \in A$, then (2.36) implies that it exists a neighborhood U of 0 and a neighborhood V of ξ such that for every $\eta \in V$, $\eta = \hat{C}(s)(\xi), s \in U$. This means it exists an admissible integral curve γ_η which connects η and ξ.
Notice that we have chosen $\xi \in A$, then it exists an admissible integral curve γ which connects ξ and ξ_0. If we compose together γ and γ_η we obtain an admissible integral curve which connects η and ξ_0, then $\eta \in A$.
We have then proved that $\forall \xi \in A, \exists V$ neighborhood of ξ such that $V \subseteq A$. Then A is an open subset of M.
Let's now prove that A is a closed subset of M.
Let (ξ_n) be a sequence in A such that $\xi_n \to \bar{\xi}$. We want to prove that $\bar{\xi} \in A$. Thanks to (2.36) we know that $\exists V$ neighborhood of $\bar{\xi}$ such that $\forall \eta \in V$ it exists an admissible integral curve $\hat{C}(s)(\xi)$ of extremes $\bar{\xi}, \eta$.
We have supposed $\xi_n \to \xi$, then $\exists \bar{n} : \forall n \geq \bar{n}, \xi_n \in V$. This implies that it exists an admissible integral curve $\bar{\gamma}$ which connects ξ_n and $\bar{\xi}$.
$\xi_{\bar{n}} \in A$ then it exists γ admissible integral curve of extremes $\xi_0, \xi_{\bar{n}}$. If we compose together γ and $\bar{\gamma}$ we obtain an admissible integral curve that connects ξ_0 and $\bar{\xi}$. Then $\bar{\xi} \in A$. This completes the proof. □

The connectivity property permits us to define a distance on the sub-Riemannian manifold.

Definition 2.2.30. Let (M, Δ, g) be a connected sub-Riemannian manifold, with Δ verifying the Hörmander condition. Then, if $\xi, \xi_0 \in M$ we define the distance between ξ and ξ_0 as

$$d(\xi, \xi_0) = \inf\{l(\gamma) : \gamma \textit{ is an admissible integral curve of extremes } \xi_0, \xi\} \tag{2.38}$$

Notice that

$$\forall \xi, \xi_0 \in M, d(\xi, \xi_0) < +\infty$$

We end the discussion on the distance defined on a sub-Riemannian manifold giving an estimation of the distance between two points which are connected with not only admissible integral curves.

Definition 2.2.31. If $\xi_0 \in M$ is fixed, we define the canonical coordinates of ξ around ξ_0 as the coefficients t such that

$$\xi = exp(\sum_{j=1}^{n} t_j X_j)(\xi_0)$$

Theorem 2.2.32. *Let $\xi_0 \in M$ be fixed and consider V a neighborhood of ξ_0. Then for every $\xi \in V$ there exist two constants $c_1, c_2 > 0$ such that*

$$c_1 \sum_{i=1}^{n} |t_i|^{1/deg(X_i)} \leq d(\xi, \xi_0) \leq c_2 \sum_{i=1}^{n} |t_i|^{1/deg(X_i)} \tag{2.39}$$

Proof: We want to prove that $E(t) = exp(t_1 X_1 + ... + t_n X_n)(\xi_0)$ is a local diffeomorphism, with $t = (t_1, ..., t_n)$.
First of all notice that $E(0) = \xi_0$ and that if we consider the derivative $\frac{\partial E}{\partial t_i}(0)$ this is equivalent to consider the derivative of $E(0, ...t_i, ...0)$ in $t_i = 0$. This fact yields:

$$\frac{d}{dt_i} E(0, .., t_i, ..0)(\xi_0) = \frac{d}{dt_i} exp(t_i X_i)(\xi_0) = X_i I(exp(t_i X_i)(\xi_0))$$

where the last equality follows from definition (2.2.16).
Then for $t_i = 0$ we obtain $\frac{\partial}{\partial t_i} E(0) = X_i I(\xi_0)$, where $X_1, ..., X_n$ is a basis of M. This implies $(\frac{\partial E}{\partial t_i}(0))_{i=1,...,n}$ are linearly indipendent, then $det(J_E(0)) = det(\frac{\partial E}{\partial t_1}, ..., \frac{\partial E}{\partial t_n}) \neq 0$. This completes the proof. □

We end this section by giving the main definitions and results on differential calculus in a sub-Riemannian setting. Clearly, if f is C^1, then the Lie derivative coincides with the directional derivative. Then, it is a more general definition: the Lie derivative can exist even though the directional derivative doesn't exist.

Definition 2.2.33. Let $M \subset \mathbb{R}^n$ be an open set, let $(X_j), j = 1, ..., m$ be a family of smooth vector fields defined on M, and let $f : M \to \mathbb{R}$. If there exists $X_j f$ for every $j = 1, ..., m$ we call horizontal gradient of the function f with respect to this set of vectors

$$\nabla_H f = (X_1 f, ..., X_m f)$$

Definition 2.2.34. Suppose the same hypothesis of the previous definition. A function f is of class C^1_H if $\nabla_H f$ is continuous with respect to the distance defined in (2.38). f is of class C^2_H if $\nabla_H f$ is of class C^1_H.Analougously all C^k_H classes are defined.

Note that a C^1_H function is not differentiable with respect to X_j if $j > m$. It follows that a function of class C^1_H is not of class C^1_E, in the standard Euclidean sense.

Remark 2.2.35. *If the vector fields* (X_j),$j = 1, ..., m$ *satisfy the Hörmander condition,* f *is* C^∞_H *if and only if is of class* C^∞_E *in the standard sense.*

From the definition of Lie derivative and the properties of integral curves, the following result follows:

Theorem 2.2.36. *Let* $M \subset \mathbb{R}^n$*, let* X *and* Y *be continuous vector fields defined on* M *and let* $f : M \to \mathbb{R}$*. Assume that at every point* $\xi \in M$ *there exists* $Xf(\xi)$ *and* $Y f(\xi)$*, and these derivatives are continuous. If* $\gamma(t) = exp(tX)(exp(tY)(\xi))$*, then there exists*

$$(f \circ \gamma)'(0) = Xf(\xi) + Y f(\xi)$$

Proof:

$$\begin{aligned} &\frac{1}{t}(f(\gamma(t)) - f(\gamma(0))) = \\ &\qquad = \frac{1}{t}(f(exp(tX)(exp(tY)(\xi))) - f(exp(tY)(\xi)))+ \\ &\qquad + \frac{1}{t}(f(exp(tY)(\xi)) - f(\xi)) = \end{aligned} \tag{2.40}$$

by the mean value theorem

$$= Xf(exp(t_1X)(exp(tY)(\xi)) + Y f(exp(t_2Y)(\xi))) \to Xf(\xi) + Y f(\xi)$$

as $t \to 0$. □

Remark 2.2.37. *Consider*

$$K(t) = exp(-tY)exp(-tX)exp(tY)exp(tX)(\xi)$$

and $f \in C^1_H(M)$*, then there exists*

$$\frac{d}{dt}(f \circ K)(0) = 0.$$

Theorem 2.2.38. *Let $M \subset \mathbb{R}^n$, let $f : M \to \mathbb{R}$ be a continuous function such that there exist the Lie derivatives Xf and Yf and they are continuous functions. Then there also exists $(X+Y)f = Xf + Yf$ in M.*

Proof: Arguing as in theorem 2.2.21 we immediately see that

$$|exp(tX)exp(tY)(\xi) - exp(t(X+Y))(\xi)| = O(t^2),$$

locally uniformly in ξ. It follows that

$$\begin{aligned} &\frac{1}{t}(f(exp(t(X+Y))(\xi)) - f(\xi)) = \\ &= \frac{1}{t}(f(exp(tX)(exp(tY)(\xi))) - f(\xi)) + O(t) \to Xf(\xi) + Yf(\xi), \end{aligned} \tag{2.41}$$

as $t \to 0$ by theorem 2.2.36. □

Definition 2.2.39. Let M be an open set in $\mathbb{R}^n$, and assume that on M is defined a family of vector fields X_j $j = 1, ..., m$, satisfying the Hörmander condition. A function $f : M \to \mathbb{R}$ is differentiable at a point $\xi \in M$ in the intrinsic sense if

$$f(\sum_{j=1}^{m} exp(t_j X_j)(\xi)) - f(\xi) = \sum_{j=1}^{m} t_j X_j f(\xi) + o(\| t \|) \tag{2.42}$$

as $\| t \| \to 0$.

Note that in the above definition appears only vector fields of degree 1. As a direct consequence of the previous theorem it follows the following theorem

Theorem 2.2.40. *Let $M \subset \mathbb{R}^n$, and assume that on M is defined a family of vector fields X_j, $j = 1, ...m$ satisfying the Hörmander condition. If f is of class $C^1_H(M)$, then it is differentiable.*

Finally, we introduce the notion of subelliptic operator and we recall the Hörmander theorem.

Definition 2.2.41. If $\phi = (\phi_1, ..., \phi_m)$ is a C^1_H section of the horizontal tangent plane, we call divergence of ϕ

$$div_H(\phi) = \sum_{j=1}^{m} X_j^* \phi_j$$

where X_j^* is the formal adjoint of the vector field X_j.

Definition 2.2.42. Let $M \subset \mathbb{R}^n$ be an open set, let (X_j),$j = 1, ..., m$ be a family of smooth vector fields defined on M, and let $f : M \to \mathbb{R}$, $f \in C^1_H$. We define the canonical Sublaplacian operator as

$$\Delta_H(f) = div_H(\nabla_H)(f).$$

where ∇_H is the horizontal gradient defined in Def. 2.2.33.

Definition 2.2.43. Let A be an $m \times m$ matrix (A_{ij}). A is uniformly elliptic if there exist two positive numbers λ, Λ such that

$$\lambda |x_i|^2 \leq \sum_{j=1}^{m} A_{ij}\xi_i\xi_j \leq \Lambda |\xi|^2 .$$

We call the operator induced by the matrix A

$$L_A = \sum_{i,j=1}^{m} A_{ij} X_i X_j$$

a uniformly elliptic operator.

We then define the subcaloric equation, the natural analogous of the heat equation, expressed in terms of the subelliptic operator:

$$\partial_t = L_A$$

Definition 2.2.44. Consider L_A an operator. L_A is called hypoelliptic if for every $u \in C_0^\infty$, u is in C_H^∞ in every open set where Lu is C_H^∞.

We end this section by stating the well known Hörmander theorem [19]

Theorem 2.2.45. *If $X_1, ..., X_m$ satisfy the Hörmander condition, then the associated subelliptic operator and the heat operator are hypoelliptic operators.*

2.3 The visual cortex as the SE(2) group

The visual cortex at a certain level is naturally modelled as the Rototranslation group with a sub-Riemannian metric. In the literature the Rototranslation group is also known as the *2D Eucliedean motion group SE(2)*. The aim of this section is to study the law group of the Rototranslation group and find the Lie algebra. Within this thesis we won't focus on how to model the visual cortex, for more details see [15]. In [15] is described the functional architecture of the visual cortex, in particular the fibration resulting from the vector product $\mathbb{R}^2 \times S^1$ describes the space of parameters of simple V1 cells. The parametrization is given by their retinotopic position (x_1, y_1) and their orientation reference θ.

We now want to define the group law. First of all we denote T_{x_1,y_1} the translation of the vector (x_1, y_1) and R_θ a rotation matrix of angle θ:

$$R_\theta = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \tag{2.43}$$

Then a general element of the $SE(2)$ group is of the form $A_{x_1,y_1,\theta} = T_{x_1,y_1} \circ R_\theta$, and applied to a point (x, y) it yields:

$$A_{x_1,y_1,\theta}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + R_\theta \begin{pmatrix} x \\ y \end{pmatrix} \tag{2.44}$$

All the profiles can be interpreted as: $\phi(x_1, y_1, \theta) = \phi_0 \circ A_{x_1,y_1,\theta}$. The set of parameters $g_1 = (x_1, y_1, \theta_1)$ form a group with the operation induced by the composition $A_{x_1,y_1,\theta_1} \circ A_{x_2,y_2,\theta_2}$. This turns out to be:

$$g_1 \cdot g_2 = (x_1, y_1, \theta_1) +_R (x_2, y_2, \theta_2) = \left(\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + R_{\theta_1} \begin{pmatrix} x \\ y \end{pmatrix} \right)^T, \theta_1 + \theta_2 \right) \tag{2.45}$$

Being induced by the composition law, one can easily check that $+_R$ verifies the group operation axioms, where the inverse of a point $g_1 = (x_1, y_1, \theta_1)$ is induced by the rototranslation:

$$A^{-1}_{x_1,y_1,\theta_1} = R^{-1}_\theta \circ T^{-1}_{x_1,y_1},$$

and the identity element is given by the trivial point $e = (0, 0, 0)$. Then, the group generated by the operation $+_R$ in the space $\mathbb{R}^2 \times S^1$ is called the rototranslation group or equivalently $SE(2)$.

In the case of V1 Citti and Sarti in [13] and [14] proposed to endow $\mathbb{R}^2 \times S^1$ with a subriemannian structure. In the standard Euclidean setting the tangent space to $\mathbb{R}^2 \times S^1$ has dimension 3: they selected a bi-dimensional subset of the tangent space at each point, called the horizontal plane, as a model of the connectivity in V1.
The horizontal plane is generated by the first order operators X_1, X_2 defined as follows:

$$X_1 = cos\theta\partial_x + sin\theta\partial_y \quad X_2 = \partial_\theta. \tag{2.46}$$

They then proved that only curves of the space are integral curves of the two vector fields X_1, X_2, while there is no natural curve with a non vanishing component in the direction

$$X_3 = -sin\theta\partial_x + cos\theta\partial_y$$

We explicitly note that the vector fields $\vec{X_1}, \vec{X_2}, \vec{X_3}$ are left invariant with respect to the group law of rotations and translations, so that they are the generators of the associated Lie algebra.
Finally note that $rank(\mathcal{L}(X_1, X_2)) = 3$, infact

$$[X_2, X_1] = X_3$$

then a basis of the space $\mathbb{R}^2 \times S^1$ is

$$\begin{aligned} X_1 &= cos\theta\partial_x + sin\theta\partial_y \\ X_2 &= \partial_\theta \\ X_3 &= -sin\theta\partial_x + cos\theta\partial_y \end{aligned} \tag{2.47}$$

We are now ready to illustrate the PDE which is the starting point of discussion of this thesis.

Chapter 3

A convection-diffusion-dilation/erosion equation

In this chapter we will study the following equation introduced by Duits, M.N.Smets, Portegies and Bekkers in [1]:

$$\begin{cases} \frac{\partial W}{\partial t}(p,t) = \overbrace{-cW(p,t)}^{Convection} \overbrace{-(-\Delta_{\mathcal{G}_1})^{\alpha} W(p,t)}^{Fractional\ diffusion} \overbrace{\pm \parallel \nabla_{\mathcal{G}_2} W(p,t) \parallel_{\mathcal{G}_2}^{2\alpha}}^{Dilation/Erosion} & for\ p \in G/H, t \geq 0 \\ W(p,0) = U(p) & for\ p \in G/H \end{cases} \tag{3.1}$$

where G is a Lie group and H is a subgroup of G. This is a convection-diffusion-dilation/erosion equation composed of three different parts which will be analyzed separately. We will analyze the case the function is defined on a homogenous space, in particular we will look for our case of interest $\mathbb{M}_2$, which will be identified with $\mathbb{R}^2 \times S^1$. This kind of equations can be applied in image processing problems, see for example [1].

The chapter is organized as follows. In the first section we briefly recall the main definitions and result on homogeneous spaces which will serve us to introduce the PDE (3.1). In the second section we will then introduce more in details equation (3.1) and in the third, fourth and fifth section we will analyze each part of the equation separately.

3.1 Groups and homogeneous spaces

In this section we will define the space in which we will work and see some important properties deriving from the request of equivariance. Finally we will see the particular case of the homogeneous space $\mathbb{M}^2$.
First of all consider X and Y two spaces and let:

$$\Psi(X) = \{f,\ f : X \to \mathbb{R}\}$$

$$\Psi(Y) = \{f\ f : Y \to \mathbb{R}\}$$

Then we want to consider operators defined from a space of functions $\Psi(X)$ defined on X to a space of functions $\Phi(Y)$ defined on Y:

$$l : \Psi(X) \longrightarrow \Phi(Y)$$

First of all we need to introduce the notion of homogeneous space.

Definition 3.1.1. We say that a group G acts smoothly on a space X if it exists a smooth map $\rho_X : G \times X \longrightarrow X$ such that for all $g, h \in G$,

$$\rho_X(gh, x) = \rho_X(g, \rho_X(h, x))$$

Notation: We write

$$\rho_X(g, x) = g.x \quad \rho_Y(h, y) = h.y$$

Definition 3.1.2. A group G acts transitively on a space X if, $\forall x, y \in X$, $\exists g \in G$ such that

$$g.x = y$$

Definition 3.1.3. Let G be a group and let X be a non-empty manifold. Then we say X is a homogeneous space for the group G if G acts smoothly and transitively on X.

In the sequel we will then make some assumptions: the first one is that all the possible transformations form a Lie group G, then G acts smoothly on the spaces.Secondly, we will assume that the group G acts transitively on the spaces: then, for any two elements of the space there exists a transformation in G that maps one to the other. This fact has the consequence that X can be seen as homogeneous space. We said we want to work on an homogeneous space and now we want to show that we can consider instead the quotient G/H, where G is a Lie group and H is a subgroup of G. In particular, this means that if we choose a reference element $x_0 \in X$, we can make the following isomorphism. Consider G a Lie group with homogeneous space X and let $x_0 \in X$ then

$$(X, G, x_0) \cong G/Stab_G(x_0)$$

using the mapping

$$x \mapsto \{g \in G | g.x_0 = x\} \tag{3.2}$$

where

$$Stab_G(x_0) = \{g \in G | g.x_0 = x_0\}$$

This mapping is a bijection due to the transitivity and the fact that $Stab_G(x_0)$ is a subgroup of G. Because of this we can represent a homogeneous space as the quotient G/H for some choice of subgroup H since all homogenous spaces are isomorphic to such a quotient by the above construction.

Definition 3.1.4. Consider $p \in G/H$. Then p is a subset of G and it is called a left coset of H since $\forall p \in G/H\ \exists g \in G$ such that $p = gH$.

We will denote the group action by an element $g \in G$ by the operator $L_g : G/H \to G/H$ given by

$$L_g p := g.p \quad \forall p \in G/H. \tag{3.3}$$

In addition, we denote the left-regular representation of G on functions f defined on G/H by $\mathcal{L}_g$ defined by

$$(\mathcal{L}_g f)(p) := f(g^{-1}.p) \tag{3.4}$$

Remark 3.1.5. *Consider $p \in G/H$ a left coset, then it is a subset of G. If we consider the equivalence relation:*

$$g_1 \equiv g_2 \Leftrightarrow g_1^{-1} g_2 \in H$$

the left cosets are a partition of G under the above equivalence relation. Infact, every left coset is expressed as $p = gH$, with $g \in G$; then if we consider $g_1, g_2 \in p$ we have $g_1 = gh_1$ and $g_2 = gh_2$, with $h_1, h_2 \in H$. Then $g_1^{-1} g_2 = h_1^{-1} h_2 \in H$.

We can now naturally extend the group action defined on G by considering the isomorphism defined in (3.2). We now have

$$x \mapsto p \Leftrightarrow g.x \mapsto g.p$$

where we have defined

$$g.p := gp,$$

which is again a left coset and so an element of G/H. We can see the subsets p also as atomic entities which represent some $x \in X$ by $p.x_0 = x$. With this notation the left coset that is associated with the reference element $x_0 \in X$ is H and for that reason we will indicate it by $p_0 := H$ so that the isomorphism maps $x_0 \mapsto p_0$.

Finally, notice that if we consider the p's as subsets of G and use the explicit notation

$$G_p := p \subset G \tag{3.5}$$

then we have that the group G consists of the disjoint union

$$G = \coprod_{p \in G/H} G_p$$

Definition 3.1.6. If we indicate with e the neutral element of the group G and consider $H = \{e\}$ we get $G/H \equiv G$, then the Lie group is a homogeneous space of itself. We call this space the principal homogeneous space. Notice that in this case the group action is equivalent to the group composition.

The third assumption is that we ask the operators to be equivariant.

Definition 3.1.7. Let G be a Lie group with homogeneous spaces X and Y. Let Φ be an operator from functions on X to functions on Y, then we say that Φ is equivariant with respect to G if for all functions f we have that:

$$\forall g \in G, y \in Y : (\mathcal{L}_g \circ \Phi)(y) = (\Phi \circ \mathcal{L}_g)(y). \tag{3.6}$$

The idea is to design the PDE in such a way that it is equivariant.Equivariance essentialy means that one can either transform the initial condition and then feed it through the PDE, or first feed it through the PDE and then transform the output function, and both give the same result.

The action on a homogeneous space induces an action on spaces of functions on the homogeneous space. The particular operators that we will base our framework on are vector and tensor fields, then we need to define the pushforward, a function defined on the tangent space of the homogenous space.

Definition 3.1.8. For $g \in G$ and $p \in G/H$, the pushforward

$$(L_g)_* : T_p(G/H) \to T_p(G/H) \tag{3.7}$$

of the group action L_g is defined by the condition that for all smooth functions f on G/H and all $v \in T_p(G/H)$ we have that

$$((L_g)_* v)f := v(f \circ L_g). \tag{3.8}$$

As we've already seen we can consider a tangent vector $v \in T_p(G/H)$ as a differential operator acting on functions defined on G/H.

In order to complete this section we need to define left-invariant vector fields and left-invariant metric tensor fields on homogeneous spaces.

Definition 3.1.9. A vector field v on G/H is left invariant with respect to G if it satisfies

$$\forall g \in G, \forall p \in G/H, v(g.p) = (L_g)_* v(p). \tag{3.9}$$

Lemma 3.1.10. *A vector field v on G/H is left invariant if and only if*

$$\forall g \in G, \forall p \in G/H : v(p)f = v(g.p)[\mathcal{L}_g f]. \tag{3.10}$$

Proof:Consider f a differentiable function defined on G/H.
First of all notice that for $p \in G/H$ and $g \in G$ it holds:

$$\mathcal{L}_g(f \circ L_g)(p) = \mathcal{L}_g(f)(g.p) = f(g^{-1}g.p) = f(p) \tag{3.11}$$

Then we have $\mathcal{L}_g(f \circ L_g) = f$.
Now, suppose v is a left invariant tensor field and consider

$$v(g.p)(\mathcal{L}_g f) = (L_g)_* v(p)(\mathcal{L}_g f) = v(p)(\mathcal{L}_g(f \circ L_g)) = v(p)f \tag{3.12}$$

where we have used definition 3.9 and 3.8.
On the other hand, if we suppose 3.10 holds for every $g \in G, p \in G/H$ then

$$(L_g)_* v(p) f = v(p)(f \circ L_g) = v(g.p)(\mathcal{L}_g(f \circ L_g)) = v(g.p)f \tag{3.13}$$

this completes the proof. □

The lemma immediately implies the following result

Theorem 3.1.11. *On a homogeneous space G/H with reference element p_0 the left invariant vector fields have the following properties*

1. *they are fully determined by their value $v(p_0)$ in p_0,*
2. $\forall h \in H, \forall v \in T_{p_0}(G/H) : (L_h)_* v = v$.

Finally, we introduce left-invariant metric tensor fields.

Definition 3.1.12. Let G be a Lie group and G/H a homogeneous space then the metric tensor field $\mathcal{G}$ on G/H is left-invariant with respect to G if and only if

$$\forall g \in G, \forall p \in G/H, \forall v, w \in T_p(G/H) : \mathcal{G}\mid_p (v, w) = \mathcal{G}\mid_{g.p} ((L_g)_* v, (L_g)_* w). \tag{3.14}$$

Recall that $L_g p := g.p$ and so the push-forward $(L_g)_*$ maps tangent vector from T_p to $T_{g.p}$. Again it follows immediately from this definition that a left-invariant metric has similar properties as a left-invariant vector field

Theorem 3.1.13. *On a homogeneous space G/H with reference element p_0 a left-invariant metric tensor field $\mathcal{G}$ has the following properties:*

1. *it is fully determined by its metric tensor $\mathcal{G}\mid_{p_0}$ at p_0*
2. $\forall h \in H, \forall v, w \in T_{p_0}(G/H) : \mathcal{G}\mid_{p_0} (v, w) = \mathcal{G}\mid_{p_0} ((L_h)_* v, (L_h)_* w)$.

Or in other words, the metric has to be invariant with respect to the subgroup H. We end this theoretical preliminaries by analyzing the particular case of the space of positions and orientations $\mathbb{M}_2$.

3.1.1 The group $SE(2)$ and the homogenous space $\mathbb{M}_2$

The particular case we want to consider is the one of the homogenous space $\mathbb{M}_2$. Consider the *Special Euclidean* group $SE(2)$ of rotations and translations of $\mathbb{R}^2$ introduced in section (2.3). When we take $H = \{0\} \times SO(1)$, by taking the quotient $SE(2)/H$, we obtain the space of positions and orientations $\mathbb{M}_2$.

As a set we identify $\mathbb{M}_2$ with $\mathbb{R}^2 \times S^1$ then we can represent elements of $\mathbb{M}_2$ with $(x, y, \theta) \in \mathbb{R}^3$ where x, y are the usual Cartesian coordinates and θ the angle with respect to the x-axis. The reference element is then simply denoted by $(0, 0, 0)$. If we denote elements of $SE(2)$ as translation/rotation pairs $(y, R) \in \mathbb{R}^2 \times SO(2)$ then the group multiplication is given by

$$g_1 = (y_1, R_1), g_2 = (y_2, R_2) \in SE(2) : g_1 g_2 = (y_1, R_1)(y_2, R_2) = (y_1 + R_1 y_2, R_1 R_2) \tag{3.15}$$

and the group action on elements $p = (x, n) \in \mathbb{R}^2 \times S^1 \equiv \mathbb{M}_2$ is given as

$$g.p = (y, R).(x, n) = (y + Rx, Rn). \tag{3.16}$$

We end this brief introduction to the $\mathbb{M}_2$ space by stating the following theorem about Riemannian metric tensor fields.

Theorem 3.1.14. *The only Riemannian metric tensor fields on $\mathbb{M}_2$ that are left-invariant with respect to $SE(2)$ are of the form:*

$$\begin{aligned} \mathcal{G} \mid_{(x,y,\theta)} ((\dot{x}, \dot{y}, \dot{\theta}), (\dot{x}, \dot{y}, \dot{\theta})) = D_M(|\dot{x}cos\theta|^2 + |\dot{y}sin\theta|^2) + \\ + D_L(|\dot{x}sin\theta|^2 + |\dot{y}cos\theta|^2) + D_A|\dot{\theta}|^2. \end{aligned} \tag{3.17}$$

with $D_M, D_L, D_A > 0$ weighing the main, lateral and angular motion respectively.

Proof: See [1],Th.2.8.

3.2 Convection-diffusion-dilation/erosion equation

We want to find the solution to the PDE:

$$\begin{cases} \frac{\partial W}{\partial t}(p, t) = -cW(p, t) - (-\Delta_{\mathcal{G}_1})^\alpha W(p, t) \pm \parallel \nabla_{\mathcal{G}_2} W(p, t) \parallel_{\mathcal{G}_2}^{2\alpha} & for\ p \in G/H, t \geq 0 \\ W(p, 0) = U(p) & for\ p \in G/H. \end{cases} \tag{3.18}$$

Here, c is a left-invariant vector field on G/H, $\alpha \in [1/2, 1]$, $\mathcal{G}_1$ and $\mathcal{G}_2$ are left-invariant metric tensor fields on G/H, U is the initial condition and $\Delta_{\mathcal{G}}$ and $\| \cdot \|_{\mathcal{G}}$ denote the Laplacian operator and norm induced by the metric tensor field $\mathcal{G}$.
As we can clearly see the PDE is divided in three parts:

1. The first one $-cW(p,t)$ is the convection part, it takes care of equivariant transport and depends on the left-invariant vector field c.

2. The second part of the PDE $-(-\Delta_{\mathcal{G}_1})^\alpha W(p,t)$ is the fractional diffusion

3. The third part of the PDE $\pm \parallel \nabla_{\mathcal{G}_2} W(p,t) \parallel^{2\alpha}_{\mathcal{G}_2}$ is the dilation (+ sign) and erosion (- sign).

Since the convection vector field c and the metric tensor fields $\mathcal{G}_1$ and $\mathcal{G}_2$ are left-invariant, the PDE is automatically equivariant.

The convection, diffusion and dilation/erosion steps are implemented with respectively a shift, convolution and morphological convolution. In the next sections we will analyze each step separately.

3.3 Convection

Suppose we have an input function $U^1 : G/H \to \mathbb{R}$, then the convection step takes this function as initial condition of the following PDE

$$\begin{cases} \frac{\partial W^1}{\partial t}(p,t) = -c(p)W^1(\cdot,t) & for\ p \in G/H, t \geq 0, \\ W^1(p,0) = U^1(p) & for\ p \in G/H \end{cases} \tag{3.19}$$

The output is the function $p \mapsto W^1(p,T)$. In fact, as we've said before, we want to find the solution of the PDE at time $t = T$. In this case the solution of the PDE is a shift of the initial condition, in particular it holds the following theorem

Theorem 3.3.1. *The solution of the convection PDE is found by method of characteristics, and is given by*

$$W^1(p,t) = (\mathcal{L}_{g_p^{-1}} U^1)(\gamma_c(t)^{-1}.p_0) = U^1(g_p \gamma_c(t)^{-1}.p_0), \tag{3.20}$$

where $g_p \in G_p$, (then $g_p.p_0 = p$) and $\gamma_c : \mathbb{R} \to G$ is the exponential curve that satisfies $\gamma_c(0) = e$ and

$$\frac{\partial}{\partial t}(\gamma_c(t).p)(t) = c(\gamma_c(t).p), \tag{3.21}$$

This means that γ_c is the exponential curve in the group G that induces the integral curves of the left-invariant vector field c on G/H when acting on elements of the homogeneous space.

Note that in general such exponential curves do not exist for general convection vector fields. Then the previous theorem is a consequence of the vector field c being left-invariant.

Proof: We want to show that the solution (3.20) solves the convection PDE. Consider:

$$\begin{aligned}\frac{\partial W^1}{\partial t}(p,t) &= \lim_{h\to 0}\frac{W^1(p,t+h)-W^1(p,t)}{h}\\ &= \lim_{h\to 0}\frac{U^1(g_p\gamma_c(t+h)^{-1}.p_0)-U^1(g_p\gamma_c(t)^{-1}.p_0)}{h}\\ &= \lim_{h\to 0}\frac{U^1(g_p\gamma_c(t)^{-1}\gamma_c(h)^{-1}.p_0)-U^1(g_p\gamma_c(t)^{-1}.p_0)}{h}\end{aligned} \tag{3.22}$$

where we have used the fact that γ_c is an exponential curve, then

$$\gamma_c(t+h)^{-1} = \gamma_c(t)^{-1}\gamma_c(h)^{-1}$$

Now, let $\bar{U} = \mathcal{L}_{\gamma_c(t)g_p^{-1}}U^1$.
Recall that $(\mathcal{L}_g U^1)(p) = U^1(g^{-1}.p)$, then

$$\bar{U}(p) = \mathcal{L}_{\gamma_c(t)g_p^{-1}}U^1(p) = U^1(g_p\gamma_c(t)^{-1}.p)$$

Substituting, it yields:

$$\begin{aligned}&= \lim_{h\to 0}\frac{\bar{U}(\gamma_c(h)^{-1}.p_0)-\bar{U}(p_0)}{h}\\ &= -c(p_0)\bar{U}\\ &= -(L_{g_p})_* c(p_0)\mathcal{L}_{g_p}\bar{U}\end{aligned} \tag{3.23}$$

where in the last equality we have used definitions (3.8) and (3.4).
The vector field c is left invariant, which means that it holds the following equality

$$c(g.p) = (L_g)_* c(p) \tag{3.24}$$

Then

$$(L_{g_p})_* c(p_0) = c(g_p.p_0) = c(p) \tag{3.25}$$

This yields:

$$\begin{aligned}&= -c(p)\mathcal{L}_{g_p}\mathcal{L}_{\gamma_c(t)g_p^{-1}}U^1\\ &= -c(p)[p\mapsto U^1(g_p\gamma_c(t)^{-1}g_p^{-1}.p)]\\ &= -c(p)[p\mapsto U^1(g_p\gamma_c(t)^{-1}.p_0)]\\ &= -c(p)W^1(\cdot,t).\end{aligned} \tag{3.26}$$

Where we have used the fact that $g_p^{-1}.p = p_0$. This completes the proof. □

In this first step we do not need to do any further approximations because the solution is a shift of the initial condition U^1.

3.4 Fractional diffusion

We will indicate the initial condition of the fractional diffusion PDE as the function $U^2 : G/H \to \mathbb{R}$. The fractional diffusion step solves the PDE

$$\begin{cases} \frac{\partial W^2}{\partial t} = -(-\Delta_{\mathcal{G}_2})^\alpha W^2(p,t) & for\ p \in G/H, t \geq 0 \\ W^2(p,0) = U^2(p) \end{cases} \tag{3.27}$$

Before analyzing the solution of the PDE we want to give the main definitions and results about the fractional laplacian. We refer to [2] for details.

3.4.1 Fractional laplacian

In this section we will introduce the notion of fractional Laplacian $L^s = (-\Delta)^s, 0 < s < 1$ and we will show how we can obtain the pointwise formula for $(-\Delta)^s u(x)$ starting from the semigroup formula. We will then give the formula for the solution to the Poisson problem $(-\Delta)^s u = f$ found through the semigroup approach as the inverse of the fractional Laplacian $u(x) = (-\Delta)^{-s} f(x)$.

First of all, let u be a function in the Schwartz class $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$, the set of all functions on $\mathbb{R}^n$ which decrease rapidly.

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n) \mid \forall \alpha, \beta \in \mathbb{N}_0^n, \sup_{x \in \mathbb{R}^n} |x^\alpha \partial_\beta f(x)| < \infty\}$$

Then the Fourier transform of u, denoted by $\hat{u}$, is also in $\mathcal{S}$. For the Laplacian $-\Delta$ on $\mathbb{R}^n$ we have:

$$\widehat{-(\Delta)u}(\xi) = |\xi|^2\, \hat{u}(\xi) \tag{3.28}$$

The Fractional Laplacian is then defined in a natural way as

Definition 3.4.1. Let $u \in \mathcal{S}(\mathbb{R}^n)$, then we define the fractional Laplacian $-(\Delta)^s, 0 < s < 1$ as

$$-(\Delta)^s u(\xi) = \widehat{|\xi|^{2s}\, \hat{u}(\xi)}. \tag{3.29}$$

It is obvious from the above definition that

$$(-\Delta)^0 u = u, \quad (-\Delta)^1 u = -\Delta u, \quad (-\Delta)^{s_1} \circ (-\Delta)^{s_2} u = (-\Delta)^{s_1+s_2} u \quad \forall s_1, s_2$$

Even though $|\xi|^{2s}\, \hat{u}(\xi)$ is a well defined function of $\xi \in \mathbb{R}^n$, we still have

$$(-\Delta)^s \notin \mathcal{S}$$

because $|\xi|^{2s}$ creates a singularity at $\xi = 0$. On the other hand we have that if $u \in \mathcal{S}$ then $(-\Delta)^s u \in C^\infty(\mathbb{R}^n)$.

In order to find the pointwise formula for the fractional Laplacian and give the solution to the Poisson problem we start stating the following numerical identities:

$$\lambda^s = \frac{1}{\Gamma(-s)} \int_0^\infty (e^{-t\lambda} - 1) \frac{dt}{t^{1+s}} \quad for\ any\ \lambda \geq 0 \tag{3.30}$$

$$\lambda^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty e^{-t\lambda} \frac{dt}{t^{1-s}} \quad for\ any\ \lambda > 0, s > 0 \tag{3.31}$$

which can be easily checked with a simple change of variables.

If we now choose $\lambda = |\xi|^2$, for $\xi \in \mathbb{R}^n$ and substitute in the numerical formula (3.30), then

$$\widehat{(-\Delta)^s} u(\xi) = |\xi|^{2s}\, \hat{u}(\xi) = \frac{1}{\Gamma(-s)} \int_0^\infty (e^{-t|\xi|^2} \hat{u}(\xi) - \hat{u}(\xi)) \frac{dt}{t^{1+s}}. \tag{3.32}$$

Thus by inverting the Fourier transform, we obtain the semigroup formula for the fractional Laplacian:

$$(-\Delta)^s u(x) = \frac{1}{\Gamma(-s)} \int_0^\infty (e^{t\Delta} u(x) - u(x)) \frac{dt}{t^{1+s}} \tag{3.33}$$

The family of operators $\{e^{t\Delta}\}_{t\geq 0}$ is the classical heat diffusion semigroup generated by Δ. Consider the solution $v = v(x,t)$, for $x \in \mathbb{R}^n$ and $t \geq 0$, of the heat equation on the whole space $\mathbb{R}^n$ with initial temperature u

$$\begin{cases} \partial_t v = \Delta v & for\ x \in \mathbb{R}^n, t > 0 \\ v(x,0) = u(x) & for\ x \in \mathbb{R}^n \end{cases} \tag{3.34}$$

If we apply the Fourier transform in the variable x for each fixed time t then

$$\hat{v}(x,t) = e^{-t|\xi|^2} \hat{u}(\xi) = \widehat{e^{t\Delta} u(\xi)} \tag{3.35}$$

so that $u \mapsto e^{t\Delta} u$ is the solution operator. It is well known that

$$v(x,t) \equiv e^{t\Delta} u(x) = G_t * u(x) = \int_{\mathbb{R}^n} G_t(x-z) u(z) dz$$

where $G_t(x)$ is the Gauss-Weierstrass heat kernel:

$$G_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-|x|^2/(4t)}. \tag{3.36}$$

The semigroup formula and the positivity of the heat kernel (3.36) easily imply the maximum principle for the fractional Laplacian. They also permit us to compute the pointwise formula for the fractional Laplacian. Then it holds the following theorem

Theorem 3.4.2. *Let $u \in \mathcal{S}$,$x \in \mathbb{R}^n$ and $0 < s < 1$.*

1. *If $0 < s < 1/2$ then*

$$(-\Delta)^s u(x) = c_{n,s} \int_{\mathbb{R}^n} \frac{u(x) - u(z)}{|x - z|^{n+2s}} dz$$

and the integral is absolutely convergent.

2. *If $1/2 \leq s < 1$ then, for any $\delta > 0$,*

$$(-\Delta)^s u(x) = c_{n,s} \lim_{\varepsilon \to 0^+} \int_{|x-z|>\varepsilon} \frac{u(x) - u(z)}{|x - z|^{n+2s}} dz$$

$$= c_{n,s} \int_{\mathbb{R}^n} \frac{u(x) - u(z) - \nabla u(x) \cdot (x - z)\chi_{|x-z|<\delta}(z)}{|x - z|^{n+2s}} dz$$

where the second integral is absolutely convergent.

Proof:See [2, Theorem1]

Consider the Poisson problem. If we apply the Fourier transform to solve the Poisson equation

$$(-\Delta)^s u = f \quad in\ \mathbb{R}^n \tag{3.37}$$

we find that $|\xi|^{2s}\,\hat{u}(\xi) = \hat{f}(\xi)$.

Definition 3.4.3. Let $u \in \mathcal{S}(\mathbb{R}^n)$. The inverse of the fractional Laplacian, or negative power of the Laplacian $(-\Delta)^{-s}, s > 0$ is defined for $f \in \mathcal{S}(\mathbb{R}^n)$ as

$$(-\Delta)^{-s} f(\xi) = \widehat{|\xi|^{-2s}\,\hat{f}}(\xi), \quad for\ \xi \neq 0. \tag{3.38}$$

If we now choose $\lambda = |\xi|^2, \xi \neq 0$, substituting in the numerical formula (3.31) we obtain, for $\xi \in \mathbb{R}^n$,

$$\widehat{(-\Delta)^{-s}} f(\xi) = \frac{1}{\Gamma(s)} \int_0^\infty e^{-t|\xi|^2} \hat{f}(\xi) \frac{dt}{t^{1-s}}. \tag{3.39}$$

Therefore, by inverting the Fourier transform, we obtain the semigroup formula for the inverse fractional Laplacian:

$$(-\Delta)^{-s} f(x) = \frac{1}{\Gamma(s)} \int_0^\infty e^{t\Delta} f(x) \frac{dt}{t^{1-s}}. \tag{3.40}$$

Finally, it holds the following theorem on the foundamental solution for the Poisson problem considered in (3.37)

Theorem 3.4.4. *Let $f \in \mathcal{S}, x \in \mathbb{R}^n$ and $0 < s \leq n/2$. In case $s = n/2$ assume in addition that $\int_{\mathbb{R}^n} f = 0$. Then*

$$(-\Delta)^{-s} f(x) = \int_{\mathbb{R}^n} K_{-s}(x-z) f(z) dz. \tag{3.41}$$

Here

$$K_{-s}(x) = \begin{cases} c_{n,-s} \frac{1}{|x-z|^{n-2s}} & if\ 0 < s < n/2 \\ \frac{1}{\Gamma(n/2)(4\pi)^{n/2}}(-2log\,|x| - \gamma) & if\ s = n/2 \end{cases} \tag{3.42}$$

where γ is the Euler-Mascheroni constant and

$$c_{n,-s} = \frac{\Gamma(n/2 - s)}{4^s \Gamma(s) \pi^{n/2}} \quad for\ 0 < s < n/2.$$

Proof:See [2, Theorem5].

If we now consider the fractional diffusion PDE on $\mathbb{R}^n$

$$\begin{cases} \partial_t v(x,t) = -(-\Delta)^s v(x,t) & in\ \mathbb{R}^n \times (0,\infty) \\ v(x,0) = f(x) & on\ \mathbb{R}^n \end{cases} \tag{3.43}$$

it can be demonstrated that the solution is given by the linear convolution of f with a function called the foundamental solution of the fractional diffusion equation. The solution v is then given by

$$v(x,t) = (K^s * f)(x,t) \tag{3.44}$$

where

$$K^s : (0,\infty) \times \mathbb{R}^n \to [0,\infty)$$

Now that we have briefly introduced the notion of fractional Laplacian we can continue our discussion on the solution to the PDE (3.27).

Now that we have defined the fractional Laplacian we need to introduce the notion of linear group convolution. This will serve us to define the solution to the fractional diffusion equation.

Definition 3.4.5. Let G be a topological group. A left-invariant Haar measure on G is a Radon measure μ on G (then a regular measure which is finite on compact sets) such that $\mu(gA) = \mu(A)$ for all $g \in G$ and all measurable subsets A of G.

Definition 3.4.6. Let $H = Stab_G(p_0)$ be compact with reference element $p_0 \in G/H$, let $f \in L^2(G/H)$ and $k \in L^1(G/H)$ such that

$$\forall h \in H, p \in G/H : k(h.p) = k(p) \tag{3.45}$$

then we define the linear group convolution as

$$(k *_G f)(p) := \int_G k(g^{-1}.p) f(g.p_0) d\mu_G(g) \tag{3.46}$$

where μ_G is the left-invariant Haar measure on the group.

As with fractional diffusion on $\mathbb{R}^n$, there exists a smooth function

$$K^\alpha : (0, \infty) \times (G/H) \to [0, \infty), \tag{3.47}$$

called the foundamental solution of the $\alpha-$diffusion equation, such that for every initial condition U^2, the solution to the PDE (3.27) is given by the convolution of the function U^2 with the foundamental solution K_t^α

$$W^2(p,t) = (K_t^\alpha *_G U^2(\cdot, t))(p) \tag{3.48}$$

For details on the analytic expression of K_t^α see [23].

In chapter 6 we will analyze more in details the foundamental solution K_t^α. In order to find approximation of the fractional diffusion kernel we need to illustrate the parametrix method, which will solve the fractional diffusion equation. Let's now focus on the last substep of the PDE unit: the dilation/erosion PDE.

3.5 Dilation and erosion

Suppose U^3 is the input given to the dilation/erosion PDE. Then the output $W^3(p,t)$ will be the solution at time $t = T$ of the following dilation/erosion PDE:

$$\begin{cases} \frac{\partial W^3}{\partial t} = \pm \parallel \nabla_{\mathcal{G}_2} W^3(p,t) \parallel_{\mathcal{G}_3}^{2\alpha} & for\ p \in G/H, t \geq 0 \\ W^3(p,0) = U^3(p) & for\ p \in G/H \end{cases} \tag{3.49}$$

where we have chosen U^3 as initial condition of the PDE.

In order to give the expression of the solution to the dilation/erosion PDE we need to introduce the definition of morphological convolution.

Definition 3.5.1. Let $k : G/H \to \mathbb{R} \cup \{\infty\}$. k is a proper function if and only if there exists at least one $x \in G/H$ such that:

$$k(x) < +\infty \quad and \quad k(x) > -\infty$$

then the function k is not everywhere equal to ∞.

Definition 3.5.2. Let $f \in L^\infty(G/H)$, let $k : G/H \to \mathbb{R} \cup \{\infty\}$ be proper and let $p_0 \in G/H$ be the reference element of the homogeneous space, then we define the morphological convolution between f and k as:

$$(k \square_G f)(p) := \inf_{g \in G} k(g^{-1}.p) + f(g.p_0) \tag{3.50}$$

Example 3.5.3. *For $m \in \mathbb{R}$ let us define the convex Dirac function:*

$$\delta_m^c(x) = \begin{cases} +\infty & for\ x \neq m \\ 0 & for\ x = m, \end{cases} \tag{3.51}$$

and consider the function $\mathcal{M}_{m,\sigma}^p : \mathbb{R} \to \bar{\mathbb{R}}$, $\mathcal{M}_{m,\sigma}^p(x) = \frac{1}{p}(|x-m|/\sigma)^p$ for $p \geq 1$ with $\mathcal{M}_{m,0}^p = \delta_m^c$. We have the formula

$$\mathcal{M}_{m,\sigma}^p \Box \mathcal{M}_{\bar{m},\bar{\sigma}}^p = \mathcal{M}_{m+\bar{m},[\sigma^{p'}+\bar{\sigma}^{p'}]^{1/p'}}^p \quad with\ 1/p + 1/p' = 1$$

This result is the analogue of

$$\mathcal{N}_{m,\sigma} * \mathcal{N}_{\bar{m},\bar{\sigma}} = \mathcal{N}_{m+\bar{m},\sqrt{\sigma^2+\bar{\sigma}^2}}$$

in the particular case $p = 2$, where $\mathcal{N}_{m,\sigma}$ denotes the Gaussian law of mean m and standard deviation σ and $*$ the convolution operator.

Therefore there exists a morphism between the set of quadratic forms endowed with the morphological convolution operator and the set of exponentials of quadratic forms endowed with the convolution operator. This morphism is a particular case of the Cramér Fourier transform that we will define later.

Example 3.5.4. *Consider f_1, f_2 two proper functions on $\mathbb{R}^d$. The grayscale morphology operations $\oplus$ (dilation) and $\ominus$ (erosion) on $\mathbb{R}^d$ are defined as follows:*

$$\begin{aligned} (f_1 \oplus f_2)(x) &= \sup_{y \in \mathbb{R}^d} (f_1(y) + f_2(x-y)) \\ (f_1 \ominus f_2)(x) &= \inf_{y \in \mathbb{R}^d} (f_1(y) - f_2(x-y)) \end{aligned} \tag{3.52}$$

The morphological convolution is related to these grayscale morphology operations as follows. Consider the following function:

$$\nu(x) = -f_2(-x)$$

then

$$\begin{aligned} f_1 \oplus f_2 &= -(-f_1 \Box_{\mathbb{R}^d} - f_2), \\ f_1 \ominus f_2 &= f_1 \Box_{\mathbb{R}^d} \nu. \end{aligned} \tag{3.53}$$

By a generalization of the Hopf-Lax formula [21], the solution to (3.49) is given by morphological convolution

$$W^3(p,t) = -(k_t^\alpha \Box_G - U^3)(p) \tag{3.54}$$

for the (+) dilation variant and

$$W^3(p,t) = (k_t^\alpha \Box_G U^3)(p) \tag{3.55}$$

for the (−) erosion variant, where the kernel k_t^α is a proper lower semi-continuous function of the type

$$k_\cdot^\alpha : (0,\infty) \times (G/H) \to \mathbb{R} \cup \{\infty\} \tag{3.56}$$

In chapter 6 we will see in details the approximations for the fractional diffusion and dilation/erosion kernels. In order to solve the fractional diffusion equation we have decided to use the parametrix method. This method is analyzed for $\alpha = 1$ for diffusion equations on $\mathbb{R}^d$ and on Lie groups. Using the parametrix method allow us to find approximations for the diffusion kernel also in case α is different from 1.

Concerning the dilation/erosion equation we will give later the definition of Cramér-Fourier transform. This particular operator will permit us to relate linear convolution and morphological convolution. Thanks to this fact we will be able to relate the approximate diffusion kernel (found thanks to the parametrix method) to the approximate dilation/erosion kernel.

Therefore in the next chapter we will analyze the parametrix method for parabolic operators on $\mathbb{R}^d$.

Chapter 4

Parametrix method

The objective of this chapter is to determine the foundamental solutions of parabolic equations with the parametrix method.

The parabolic operator is defined as the operator:

$$L_A = \sum_{i,j=1}^{n} a_{i,j}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} - \partial_t$$

where $(a_{i,j})$ is a symmetric and positive definite matrix defined on

$$\bar{\Omega} \times [T_1, T2]$$

,where Ω is a bounded open set of $\mathbb{R}^n$.

With Γ_A we will indicate the foundamental solution of the L operator. For constant coefficients a_{ij} the fundamental solution can be easily obtained, through a change of variables, from the fundamental solution of the heat operator. The foundamental solution of the heat operator is:

$$\Gamma(x,t) = \begin{cases} 0 & t \leq 0 \\ \frac{1}{(4\pi t)^{n/2}} e^{-\frac{\|x\|^2}{4t}} & t > 0 \end{cases} \tag{4.1}$$

As a result, the foundamental solution Γ_A will be:

$$\Gamma_A(x,t) = \frac{\sqrt{det(A^{-1})}}{(4\pi t)^{n/2}} e^{-\frac{\|x\|_A^2}{4t}} \tag{4.2}$$

where $\|x\|_A$ is the norm associated with the matrix A.

For variable coefficients $a_{i,j}$, in order to find the foundamental solution Γ_A, we have to use the parametrix method. This method is an iterative method which consists in approximating the operator L with operators with constant coefficients whose fundamental solution is explicitly known.

Let $(\xi, \tau) \in Q$, we define:

$$L_{\xi,\tau} = \sum_{i,j=1}^{n} a_{i,j}(\xi, \tau) \frac{\partial^2}{\partial x_i \partial x_j} - \partial_t$$

Let $\Gamma_{\xi,\tau}$ be the foundamental solution of $L_{\xi,\tau}$, which is known from 4.2. The foundamental solution Γ of the parabolic operator L will be found as:

$$\Gamma(x,t) = \Gamma_{\xi,\tau}(x,t) - (\Gamma_{\xi,\tau} * \sum_{j=1}^{+\infty} R_j^{\xi,\tau})(x,t)$$

Where:

$$\begin{cases} R_1^{\xi,\tau}(x,t) = \int_{T_0}^{t} \int_{\bar{\Omega}} (L - L_{y,\sigma}) \Gamma_{y,\sigma}(x-y, t-\sigma) dy d\sigma \\ R_j^{\xi,\tau}(x,t) = (-R_1^{\xi,\tau} * R_{j-1}^{\xi,\tau})(x,t) \end{cases} \tag{4.3}$$

It will therefore be demonstrated that Γ defined in this way is the foundamental solution of the L operator.

The chapter is organized as follows. In the first section we will prove some main properties of the foundamental solution Γ of a parabolic operator. These properties will serve us to demonstrate the subsequent results. In the second section we will focus on the characteristics of the function $R_j^{\xi,\tau}$. Finally,in the third section, we will construct the parametrix, we will demonstrate the convergence of the series $\sum_{i,j=1}^{+\infty} R_j^{\xi,\tau}$ and verify that Γ is the foundamental solution of the L operator.

4.1 Parabolic operators with constant coefficients

In this section we will see some of the main properties of the foundamental solution of a parabolic operator.

Definition 4.1.1. Let A be a symmetric and positive definite matrix. The dot product associated with A is defined as:

$$\langle v, w \rangle_A := \langle A^{-1} v, w \rangle$$

Therefore the associated norm is:

$$\|x\|_A = \sqrt{\langle x, x \rangle_A}$$

Remark 4.1.2. *: Since A is a symmetric and positive definite matrix,it holds:*

$$\lambda_{min} \|x\|^2 \leq \|x\|_A^2 \leq \lambda_{max} \|x\|^2$$

Let's consider the following parabolic operator:

$$L_A = \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} - \partial_t$$

where $A(x,t) = (a_{ij}(x,t))_{1 \le i,j \le n}$ is a symmetric and positive definite matrix.

Definition 4.1.3. The foundamental solution of the operator L_A is a function Γ_A such that:

$$L_A \Gamma_A = \delta$$

Example 4.1.4. *Let Q be $Q = \bar{\Omega} \times [T_0, T_1]$, where Ω is an open subset of $\mathbb{R}^n$. Let $L = \partial_t - M$, where M is an operator which does not depend on the time variable t. Let $\Gamma(x,t,\xi,\tau)$ be the foundamental solution of the equation $Lu = \delta$ on Q.Then*

- *$\forall(\xi,\tau) \in Q$ fixed, $\Gamma(x,t,\xi,\tau)$ is a solution of the equation $Lu = \delta(\xi,\tau)$, $\forall(x,t) \in Q$ with $(x,t) \neq (\xi,\tau)$*
- *For every function f continuous on $\bar{\Omega}$ it holds:*

$$\lim_{t \to \tau^+} \int_{\bar{\Omega}} \Gamma(x,t,\xi,\tau) f(\xi) d\xi = f(x) \quad \forall x \in \bar{\Omega} \tag{4.4}$$

Remark 4.1.5. *: Since A is symmetric and positive definite , also A^{-1} is symmetric and positive definite, and there exists N symmetric and positive definite such that $N^t N = A^{-1}$.*
If u is a solution of $L_{heat} u(x) = f(x)$, then $v(x,t) = u(Nx,t)$ is a solution of $L_A v(x,t) = f(Nx,t)$. Here L_{heat} indicates the heat operator.

Proof:

$$\begin{aligned}
L_A(v(x,t)) = L_A(u(Nx,t)) &= \sum_{i,j=1}^{n} a_{i,j} \frac{\partial^2 u(Nx,t)}{\partial x_i \partial x_j} - \frac{\partial u(Nx,t)}{\partial t} = \\
&= \sum_{i,j=1}^{n} a_{i,j} \langle Hess(u(Nx,t)) Ne_i, Ne_j \rangle - \frac{\partial u(Nx,t)}{\partial t} = \\
&= \sum_{i,j=1}^{n} \sum_{s=1}^{n} \sum_{t=1}^{n} a_{i,j} \frac{\partial^2 u(Nx,t)}{\partial x_s \partial x_t} n_{s_i} n_{t_i} - \frac{\partial u(Nx,t)}{\partial t} = \\
&= \sum_{s=1}^{n} \sum_{t=1}^{n} \frac{\partial^2 u(Nx,t)}{\partial x_s \partial x_t} \delta_{s,t} - \frac{\partial u(Nx,t)}{\partial t} = \\
&= \Delta u(Nx,t) - \frac{\partial u(Nx,t)}{\partial t} = f(Nx,t)
\end{aligned} \tag{4.5}$$

□From remark (4.1.5) it clearly follows that:

$$\Gamma_A(x,t) = \frac{1}{\sqrt{det(A)}}\Gamma(Nx,t)$$

Where Γ is the foundamental solution of the heat operator. Therefore the foundamental solution of the L_A operator can now be explicity expressed as:

$$\Gamma_A(x,t) = \begin{cases} 0 & if\ t \leq 0 \\ \frac{\sqrt{det(A^{-1})}}{(4\pi t)^{n/2}} e^{-\frac{\|x\|_A^2}{4t}} & if\ t > 0 \end{cases} \tag{4.6}$$

Remark 4.1.6. *: The derivatives of Γ_A are*

$$\begin{aligned} \frac{\partial \Gamma_A(x,t)}{\partial x_i} &= -\Gamma_A(x,t)\frac{\langle x, e_i\rangle_A}{2t} \\ \frac{\partial^2 \Gamma_A(x,t)}{\partial x_i \partial x_j} &= \Gamma_A(x,t)\frac{\langle x, e_i\rangle_A \langle x, e_j\rangle_A - 2t\langle e_j, e_i\rangle_A}{4t^2} \\ \frac{\partial \Gamma_A(x,t)}{\partial t} &= \Gamma_A(x,t)\frac{\|x\|_A^2 - 2nt}{4t^2} \end{aligned} \tag{4.7}$$

Theorem 4.1.7. *The following inequalities hold for the foundamental solution Γ_A*

$$\begin{aligned} \Gamma_A(x,t) &\leq \frac{\sqrt{det(A^{-1})}}{(4\pi t)^{n/2}} e^{-\frac{\lambda_{min}\|x\|^2}{4t}}, \\ \left|\frac{\partial \Gamma_A(x,t)}{\partial x_i}\right| &\leq \frac{\lambda_{max}}{\sqrt{t}}\Gamma_A^{(\lambda_{min})}(x,t) \\ \left|\frac{\partial^2 \Gamma_A(x,t)}{\partial x_i \partial x_j}\right| &\leq \frac{\lambda_{max} M}{t}\Gamma_A^{(\lambda_{min}}(x,t) \\ \left|\frac{\partial \Gamma_A(x,t)}{\partial t}\right| &\leq \frac{\lambda_{max} M}{t}\Gamma_A^{(\lambda_{min})}(x,t) \end{aligned} \tag{4.8}$$

Where:

$$M = \max_{1\leq i,j\leq n} \max_{(x,t)\in Q} a_{ij}(x,t)$$

Proof: For the first inequality we have used only remark (4.1.2). For the second inequality we have:

$$\begin{aligned} &\left|\frac{\partial \Gamma_A(x,t)}{\partial x_i}\right| \leq \Gamma_A(x,t)\frac{|\langle x, e_i\rangle_A|}{2t} \leq \Gamma_A(x,t)\frac{1}{\sqrt{t}}\frac{\lambda_{max}\|x\|_A}{\sqrt{t}} \leq \\ &\leq \frac{\lambda_{max}}{\sqrt{t}}\Gamma_A^{(\lambda_{min})}(x,t) \end{aligned} \tag{4.9}$$

where the last inequality follows from the following result: let $k \in \mathbb{R}$ be fixed, then there exists $C > 0$ such that:

$$s^k e^{-s^2} \leq C e^{-s^2/2} \quad \forall s \geq 0 \tag{4.10}$$

For the third inequality we have:

$$\begin{aligned} \left| \frac{\partial^2 \Gamma_A(x,t)}{\partial x_i \partial x_j} \right| &\leq \Gamma_A(x,t) \frac{|\langle x, e_i \rangle_A \langle x, e_j \rangle_A - 2t \langle e_i, e_j \rangle_A|}{4t^2} \leq \\ &\leq \Gamma_A(x,t) \frac{1}{t} \frac{M(\lambda_{max} \|x\|^2 - 2t)}{4t} \leq \Gamma_A(x,t) \frac{1}{t} \frac{M \lambda_{max} \|x\|^2}{4t} \leq \\ &\leq \frac{M \lambda_{max}}{t} \Gamma_A^{(\lambda_{min})}(x,t). \end{aligned} \tag{4.11}$$

Finally, similar calculations are made to find the last inequality. □

4.2 Parabolic operators with Hölder-continuous coefficients

The idea of the parametrix method is to find the foundamental solution of a parabolic operator with Hölder-continuous coeficcients starting from the foundamental solution of an operator with constant coefficients. In particular, the solution will be written as a convolutional operator with the fundamental solution of the operator with constant coefficients. It is therefore necessary, before illustrating the method, to study some important continuity properties of the convolutional operators.

Definition 4.2.1. A function f defined on $\bar{\Omega}$ is Hölder-continuous of order α, with $0 < \alpha \leq 1$, if there exists $c \in \mathbb{R}$, $c > 0$ such that:

$$|f(x) - f(y)| < c\,|x - y|^\alpha \quad \forall x, y \in \bar{\Omega}$$

We want to study the following parabolic operator:

$$Lu(x,t) = \sum_{i,j=1}^{n} a_{i,j}(x,t) \frac{\partial^2 u}{\partial x_i \partial x_j} - \frac{\partial u}{\partial t}$$

where $a_{i,j}(x,t)$ are functions defined on Q that satisfy the following Hölder-continuity conditions:

$$|a_{i,j}(x,t) - a_{i,j}(y,\tau)| \leq K(|x-y|^\alpha + |t-\tau|^{\alpha/2} \tag{4.12}$$

Moreover, we consider the matrix $A = (a_{i,j})_{1 \leq i,j \leq n}$ to be symmetric and positive definite.

If we consider $(\xi, \tau) \in Q$ fixed the parabolic operator L is an operator with constant coefficients then, for what we have seen in the previous section, the foundamental solution of:

$$L_{\xi,\tau}u = \sum_{i,j=1}^{n} a_{i,j}(\xi,\tau)\frac{\partial^2 u}{\partial x_i \partial x_j} - \frac{\partial u}{\partial t}$$

is:

$$\Gamma_{\xi,\tau}(x,t) = \begin{cases} 0 & if \quad t \leq 0 \\ \frac{\sqrt{det(A^{-1}(\xi,\tau))}}{(4\pi t)^{n/2}} e^{-\frac{\|x\|^2_{\xi,\tau}}{4t}} & if \quad t > 0 \end{cases} \tag{4.13}$$

where $\|x\|_{\xi,\tau} = \|x\|_{A(\xi,\tau)}$. Furthermore, the following inequality holds:

$$\lambda_{min}\|x\|^2 \leq \|x\|^2_{\xi,\tau} \leq \lambda_{max}\|x\|^2 \tag{4.14}$$

As we said before, the foundamental solution of the L operator will be find as:

$$\Gamma(x,t) = \Gamma_{\xi,\tau}(x,t) - (\Gamma_{\xi,\tau} * \sum_{j=1}^{+\infty} R_j^{\xi,\tau})(x,t)$$

Definition 4.2.2. Let f be a Hölder-continuous function defined on Q. We define:

$$V(x,t) = \int_{T_0}^{t}\int_{\bar{\Omega}} \Gamma_{\xi,\tau}(x-\xi, t-\tau) f(\xi,\tau) d\xi d\tau = \Gamma_{\xi,\tau} * f(x,t) \tag{4.15}$$

We want to study the regularity properties of $V(x,t)$.

Theorem 4.2.3. *Let $f(x,t)$ be a continuos function defined on Q. Then $\frac{\partial V(x,t)}{\partial x_i}$ is continuous and it holds:*

$$\frac{\partial V(x,t)}{\partial x_i} = \int_{T_0}^{t}\int_{\bar{\Omega}} \frac{\partial \Gamma_{\xi,\tau}(x-\xi, t-\tau)}{\partial x_i} f(\xi,\tau) d\xi d\tau \tag{4.16}$$

Proof: $\Gamma_{\xi,\tau}(x-\xi, t-\tau)$ is a gaussian function, from (4.8) it holds:

$$\int_{\bar{\Omega}} \frac{\partial \Gamma_{\xi,\tau}(x-\xi, t-\tau)}{\partial x_i} d\xi \leq \int_{\bar{\Omega}} \frac{\lambda_{max}}{\sqrt{t-\tau}} \Gamma^{(\lambda_{min})}_{\xi,\tau}(x-\xi, t-\tau) d\xi \leq \frac{C}{\sqrt{t-\tau}} \tag{4.17}$$

which is an L^1 function with respect to the variable t. The thesis follows from the Lebesgue's dominated convergence theorem. Moreover,

$$\frac{\partial \Gamma_{\xi,\tau}(x-\xi, t-\tau)}{\partial x_i}$$

is a continuous function, then, from the dominated convergence theorem,also $\frac{\partial V(x,t)}{\partial x_i}$ is continuous. □

Theorem 4.2.4. *It holds*

$$\left|\frac{\partial^2(\Gamma_{\xi,\tau}-\Gamma_{y,t})(x-\xi,t-\tau)}{\partial x_i \partial x_j}\right| \leq \frac{C}{|t-\tau|^{1-\alpha/2}}\Gamma^{\lambda_{min}}_{\xi,\tau}(x-\xi,t-\tau) \qquad (4.18)$$

Proof: First of all we remember that, for the mean value theorem, $\exists c \in [a,b]$ such that:

$$\left|\frac{e^b-e^a}{b-a}\right| \leq e^c$$

Then, from (4.14) we get:

$$\left|e^{-\frac{\|x-\xi\|_{\xi,\tau}}{4(t-\tau)}} - e^{-\frac{\|x-\xi\|_{y,t}}{4(t-\tau)}}\right| \leq e^c \left|\frac{\|x-\xi\|_{y,t}-\|x-\xi\|_{\xi,\tau}}{4(t-\tau)}\right|$$

where

$$min(-\frac{\|x-\xi\|_{\xi,\tau}}{4(t-\tau)},-\frac{\|x-\xi\|_{y,t}}{4(t-\tau)}) \leq c \leq max(-\frac{\|x-\xi\|_{\xi,\tau}}{4(t-\tau)},-\frac{\|x-\xi\|_{y,t}}{4(t-\tau)}) \leq \frac{-\lambda_{min}\|x-\xi\|^2}{4\,|t-\tau|}$$

. Then we obtain:

$$\begin{aligned}\left|e^{-\frac{\|x-\xi\|_{\xi,\tau}}{4(t-\tau)}} - e^{-\frac{\|x-\xi\|_{y,t}}{4(t-\tau)}}\right| &\leq e^{\frac{-\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}} \left|\frac{\|x-\xi\|_{y,t}-\|x-\xi\|_{\xi,\tau}}{4(t-\tau)}\right| \leq \\ \leq e^{\frac{-\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}} \frac{\left|\langle(A^{-1}(\xi,\tau)-A^{-1}(y,t))x-\xi,x-\xi\rangle\right|}{4\,|t-\tau|} &\leq \qquad (4.19) \\ \leq e^{\frac{-\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}} \frac{\|A^{-1}(\xi,\tau)-A^{-1}(y,t)\|_1\|x-\xi\|^2}{4\,|t-\tau|}&.\end{aligned}$$

The coefficients of A are Hölder-continuous, then we have

$$\|A^{-1}(\xi,\tau)-A^{-1}(y,t)\|_1 \leq n^2(\|x-\xi\|^\alpha+|t-\tau|^{\alpha/2})$$

Then, using (4.10) with $k=2$ we get:

$$\left|e^{-\frac{\|x-\xi\|_{\xi,\tau}}{4(t-\tau)}} - e^{-\frac{\|x-\xi\|_{y,t}}{4(t-\tau)}}\right| \leq e^{\frac{-\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}} \frac{n^2}{4}(\|x-\xi\|^\alpha+|t-\tau|^{\alpha/2})$$

Now we can demonstrate the thesis, infact:

$$\left|\frac{\partial^2(\Gamma_{\xi,\tau}-\Gamma_{y,t})(x-\xi,t-\tau)}{\partial x_i\partial x_j}\right| \leq \left|C\frac{(\Gamma_{\xi,\tau}-\Gamma_{y,t})(x-\xi,t-\tau)}{t-\tau}\right| \leq$$
$$\leq \frac{C}{(t-\tau)(4\pi(t-\tau))^{n/2}}\left|e^{-\frac{\|x-\xi\|_{\xi,\tau}}{4(t-\tau)}}-e^{-\frac{\|x-\xi\|_{y,t}}{4(t-\tau)}}\right| \leq$$
$$\leq \frac{C}{(t-\tau)(4\pi(t-\tau))^{n/2}}\left|e^{\frac{-\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}}\frac{(\|x-\xi\|^\alpha+|t-\tau|^{\alpha/2})}{4}\right| \leq$$
$$\leq \frac{C}{(4\pi(t-\tau))^{n/2}}(\left|e^{-\frac{\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}}\frac{\|x-\xi\|^\alpha}{|t-\tau|^{\alpha/2}}\frac{1}{|t-\tau|^{1-\alpha/2}}\right|+\left|e^{-\frac{\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}}\frac{1}{|t-\tau|^{1-\alpha/2}}\right|) \leq$$
$$\leq \frac{C}{(4\pi(t-\tau))^{n/2}}\left|e^{-\frac{\lambda_{min}\|x-\xi\|^2}{4|t-\tau|}}\frac{1}{|t-\tau|^{1-\alpha/2}}\right| \leq \frac{C}{|t-\tau|^{1-\alpha/2}}\Gamma^{(\lambda_{min})}_{\xi,\tau}(x-\xi,t-\tau). \tag{4.20}$$

□

Theorem 4.2.5. *Let $f(x,t)$ be function defined on Q, $f(x,t)$ Hölder-continuous of order α with respect to x. Then $\frac{\partial^2 V(x,t)}{\partial x_i\partial x_j}$ is continuous on Q and it holds:*

$$\frac{\partial^2 V(x,t)}{\partial x_i\partial x_j} = \int_{T_0}^{t}\int_{\bar\Omega}\frac{\partial^2\Gamma_{\xi,\tau}(x-\xi,t-\tau)}{\partial x_i\partial x_j}f(\xi,\tau)d\xi d\tau \tag{4.21}$$

Proof: Consider

$$J(x,t,\tau) = \int_{\bar\Omega}\Gamma_{\xi,\tau}(x-\xi,t-\tau)f(\xi,\tau)d\xi$$

Then from theorem (4.2.3) we have:

$$\frac{\partial V(x,t)}{\partial x_i} = \int_{T_0}^{t}\frac{\partial J(x,t,\tau)}{\partial x_i}d\tau$$

Let $y\in\bar\Omega$ fixed, and consider:

$$\begin{aligned}\frac{\partial J(x,t,\tau)}{\partial x_i} &= f(y,\tau)\int_{\bar\Omega}\frac{\partial\Gamma_{y,t}(x-\xi,t-\tau)}{\partial x_i}d\xi+\\ &+f(y,\tau)\int_{\bar\Omega}\frac{\partial(\Gamma_{\xi,\tau}-\Gamma_{y,t})(x-\xi,t-\tau)}{\partial x_i}d\xi+\\ &+\int_{\bar\Omega}\frac{\partial\Gamma_{\xi,\tau}(x-\xi,t-\tau)}{\partial x_i}(f(\xi,\tau)-f(y,\tau))d\xi=\\ &=f(y,\tau)[\int_{\partial K}\Gamma_{y,t}(x-\xi,t-\tau)\nu_i dS+\int_{\bar\Omega\setminus K}\frac{\partial\Gamma_{y,t}(x-\xi,t-\tau)}{\partial x_i}d\xi]+\\ &+f(y,\tau)\int_{\bar\Omega}\frac{\partial(\Gamma_{\xi,\tau}-\Gamma_{y,t})(x-\xi,t-\tau)}{\partial x_i}d\xi+\\ &+\int_{\bar\Omega}\frac{\partial\Gamma_{\xi,\tau}(x-\xi,t-\tau)}{\partial x_i}(f(\xi,\tau)-f(y,\tau))d\xi\end{aligned} \tag{4.22}$$

where $K \subseteq \bar{\Omega}$ is a closed ball and we have used the divergence theorem.
We consider now $y = x$ and derive another time $J(x, t, \tau)$:

$$\frac{\partial^2 J(x,t,\tau)}{\partial x_i \partial x_j} =$$
$$f(x,\tau)\Big[\int_{\partial K} \frac{\partial \Gamma_{y,t}(x-\xi,t-\tau)}{\partial x_i} \nu_i dS + \int_{\bar{\Omega}\setminus K} \frac{\partial^2 \Gamma_{y,t}(x-\xi,t-\tau)}{\partial x_i \partial x_j} d\xi\Big]_{y=x} +$$
$$+ f(x,\tau) \int_{\bar{\Omega}} \Big[\frac{\partial^2 (\Gamma_{\xi,\tau} - \Gamma_{y,t})(x-\xi,t-\tau)}{\partial x_i \partial x_j}\Big]_{y=x} d\xi +$$
$$+ \int_{\bar{\Omega}} \frac{\partial^2 \Gamma_{\xi,\tau}(x-\xi,t-\tau)}{\partial x_i \partial x_j} (f(\xi,\tau) - f(x,\tau)) d\xi =$$
$$= I_1 + I_2 + I_3 \tag{4.23}$$

For the dominated convergence theorem we have $\lim_{\tau \to t} I_1 = 0$, then $I_1 \in L^1([T_0, T_1])$. Using (4.8) and the Hölder-continuity condition for f we have:

$$|I_3| \leq \frac{C}{(t-\tau)^{1-\alpha/2}}$$

Finally, using theorem (4.2.4) we obtain:

$$|I_2| = |f(x,\tau)| \int_{\bar{\Omega}} \left| \frac{\partial^2 (\Gamma_{\xi,\tau} - \Gamma_{y,t})(x-\xi,t-\tau)}{\partial x_i \partial x_j} \right|_{y=x} d\xi \leq$$
$$\leq C \frac{|f(x,\tau)|}{|t-\tau|^{1-\alpha/2}} \int_{\bar{\Omega}} \Gamma^{(\lambda_{min})}(x-\xi,t-\tau) d\xi \leq C \frac{|f(x,\tau)|}{|t-\tau|^{1-\alpha/2}} \tag{4.24}$$

We have therefore shown that:

$$\left| \frac{\partial^2 J(x,t,\tau)}{\partial x_i \partial x_j} \right| \leq \frac{K}{(t-\tau)^{1-\alpha/2}} \in L^1([T_0, T_1])$$

Now we can use the dominated convergence theorem and demonstrate (4.21).The continuity of the function $\frac{\partial^2 V(x,t)}{\partial x_i \partial x_j}$ follows from the dominated convergence theorem. □

In a similar way it can be demonstrated the following:

Theorem 4.2.6. *Let $f(x,t)$ be a Hölder-continuous function of order α defined on Q. Then $\frac{\partial V(x,t)}{\partial t}$ is continuous on Q and it holds:*

$$\frac{\partial V(x,t)}{\partial t} = f(x,t) + \int_{T_0}^{t} \int_{\bar{\Omega}} \sum_{i,j=1}^{n} a_{i,j}(\xi,\tau) \frac{\partial^2 \Gamma_{\xi,\tau}(x-\xi,t-\tau)}{\partial x_i \partial x_j} f(\xi,\tau) d\xi d\tau \tag{4.25}$$

The proof of theorem (4.2.6) uses the properties of the foundamental solution $\Gamma_{\xi,\tau}$ and makes calculations similar to the previous theorems to prove the continuity of the function $\frac{\partial V(x,t)}{\partial t}$.

4.3 Construction of the parametrix

In this section we will deal with the recursive construction of the fundamental solution of the parabolic operator L.

We will indicate with $\Gamma(x,t)$ the foundamental solution of the parabolic operator:

$$L = \sum_{i,j=1}^{n} a_{i,j}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} - \frac{\partial}{\partial t}$$

so that $L\Gamma = \delta$.

Suppose that the functions $(a_{ij}(x,t))_{1\leq i,j\leq n}$ are Hölder-continuous, so that $\forall (x,t),(y,\tau) \in Q$, $(x,t) \neq (y,\tau)$ we have:

$$\mid a_{ij}(x,t) - a_{ij}(y,\tau) \mid \leq K(\mid x-y \mid^{\alpha} + \mid t-\tau \mid^{\alpha/2}) \tag{4.26}$$

For $(\xi,\tau) \in Q$ fixed, we will indicate with $\Gamma_{\xi,\tau}(x,t)$ the foundamental solution of the following parabolic operator with constant coefficients:

$$L_{\xi,\tau} = \sum_{i,j=1}^{n} a_{i,j}(\xi,\tau)\frac{\partial^2}{\partial x_i \partial x_j} - \frac{\partial}{\partial t}$$

so that $\Gamma_{\xi,\tau}L_{\xi,\tau} = \delta$.

The parametrix method consists in the construction of the foundamental solution Γ of L using the parametrix $\Gamma_{\xi,\tau}$. Consider:

$$\begin{cases} K_1 = \Gamma_{\xi,\tau} \\ R_1^{\xi,\tau} = (L - L_{\xi,\tau})\Gamma_{\xi,\tau} \end{cases} \tag{4.27}$$

We obtain:

$$LK_1 = L\Gamma_{\xi,\tau} = (L-L_{\xi,\tau})\Gamma_{\xi,\tau} + L_{\xi,\tau}\Gamma_{\xi,\tau} = R_1^{\xi,\tau} + \delta \tag{4.28}$$

Now we define:

$$\begin{cases} K_2 = K_1 - \Gamma_{\xi,\tau} * R_1^{\xi,\tau} \\ R_2^{\xi,\tau} = -R_1^{\xi,\tau} * R_1^{\xi,\tau} \end{cases} \tag{4.29}$$

If we now consider LK_2, we have:

$$\begin{aligned} LK_2 =& LK_1 - L\Gamma_{\xi,\tau} * R_1^{\xi,\tau} = R_1^{\xi,\tau} + \delta - (L-L_{\xi,\tau})\Gamma_{\xi,\tau} * R_1^{\xi,\tau} - L_{\xi,\tau}\Gamma_{\xi,\tau} * R_1^{\xi,\tau} = \\ =& R_1^{\xi,\tau} + \delta - R_1^{\xi,\tau} * R_1^{\xi,\tau} - R_1^{\xi,\tau} = \delta + R_2^{\xi,\tau} \end{aligned} \tag{4.30}$$

We define:

$$\begin{cases} K_3 = K_2 - \Gamma_{\xi,\tau} * R_2^{\xi,\tau} \\ R_3^{\xi,\tau} = -R_1^{\xi,\tau} * R_2^{\xi,\tau} \end{cases} \tag{4.31}$$

then we get:

$$LK_3 = LK_2 - L\Gamma_{\xi,\tau} * R_2^{\xi,\tau} = R_2^{\xi,\tau} + \delta - R_1^{\xi,\tau} * R_2^{\xi,\tau} - R_2^{\xi,\tau} = \delta + R_3^{\xi,\tau}$$

Iterating, we have:

$$\begin{cases} K_j = K_{j-1} - \Gamma_{\xi,\tau} * R_{j-1}^{\xi,\tau} \\ R_j^{\xi,\tau} = -R_1^{\xi,\tau} * R_{j-1}^{\xi,\tau} \end{cases} \tag{4.32}$$

and:

$$LK_j = \delta + R_j \tag{4.33}$$

From equation (4.33) we can clearly understand that, if we consider the limit for $j \to +\infty$ and we demonstrate that $\lim_{j\to+\infty} R_j^{\xi,\tau} = 0$, then $\lim_{j\to+\infty} K_j$ is the foundamental solution of the parabolic operator L.

We rewrite K_j in the following way:

$$\begin{aligned} K_j &= K_{j-1} - \Gamma_{\xi,\tau} * R_{j-1}^{\xi,\tau} = K_{j-2} - \Gamma_{\xi,\tau} * R_{j-2}^{\xi,\tau} - \Gamma_{\xi,\tau} * R_{j-1}^{\xi,\tau} = \\ &= K_1 - \Gamma_{\xi,\tau} * R_1^{\xi,\tau} - \Gamma_{\xi,\tau} * R_2^{\xi,\tau} ... - \Gamma_{\xi,\tau} * R_{j-1}^{\xi,\tau} = \Gamma_{\xi,\tau} - \Gamma_{\xi,\tau} * \phi \end{aligned} \tag{4.34}$$

with:

$$\phi = \sum_{k=1}^{j-1} R_k^{\xi,\tau} \tag{4.35}$$

$$\begin{cases} R_1^{\xi,\tau} = & (L - L_{\xi,\tau})\Gamma_{\xi,\tau} \\ R_k^{\xi,\tau} = & -R_1^{\xi,\tau} * R_{k-1}^{\xi,\tau} \quad for\ k \geq 2 \end{cases} \tag{4.36}$$

It's clear now that the foundamental solution Γ of L will be:

$$\lim_{j\to+\infty} K_j = \Gamma_{\xi,\tau} - \Gamma_{\xi,\tau} * \sum_{j=1}^{+\infty} R_j^{\xi,\tau}$$

So the problem now becomes the study of the convergence of the series ϕ. First we have to prove some inequalities which will serve us to demonstrate the convergence of the series.

Lemma 4.3.1. *It holds:*

$$\frac{\|x-\xi\|^2}{t-\tau} + \frac{\|y-\mu\|^2}{\frac{(t-\sigma)(\sigma-\tau)}{t-\tau}} = \frac{\|y-x\|^2}{t-\sigma} + \frac{\|y-\xi\|^2}{\sigma-\tau} \tag{4.37}$$

where:

$$\mu = \frac{(t-\sigma)\xi + (\sigma-\tau)x}{t-\tau}$$

Theorem 4.3.2. $\forall \sigma \in (\tau, t)$

$$\int_{\bar{\Omega}} \Gamma_{y,\sigma}(x-y,t-\sigma)\Gamma_{\xi,\tau}(y-\xi,\sigma-\tau)dy \leq C\Gamma_{\xi,\tau}^{(\lambda_{min})}(x-\xi,t-\tau) \quad (4.38)$$

with: $C = \frac{\lambda_{max}}{\lambda_{min}}$

Proof: Using Lemma (4.3.1) and (4.14) we have:

$$\begin{aligned}
&\int_{\bar{\Omega}} \Gamma_{y,\sigma}(x-y,t-\sigma)\Gamma_{\xi,\tau}(y-\xi,\sigma-\tau)dy = \\
&= \int_{\bar{\Omega}} \frac{\sqrt{|A^{-1}(y,\sigma)A^{-1}(\xi,\tau)|}}{(4\pi)^n(t-\sigma)^{n/2}(\sigma-\tau)^{n/2}} e^{-\frac{1}{4}(\frac{\|y-x\|^2_{y,\sigma}}{t-\sigma}+\frac{\|y-\xi\|^2_{\xi,\tau}}{\sigma-\tau})}dy \leq \\
&\leq \int_{\bar{\Omega}} \frac{\sqrt{|A^{-1}(y,\sigma)A^{-1}(\xi,\tau)|}}{(4\pi)^n(t-\sigma)^{n/2}(\sigma-\tau)^{n/2}} e^{-\frac{\lambda_{min}}{4}(\frac{\|y-x\|^2}{t-\sigma}+\frac{\|y-\xi\|^2}{\sigma-\tau})}dy \leq \\
&\leq \int_{\bar{\Omega}} \frac{\sqrt{|A^{-1}(y,\sigma)A^{-1}(\xi,\tau)|}}{(4\pi)^n(t-\sigma)^{n/2}(\sigma-\tau)^{n/2}} e^{-\frac{\lambda_{min}}{4}(\frac{\|x-\xi\|^2}{t-\tau}+\frac{\|y-\mu\|^2}{\frac{(t-\sigma)(\sigma-\tau)}{(t-\tau)}})}dy \leq \\
&\leq \frac{\sqrt{|A^{-1}(\xi,\tau)|}e^{-\frac{\lambda_{min}}{4}\frac{\|x-\xi\|^2}{t-\tau}}}{(4\pi)^{n/2}(t-\sigma)^{n/2}(\sigma-\tau)^{n/2}} \int_{\bar{\Omega}} \frac{\sqrt{|A^{-1}(y,\sigma)|}e^{-\frac{\lambda_{min}}{4}\frac{\|y-\mu\|^2}{\frac{(t-\sigma)(\sigma-\tau)}{t-\tau}}}}{(4\pi)^{n/2}(\frac{(t-\sigma)(\sigma-\tau)}{t-\tau})^{n/2}}dy \leq \\
&\leq \frac{\sqrt{|A^{-1}(\xi,\tau)|}e^{-\frac{\lambda_{min}}{4}\frac{\|x-\xi\|^2}{t-\tau}}\frac{\lambda_{max}}{\lambda_{min}}}{(4\pi)^{n/2}(t-\tau)^{n/2}} = \frac{\lambda_{max}}{\lambda_{min}}\Gamma_{\xi,\tau}^{(\lambda_{min})}(x-\xi,t-\tau)
\end{aligned} \quad (4.39)$$

where we have used:

$$\int_{\bar{\Omega}} \frac{\sqrt{|A^{-1}(y,\sigma)|}e^{-\frac{\lambda_{min}}{4}\frac{\|y-\mu\|^2}{\frac{(t-\sigma)(\sigma-\tau)}{t-\tau}}}}{(4\pi)^{n/2}(\frac{(t-\sigma)(\sigma-\tau)}{t-\tau})^{n/2}}dy \leq \frac{\lambda_{max}}{\lambda_{min}}$$

□

Theorem 4.3.3. *If* $\alpha < 1$, $\beta < 1$ *then* $\forall \sigma \in (\tau, t)$

$$\int_{\tau}^{t} \frac{d\sigma}{|\, t-\sigma\,|^{\alpha}|\, \sigma-\tau\,|^{\beta}} \leq \frac{cost}{|\, t-\tau\,|^{\alpha+\beta-1}} \quad (4.40)$$

Proof:

$$\int_{\tau}^{t} \frac{d\sigma}{|t-\sigma|^{\alpha}|\sigma-\tau|^{\beta}} = \int_{\tau}^{\frac{t+\tau}{2}} \frac{d\sigma}{|t-\sigma|^{\alpha}|\sigma-\tau|^{\beta}} + \int_{\frac{t+\tau}{2}}^{t} \frac{d\sigma}{|t-\sigma|^{\alpha}|\sigma-\tau|^{\beta}} \quad (4.41)$$

In the first integral we have that $|t-\sigma| \geq |t+\tau|/2$, while in the second integral it holds $|\sigma-\tau| \geq |t+\tau|/2$. Then we obtain:

$$\int_{\tau}^{t} \frac{d\sigma}{|t-\sigma|^{\alpha}|\sigma-\tau|^{\beta}} \leq \frac{2^{\alpha}}{|t-\tau|^{\alpha}} \int_{\tau}^{\frac{t+\tau}{2}} \frac{d\sigma}{|\sigma-\tau|^{\beta}} + \frac{2^{\beta}}{|t-\tau|^{\beta}} \int_{\frac{t+\tau}{2}}^{t} \frac{d\sigma}{|t-\sigma|^{\alpha}} =$$
$$= \frac{2^{\alpha}[|\sigma-\tau|^{1-\beta}]_{\tau}^{\frac{t+\tau}{2}}}{|t-\tau|^{\alpha}(1-\beta)} + \frac{2^{\beta}[|t-\sigma|^{1-\alpha}]_{\frac{t+\tau}{2}}^{t}}{|t-\tau|^{\beta}(1-\alpha)} = \frac{C}{|t-\tau|^{\alpha+\beta-1}} \tag{4.42}$$

□

Theorem 4.3.4. *It holds:*

$$\mid R_j^{\xi,\tau}(x,t) \mid \leq \frac{C^j}{\mid t-\tau \mid^{1-j\alpha/2}} \Gamma_{\xi,\tau}^{(\lambda_{min})^j}(x-\xi, t-\tau) \quad \forall j \geq 1 \tag{4.43}$$

Proof: We prove the theorem by induction: for $j=1$ we have:

$$\begin{aligned}
\mid R_1^{\xi,\tau}(x,t) \mid &= \mid (L-L_{\xi,\tau})\Gamma_{\xi,\tau}(x-\xi,t-\tau) \mid \leq \\
&\leq \sum_{i,j=1}^{n} \mid a_{ij}(x,t) - a_{ij}(\xi,\tau) \mid \frac{\partial^2 \Gamma_{\xi,\tau}(x-\xi,t-\tau)}{\partial x_i \partial x_j} \leq \\
&\leq nM\lambda_{max}\left(\frac{\mid t-\tau\mid^{\alpha/2}}{\mid t-\tau \mid}\Gamma_{\xi,\tau}^{(\lambda_{min})}(x-\xi,t-\tau)\right) \leq \\
&\leq \frac{C^1}{\mid t-\tau \mid^{1-\alpha/2}}\Gamma_{\xi,\tau}^{(\lambda_{min})}(x-\xi,t-\tau)
\end{aligned} \tag{4.44}$$

Now suppose the thesis holds for $j-1$ and consider:

$$\begin{aligned}
&\mid R_j^{\xi,\tau}(x,t) \mid = \left| R_1 * R_{j-1}^{\xi,\tau}(x,t) \right| = \left| \int_{\tau}^{t} \int_{\bar{\Omega}} (L-L_{y,\sigma})\Gamma_{y,\sigma}(x-y,t-\sigma) R_{j-1}^{\xi,\tau}(y,\sigma) dy d\sigma \right| \leq \\
&\leq \left| \int_{\tau}^{t} \int_{\bar{\Omega}} (L-L_{y,\sigma})\Gamma_{y,\sigma}(x-y,t-\sigma) \frac{C^{j-1}}{\mid \sigma-\tau \mid^{1-(j-1)\alpha/2}} \Gamma_{\xi,\tau}^{(\lambda_{min}^{j-1})}(y-\xi,\sigma-\tau) dy d\sigma \right| \leq \\
&\leq \mid \int_{\tau}^{t} \int_{\bar{\Omega}} \sum_{i,j=1}^{n} \mid a_{ij}(x,t) - a_{ij}(y,\sigma) \mid \frac{\partial^2 \Gamma_{y,\sigma}(x-y,t-\sigma)}{\partial x_i \partial x_j} \frac{C^{j-1}}{\mid \sigma-\tau \mid^{1-(j-1)\alpha/2}} \\
&\Gamma_{\xi,\tau}^{(\lambda_{min}^{j-1})}(y-\xi,\sigma-\tau) dy d\sigma \mid \leq \\
&\leq \left| \int_{\tau}^{t} \frac{C^j}{\mid t-\sigma \mid^{\alpha/2} \mid \sigma-\tau \mid^{1-(j-1)\alpha/2}} \int_{\bar{\Omega}} \Gamma_{y,\sigma}(x-y,t-\sigma) \Gamma_{\xi,\tau}^{(\lambda_{min}^{j-1})}(y-\xi,\sigma-\tau) dy d\sigma \right| \leq \\
&\leq C^j \Gamma_{\xi,\tau}^{(\lambda_{min})^j}(x-\xi,t-\tau) \int_{\tau}^{t} \frac{d\sigma}{\mid t-\sigma \mid^{1-\alpha/2} \mid \sigma-\tau \mid^{1-(j-1)\alpha/2}} \leq \\
&\leq \frac{C^j \Gamma_{\xi,\tau}^{(\lambda_{min})^j}(x-\xi,t-\tau)}{\mid t-\tau \mid^{1-\alpha j/2}}
\end{aligned} \tag{4.45}$$

where we have used the property of the functions a_{ij} and the theorems (4.3.2) and (4.3.3). $\square$

Since $\alpha > 0$, from theorem (4.3.4) we have that:

$$\exists j_0 \in \mathbb{N}: \quad 1 - \frac{j_0}{2}\alpha < 0$$

that is there exists j_0 such that $\forall j \geq j_0$

$$\left| R_j^{\xi,\tau}(x,t) \right| \leq K_0 \left| t-\tau \right|^{\gamma} \mathit{\Gamma}_{\xi,\tau}(x-\xi, t-\tau) \quad \gamma > 0 \tag{4.46}$$

We want to prove another estimation for $R_j^{\xi,\tau}$:

Theorem 4.3.5. $\forall m \in \mathbb{N}$

$$\left| R_{j_0+m}^{\xi,\tau}(x,t) \right| \leq K_0 \frac{(K(t-\tau)^{\alpha/2})^m}{\gamma(1+m\alpha/2)}$$

where $\gamma(x)$ *is the Gamma Euler function.*

Proof: The proof is made by induction. If we consider $m = 0$ the thesis follows from (4.46).
Suppose the thesis holds for m and consider $\left| R_{j_0+m}^{\xi,\tau}(x,t) \right|$.

$$\left| R_{j_0+m}^{\xi,\tau}(x,t) \right| \leq \left| \int_\tau^t \int_{\bar{\Omega}} \frac{C}{\left| t-\sigma \right|^{1-\alpha/2}} \mathit{\Gamma}_{y,\sigma}^{(\lambda_{min})}(x-y, t-\sigma) R_{j_0+m-1}^{\xi,\tau}(y,\sigma) dy d\sigma \right| \leq$$
$$\leq K_0 \frac{C^m}{\gamma(1+(m-1)\frac{\alpha}{2}} \int_\tau^t \int_{\bar{\Omega}} \mathit{\Gamma}_{y,\sigma}^{(\lambda_{min})}(x-y, t-\sigma) \frac{dy d\sigma}{\left| t-\sigma \right|^{1-\alpha/2} \left| \sigma-\tau \right|^{-(m-1)\alpha/2}} \tag{4.47}$$

We know that:

$$\int_{\bar{\Omega}} \mathit{\Gamma}_{y,\sigma}^{(\lambda_{min})}(x-y, t-\sigma) dy \leq \frac{1}{\lambda_{min}^n} \tag{4.48}$$

Using the following property of the Gamma Euler's function:

$$\int_0^1 (1-\rho)^{a-1} \rho^{b-1} d\rho = \frac{\gamma(a)\gamma(b)}{\gamma(a+b)}$$

with:

$$a = \alpha/2, \quad b = 1 + (m-1)\frac{\alpha}{2}, \quad \rho = \frac{\sigma-\tau}{t-\tau} \Rightarrow d\rho = \frac{d\sigma}{t-\tau}$$

We obtain:

$$\int_\tau^t \frac{d\sigma}{\left| t-\sigma \right|^{1-\alpha/2} \left| \sigma-\tau \right|^{-(m-1)\alpha/2}} = \frac{\gamma(\alpha/2)\gamma(1+(m-1)\frac{\alpha}{2})}{\gamma(1+m\frac{\alpha}{2})} \left| t-\tau \right|^{m\alpha/2} \tag{4.49}$$

The thesis follows from (4.48) and from the previous equation, infact we have:

$$\left|R^{\xi,\tau}_{j_0+m}(x,t)\right| \le K_0 \frac{C^m \gamma(\alpha/2)}{\lambda^n_{min}\gamma(1+m\frac{\alpha}{2})} \left|t-\tau\right|^{m\alpha/2}. \tag{4.50}$$

□We are finally ready to demonstrate the following:

Theorem 4.3.6. *The series* $\sum_{j=j_0}^{+\infty} R_j^{\xi,\tau}(x,t)$ *is convergent.*

Proof: From theorem (4.3.5) we have:

$$\sum_{j=j_0}^{+\infty} R_j^{\xi,\tau}(x,t) \le \sum_{m=0}^{+\infty} K_0 \frac{(K(t-\tau)^{\alpha/2})^m}{\gamma(1+m\alpha/2)} \tag{4.51}$$

The series is convergent for the ratio test

$$\frac{a_{m+1}}{a_m} = \frac{K(t-\tau)^{\alpha/2}\gamma(1+(m-1)\alpha/2)}{\gamma(1+m\alpha/2)} \to_{m\to+\infty} 0$$

thanks to the property of the Gamma Euler function. □

Theorem 4.3.7. *The following inequality holds:*

$$\phi_{\xi,\tau}(x,t) \le \begin{cases} \frac{2K0}{\alpha} e^{K2/\alpha} + \frac{C}{|t-\tau|^{1-\alpha/2}} & if \mid t-\tau \mid \le 1 \\ \frac{2K_0 K^{2/\alpha}(t-\tau)}{\alpha} e^{K2(t-\tau)/\alpha} + C^{j_0} & if \mid t-\tau \mid > 1 \end{cases} \tag{4.52}$$

Proof: Let r be the integer part of $m\alpha/2$. Then for the properties of the Gamma Euler's function we have:

$$\gamma(1+m\alpha/2) \ge \gamma(1+r) = r!$$

Using the inequalities found in theorems (4.3.4) and (4.3.5) we have:

$$\begin{aligned} \phi_{\xi,\tau}(x,t) &= \sum_{j=1}^{j_0-1} R_j^{\xi,\tau}(x,t) + \sum_{j=j_0}^{+\infty} R_j^{\xi,\tau}(x,t) \le \\ &\le \sum_{j=1}^{j_0-1} \frac{C^j}{\left|t-\tau\right|^{1-j\alpha/2}} + \sum_{j=j_0}^{+\infty} K_0 \frac{(K(t-\tau)^{\alpha/2})^m}{\gamma(1+m\alpha/2)} \le \\ &\le \sum_{j=1}^{j_0-1} \frac{C^j}{\left|t-\tau\right|^{1-j\alpha/2}} + \sum_{j=j_0}^{+\infty} K_0 \frac{(K^{2/\alpha}(t-\tau))^{m\alpha/2}}{r!} \end{aligned} \tag{4.53}$$

Then if $|t-\tau| \leq 1$:

$$
\begin{aligned}
\phi_{\xi,\tau}(x,t) &\leq \frac{C}{|t-\tau|^{1-\alpha/2}} + \sum_{j=j_0}^{+\infty} K_0 \frac{(K^{2/\alpha}(t-\tau))^{m\alpha/2}}{r!} \leq \\
&\leq \frac{C}{|t-\tau|^{1-\alpha/2}} + \frac{2K_0}{\alpha} \sum_{r=0}^{+\infty} \frac{(K^{2/\alpha}(t-\tau))^{r}}{r!} = \\
&= \frac{C}{|t-\tau|^{1-\alpha/2}} + \frac{2K_0}{\alpha} e^{K^{2/\alpha}(t-\tau)} \leq \frac{C}{|t-\tau|^{1-\alpha/2}} + \frac{2K_0}{\alpha} e^{K^{2/\alpha}}
\end{aligned}
\tag{4.54}
$$

If $|t-\tau| > 1$ we have:

$$
\begin{aligned}
\phi_{\xi,\tau}(x,t) &\leq \frac{C^{j_0}}{|t-\tau|^{1-j_0\alpha/2}} + \sum_{j=j_0}^{+\infty} K_0 \frac{(K^{2/\alpha}(t-\tau))^{m\alpha/2}}{r!} \leq \\
&\leq C^{j_0} + \frac{2K_0}{\alpha} \sum_{r=0}^{+\infty} \frac{(K^{2/\alpha}(t-\tau))^{r+1}}{r!} = C^{j_0} + \frac{2K_0 K^{2/\alpha}(t-\tau)}{\alpha} e^{K^{2/\alpha}(t-\tau)}.
\end{aligned}
\tag{4.55}
$$

□Now we have to check if the function we have defined is actually the foundamental solution of the parabolic operator L.

Theorem 4.3.8. *If we define $\bar{\Gamma}(x,t,\xi,\tau) = \Gamma_{\xi,\tau} - \Gamma_{\xi,\tau} * \phi$ then $\bar{\Gamma}$ is the foundamental solution of the parabolic operator L.*

Proof:

$$
\begin{aligned}
&L\bar{\Gamma}_{\xi,\tau}(x,t) = L\Gamma_{\xi,\tau}(x,t) - L\Gamma_{\xi,\tau} * \phi(x,t) = \\
&= R_1^{\xi,\tau} + \delta - L\int_\tau^t \int_{\bar{\Omega}} \Gamma_{y,\sigma}(x-y,t-\sigma) \sum_{j=1}^{+\infty} R_j^{\xi,\tau}(y,\sigma)dyd\sigma = \\
&= R_1^{\xi,\tau} + \delta - \lim_{N\to+\infty} L\int_\tau^t \int_{\bar{\Omega}} \Gamma_{y,\sigma}(x-y,t-\sigma) \sum_{j=1}^{N} R_j^{\xi,\tau}(y,\sigma)dyd\sigma
\end{aligned}
\tag{4.56}
$$

If we focus on the last part of the equation:

$$\begin{aligned}
&\lim_{N\to+\infty} L\int_{\tau}^{t}\int_{\bar{\Omega}} \mathit{\Gamma}_{y,\sigma}(x-y,t-\sigma)\sum_{j=1}^{N} R_j^{\xi,\tau}(y,\sigma)dyd\sigma = \\
&= \lim_{N\to+\infty} L(\mathit{\Gamma}_{\xi,\tau} * \sum_{j=1}^{N} R_j^{\xi,\tau})(x,t) \\
&= \lim_{N\to+\infty} (L-L_{\xi,\tau})(\mathit{\Gamma}_{\xi,\tau} * \sum_{j=1}^{N} R_j^{\xi,\tau})(x,t) + L_{\xi,\tau}(\mathit{\Gamma}_{\xi,\tau} * \sum_{j=1}^{N} R_j^{\xi,\tau})(x,t) = \\
&= \lim_{n\to+\infty} R_1^{\xi,\tau} * \sum_{j=1}^{N} R_j^{\xi,\tau} + \sum_{j=1}^{N} R_j^{\xi,\tau} = \lim_{N\to+\infty} R_{N+1}^{\xi,\tau} + \sum_{j=1}^{N} R_j^{\xi,\tau} = \\
&= \sum_{j=1}^{+\infty} R_j^{\xi,\tau}
\end{aligned} \tag{4.57}$$

Then we get:

$$L\bar{\mathit{\Gamma}}_{\xi,\tau}(x,t) = R_1^{\xi,\tau} + \delta - \sum_{j=1}^{+\infty} R_j^{\xi,\tau} = \delta - \lim_{j\to+\infty} R_j^{\xi,\tau} = \delta. \tag{4.58}$$

□

Chapter 5

Parametrix method in Lie groups

In this chapter we want to extend the results found in chapter (4) in a more general set. In particular we will analyze the case of parabolic operators in Lie group.

We will find the foundamental solution of a special parabolic operator defined on $\mathbb{R}^2 \times S^1$. This parabolic operator is defined as:

$$L = X_1^2 + X_2^2 - \partial_t$$

where X_1, X_2 are the vector fields of degree 1 defined describing the model of the visual cortex, (2.47). The main problem in this case is that we have estimates for the foundamental solution of a parabolic operator in a Lie Group only if the group is an homogeneous Carnot group. In our case $\mathbb{R}^2 \times S^1$ it's not an homogeneous Carnot group, so we have to apply the parametrix method in order to find the foundamental solution of the parabolic operator. Analogously to what was done in the Euclidean case we will consider an approximate parabolic operator, whose vector fields will be defined on an homogeneous Carnot group. This will permit us to illustrate some important inequalities for the foundamental solution of the considered operator. After that, similarly to what we have done in the Euclidean case, we will write the foundamental solution of the parabolic operator L as:

$$\Gamma = \Gamma_{\xi_0} - \Gamma_{\xi_0} * \phi$$

where ϕ is the series $\sum_{j=0}^{+\infty} R_j(\xi_0, \tau, \cdot)$. In order to prove that Γ is the fundamental solution of L we'll have to prove that the series ϕ is convergent.

The chapter is organized as follows: in the first section we will give the main definitions that have been mentioned above and we will recall the definition of the group law which we will consider. In the second section we will define the approximate vector fields and we will define the approximate parabolic operator which will be used in the parametrix method. In particular we will see that the vector fields frozen at a point ξ are defined on a Carnot group. In the third section we will see some of the main properties of the foundamental solution of a parabolic operator on a Carnot Group, while in the fourth section we will demonstrate the continuity properties of the fucntions $R_j^{\xi_0}$ and of the convolutional operator $\Gamma_{\xi_0} * R_j^{\xi_0}$. Finally in the last section we will construct the parametrix and demonstrate that the function Γ defined as limit of the functions K_j is the foundamental solution of the L operator.

5.1 Definition of the space

In order to study the parametrix method in a Lie group we need first of all to introduce some main definitions.

Definition 5.1.1. A Carnot group (or stratified group) $\mathbb{H}$ is a connected and simply connected Lie group whose Lie algebra $\mathfrak{h}$ admits a stratification, i.e. a direct sum decomposition:

$$\mathfrak{h} = V_1 \oplus ... \oplus V_r$$

such that $[V_1, V_{i-1}] = V_i$ if $2 \leq i \leq r$ and $[V_i, V_r] = 0$

The stratification implies that the Lie Algebra $\mathfrak{h}$ is nilpotent of step r. Any Carnot group is isomorphic to an homogeneous group.
This fact permit us to give another equivalent definition of a Carnot group:

Definition 5.1.2. Let $\circ$ be an assigned Lie group law on $\mathbb{R}^N$. Suppose $\mathbb{R}^N$ is endowed with a homogeneous structure by a given family of automorphisms $\{\delta_\lambda\}_{\lambda>0}$,called *dilations*, of the form:

$$\delta_\lambda(x) = \delta_\lambda(x^{(1)}, ..., x^{(r)}) = (\lambda x^{(1)}, \lambda^2 x^{(2)}, ..., \lambda^r x^{(r)}).$$

Here $x^{(i)} \in \mathbb{R}^{N_i}$ for $i = 1, ..., r$ and $N_1 + ... + N_r = N$. We denote by $\mathfrak{g}$ the Lie algebra of $(\mathbb{R}^N, \circ)$, i.e. the Lie algebra the left invariant vector fields on $\mathbb{R}^N$. For $i = 1, ..., N_1$ let X_i be the unique vector field in $\mathfrak{g}$ that agrees at the origin with $\partial/\partial x_i$. If the Lie algebra generated by $X_1, ..., X_{N_1}$ is the whole $\mathfrak{g}$, we call $\mathbb{G} = (\mathbb{R}^N, \circ, \delta_\lambda)$ a homogeneous Carnot group.
We also say that $\mathbb{G}$ is of step r and has $m = N_1$ generators.

It is not difficult to verify that any homogeneous Carnot group is a Carnot group according to the classical definition. On the other hand, up to isomorphism, the opposite implication is also true. Recall that we denote by

$$Q = \sum_{i=1}^{r} deg(X_i)$$

the homogeneous dimension of $\mathbb{G}$. Then $L(\delta_\lambda(E)) = \lambda^Q L(E)$ for any measurable set E, where $L(.)$ denotes the Lebesgue measure on $\mathbb{R}^N$. This measure is invariant with respect to the left and right translations on $\mathbb{G}$. If $Q \leq 3$, then $\mathbb{G}$ is the ordinary Euclidean group $(\mathbb{R}^Q, +)$ and $\Delta_{\mathbb{G}}$ is the classical Laplace operator.

Let's now consider our particular case. In chapter (2) we have considered the group $\mathbb{R}^2 \times S^1$ with the group law defined in (2.45) and we have defined the following vector fields:

$$\begin{aligned} X_1 &= cos\theta\partial_x + sin\theta\partial_y \\ X_2 &= \partial_\theta \end{aligned} \tag{5.1}$$

We have then proved that $Lie(X_1, X_2)$ verifies the Hörmander condition. Infact, if we consider the commutator X_3:

$$X_3 = [X_1, X_2] = -sin\theta\partial_x + cos\theta\partial_y$$

From the formula above we can easily conclude that X_1, X_2, X_3 are linearly indipendent then $dim(Lie(X_1, X_2)) = 3$. In particular these vector fields have the following degrees:

$$\begin{aligned} deg(X_1) &= 1 \\ deg(X_2) &= 1 \\ deg(X_3) &= 2 \end{aligned} \tag{5.2}$$

then $Q = 4$.

Let's now consider the sub-Laplacian operator defined in (2.2.42).The operator we want to consider on $\mathbb{R}^2 \times S^1 = \mathbb{G}$ is:

$$\Delta_{\mathbb{G}} = X_1^2 + X_2^2$$

Then, the associated heat operator is defined as:

$$L = \Delta_{\mathbb{G}} - \partial_t = X_1^2 + X_2^2 - \partial_t$$

The results we will show in the next sections are based on the fact that the parabolic operator is defined on a Carnot group. In our case the issue is that the vector fields X_1, X_2, X_3 are defined on $\mathbb{R}^2 \times S^1$, which is not a Carnot group, then is not equipped with a homogeneous structure.
Therefore, it results necessary to introduce an approximate operator which will be defined on a Carnot group. We will then look towards the homogeneity behavioural of vector fields in $\mathbb{R}^2 \times S^1$ in terms of the homogeneity properties of the stratified group of same dimension, namely the Heisenberg group $\mathbb{H}$. This procedure is similar to the one we have done in the Euclidean

case: in that case we had considered as approximate heat operator, the operator with constant coefficients. Thanks to the properties already known for the operators with constant coefficients, we had been able to apply the parametrix method and find the foundamental solution of the original heat operator. Therefore in the next section we will introduce an approximate heat operator, which will be defined, up to isomorphism, on a Carnot group.

5.2 Approximate vector fields

In this section we want to define, for a fixed point ξ, the approximate vector fields. Then we will show that , up to diffeomorphism, these approximate vector fields are defined on the Heisenberg group, which is known to be a Carnot group.

Let's consider the vector fields $\{X_i\}_{i=1,2}$ defined in (5.1) and let $\xi_0 = (x_0, y_0, \theta_0) \in \mathbb{R}^2 \times S^1$ be a fixed point.

Definition 5.2.1. The approximate vector fields frozen at ξ_0 are:

$$\begin{aligned} X_{1,\xi_0} &= (cos\theta_0 - (\theta-\theta_0)sin\theta_0)\partial_x + (sin\theta_0 + (\theta-\theta_0)cos\theta_0)\partial_y \\ X_{2,\xi_0} &= X_2 = \partial_\theta \\ X_{3,\xi_0} &= [X_{1,\xi_0}, X_{2,\xi_0}] = -sin\theta_0\partial_x + cos\theta_0\partial_y. \end{aligned} \tag{5.3}$$

It's easy to verify that $dim(Lie(X_{1,\xi_0}, X_{2,\xi_0})) = 3$, then $\{X_{i,\xi_0}\}_{i=1,2,3}$ are Hörmander vector fields defined in terms of the first order Taylor development of the coefficients of the vector fields $\{X_i\}_{i=1,2,3}$.

We now want to show that the Lie algebra spanned by $X_{\xi_0,1}, X_{\xi_0,2}, X_{\xi_0,3}$ is the Heisenberg algebra.
First of all we need to introduce some main definitions.

Definition 5.2.2. The Heisenberg group $\mathbb{H}^n$ is the space $\mathbb{R}^{2n+1}$ equipped with the following group law:

$$(x,y,z)\cdot(x',y',z') = (x+x', y+y', z+z'+2\langle y,x'\rangle - 2\langle x,y'\rangle)$$

The group $\mathbb{H}^n$ equipped with the parabolic dilations $\delta_\lambda(z,t) = (\lambda z, \lambda^2 t), \lambda > 0$, carries the structure of homogeneous group.
A stratified basis of left invariant vector fields for the Lie algebra $\mathfrak{g}_n$ of $\mathbb{H}^n$ is given by:

$$\begin{aligned} X_j &= \partial_{x_j} + 2y_j\partial_t \quad j=1,...,n \\ Y_j &= \partial_{y_j} - 2x_j\partial_t \\ T &= \partial_t \end{aligned} \tag{5.4}$$

In our case we have that $n = 1$ and the vector fields are:

$$\begin{aligned} X_1 &= \partial_{e_1} - 2e_2\partial_{e_3} \\ X_2 &= \partial_{e_2} \end{aligned} \tag{5.5}$$

We want to find a diffeomorphism that turns the coordinates (x, y, θ) in (e_1, e_2, e_3). Let $\xi = (x, y, \theta)$ a point in the neighborhood of ξ_0. Let (e_1, e_2, e_3) the exponential coordinates of ξ with respect to the basis $X_{1,\xi_0}, X_{2,\xi_0}, X_{3,\xi_0}$, such that:

$$\xi = exp(e_1 X_{1,\xi_0} + e_2 X_{2,\xi_0} + e_3 X_{3,\xi_0})(\xi_0)$$

By the definition of exponential mapping $\gamma(1) = \xi = (x, y, \theta)$, where $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$ solves the Cauchy problem:

$$\begin{cases} \dot{\gamma} = e_1 X_{1,\xi_0} + e_2 X_{2,\xi_0} + e_3 X_{3,\xi_0} \\ \gamma(0) = \xi_0 \end{cases} \tag{5.6}$$

A trivial computation yields:

$$\begin{aligned} \gamma_1(1) &= x = x_0 + cos\theta_0 e_1 - \frac{1}{2} sin\theta_0 e_2 e_1 - sin\theta_0 e_3 \\ \gamma_2(1) &= y = y_0 + sin\theta_0 e_1 + \frac{1}{2} cos\theta_0 e_1 e_2 + cos\theta_0 e_3 \\ \gamma_3(1) &= \theta = \theta_0 + e_2 \end{aligned} \tag{5.7}$$

Plugging in $e_2 = \theta - \theta_0$, one obtain the linear map:

$$\begin{pmatrix} y - y_0 \\ x - x_0 \end{pmatrix} = \begin{pmatrix} sin\theta_0 + \frac{1}{2}(\theta - \theta_0) cos\theta_0 & cos\theta \\ cos\theta - \frac{1}{2}(\theta - \theta_0) sin\theta_0 & -sin\theta_0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \tag{5.8}$$

The inverse is:

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} sin\theta_0 & cos\theta_0 \\ cos\theta_0 - \frac{1}{2}(\theta - \theta_0) sin\theta_0 & -sin\theta_0 - \frac{1}{2}(\theta - \theta_0) cos\theta_0 \end{pmatrix} \begin{pmatrix} y - y_0 \\ x - x_0 \end{pmatrix} \tag{5.9}$$

The regular change of variables:

$$\begin{cases} e_1 = sin\theta_0 (y - y_0) + cos\theta_0 (x - x_0) \\ e_2 = \theta - \theta_0 \\ e_3 = cos\theta_0 (y - y_0) - sin\theta_0 (x - x_0) - \frac{1}{2} e_1 e_2 \end{cases} \tag{5.10}$$

provides a diffeomorphism:

$$\Phi_{\xi_0}(\xi) = \Phi_{x_0, y_0, \theta_0}(x, y, \theta) = (e_1, e_2, e_3)$$

turning the coordinates (x, y, θ) into the coordinates (e_1, e_2, e_3).

Definition 5.2.3. The approximate subelliptic heat operator is defined as:

$$L_{\xi_0} = X_{1,\xi_0}^2 + X_{2,\xi_0}^2 - \partial_t \tag{5.11}$$

This operator admits an explicit foundamental solution Γ_{ξ_0} which, up to the change of variables we've seen above, is the well known solution of the Heisenberg heat operator. If

$$L_{\mathbb{H}} = \Delta_{\mathbb{H}} - \partial_t = X_{\mathbb{H}}^2 + Y_{\mathbb{H}}^2 - \partial_t \tag{5.12}$$

is the Heisenberg heat operator in the standard basis, the related foundamental solution $\Gamma_{\mathbb{H}}$ is given by:

$$\Gamma_{\mathbb{H}}((x,y,\theta),t) = \frac{1}{(2\pi t)^4)} \int cos(\frac{y\tau}{t}) exp(\frac{1}{2}\tau coth 2\pi \frac{x^2+\theta^2}{t}) \frac{2\tau^4}{sinh(2\tau)^4} d\tau \tag{5.13}$$

Now that we have defined the space in which we will work and the operators we will use for the parametrix method we need, before illustrating the method, to give some known properties of the foundamental solution of a parabolic operator defined on a Carnot group. We refer to [8] for details.

5.3 Estimates for the foundamental solution of a parabolic operator on a Carnot group

In this section we will see some estimates for the foundamental solution of a parabolic operator defined on a Carnot group, with homogeneous dimension equal to Q.

Therefore, let's consider the following parabolic operator:

$$L = \sum_{i=1}^{m} X_i^2 - \partial_t \tag{5.14}$$

defined on $\mathbb{G} \times \mathbb{R}$, with $\mathbb{G}$ homogeneous Carnto group on $\mathbb{R}^N$. The first issue that need to be solved is the one of the existence of the foundamental solution. There are two main hypothesis which will permit us to demonstrate the existence and the estimates of the foundamental solution. The first hypothesis is that the vector fields $\{X_i\}$ satisfy the Hörmander condition. Infact, thanks to this hypothesis, we consider the following results.

Definition 5.3.1. If the vector fields $X_1, ..., X_m$ are of Hörmander type and selfadjoint, the associated subriemannian Laplacian and the heat operator (also called subelliptic Laplacian and Heat) are called Hörmander type operators.

Theorem 5.3.2. *A second order differential operator L is hypoelliptic if and only if there exist a fundamental solution of L which is C^∞ outside the origin.*

Furthermore, the Hörmander theorem (2.2.45) states the existence of a foundamental solution for the operator (5.14).

The second foundamental hypothesis we will need is that the heat kernel $\Gamma_{\mathbb{G}}$ satisfies some homogeneity properties, which can be stated in terms of the homogeneous dimension Q.

Definition 5.3.3. A differential operator D is homogeneous of degree r w.r.t.the dilations $\{\sigma_\lambda\}$ if and only if

$$D(u \circ \sigma_\lambda) = \lambda^r (Du) \circ \sigma_\lambda$$

for all $u \in C_0^\infty$ and $\lambda > 0$

Definition 5.3.4. $\mathbb{G} \times [0, T]$ is equipped with the family of parabolic dilations $\{\sigma_\lambda^L\}_{\lambda>0}$ given by:

$$\sigma_\lambda^L(\xi, \tau) = (\sigma_\lambda(\xi), \lambda^2 \tau)$$

The problem of the construction of the foundamental solution was solved in [7]. The foundamental solution Γ is constructed as a limit of a sequence of Green functions related to an increasing sequence of regular domains invading $\mathbb{R}^{N+1}$. Thanks to this construction we are able to prove the following:

Theorem 5.3.5. *There exists a smooth function Γ on $\mathbb{G} \times [0, T] - 0$ such that the foundamental solution of 5.14 is given by:*

$$\Gamma_{\mathbb{G}}(\xi, t, \eta, s) := \Gamma(\eta^{-1} \circ \xi, t - s)$$

The kernel $\Gamma_{\mathbb{G}}$ satisfies:

- *$\Gamma_{\mathbb{G}}$ is homogeneous of degree $-Q$ w.r.t. the parabolic dilations $\{\sigma_\lambda^L\}$ i.e.*
$$\Gamma_{\mathbb{G}}(\sigma_\lambda^L(\xi, t)) = \lambda^{-Q} \Gamma_{\mathbb{G}}(\xi, t);$$
- $\Gamma_{\mathbb{G}}(\xi, t, \eta, s) = \Gamma_{\mathbb{G}}(\eta, -s, \xi, t) = \Gamma_{\mathbb{G}}(\eta^{-1} \circ \xi, t - s, 0, 0)$
- *there exists a positive constant C such that:*
$$\Gamma_{\mathbb{G}}(\xi, t) \leq C(d_{\mathbb{G}}(\xi, 0) + |t|^{1/2})^{-Q}$$
where $d_{\mathbb{G}}$ is a control distance on $\mathbb{G}$ and Q is the homogeneous dimension of $\mathbb{G}$
- *(reproduction property)*
$$\Gamma_{\mathbb{G}}(\xi, t + s) = \int_{\mathbb{G}} \Gamma_{\mathbb{G}}(\eta^{-1} \circ \xi, t) \Gamma(\eta, s) d\eta \tag{5.15}$$

Proof: see [6, Theorem2.1].

For the foundamental solution the following Gaussian estimates hold:

Theorem 5.3.6. *There exist positive constants $C_1, C_{p,q}$ such that the foundamental solution $\Gamma_{\mathbb{G}}$ of the heat operator $L_{\mathbb{G}}$ and its derivatives satisfies:*

- $|\Gamma_{\mathbb{G}}(x,t)| \le C_1 t^{-Q/2} exp(\frac{-d^2_{\mathbb{G}}(x,0)}{C_1 t})$;
- $\left|X_{i_1}...X_{i_p}(\partial_t)^q \Gamma_{\mathbb{G}}(x,t)\right| \le C_{p,q} t^{-\frac{Q+p+2q}{2}} exp(\frac{-d^2_{\mathbb{G}}(x,0)}{C_1 t})$;

for every $p, q \in \mathbb{N}$

Proof:We set $A = \{x \in \mathbb{R}^N : d_{\mathbb{G}}(x) > 1\}$ and $\Omega = A \times (0,1)$. We want to compare $\Gamma(x,t)$ with the function $\omega(x,t) = exp(-\sigma(1-t)d^2_{\mathbb{G}}(x))$ in Ω. Here, σ is a positive constant to be chosen in the sequel. The following formula holds for radial functions $f(x) = F(d_{\mathbb{G}}(x))$.

$$Lf(x) = |\nabla_{\mathbb{G}} d_{\mathbb{G}}(x)|^2 \left(\frac{Q-1}{d_{\mathbb{G}}(x)} F'(d_{\mathbb{G}}(x)) + F''(d_{\mathbb{G}}(x))\right) \tag{5.16}$$

we have denoted with $\nabla_{\mathbb{G}}$ the subelliptic gradient $(x_1, ..., X_m)$. Hence,a direct computation shows that

$$\begin{aligned} L\omega(x,t) = \omega(x,t)\, |\nabla_{\mathbb{G}} d_{\mathbb{G}}(x)|^2 \{-2\sigma(Q-1)(1-t) + (4\sigma^2 d^2_{\mathbb{G}}(x)(1-t)^2 - 2\sigma(1-t))\}+ \\ -\sigma d^2_{\mathbb{G}}(x)\omega(x,t) \end{aligned} \tag{5.17}$$

For $(x,t) \in \Omega$, we obtain:

$$L\omega(x,t) \le (4\sigma^2 |\nabla_{\mathbb{G}} d_{\mathbb{G}}(x)|^2 - \sigma) d^2_{\mathbb{G}}(x)\omega(x,t) \le 0$$

if σ is chosen small enough (note that $|\nabla_{\mathbb{G}} d_{\mathbb{G}}(x)|$ is bounded). Recalling that $L\Gamma_{\mathbb{G}} = \delta$ in Ω, that $\Gamma_{\mathbb{G}}$ is continuous on $\bar{\Omega}$ and that $\Gamma_{\mathbb{G}}$ vanishes at infinity on $A \times \{0\}$ (note that $\Gamma_{\mathbb{G}}$ is the foundamental solution), from the maximum principle we infer that

$$\Gamma_{\mathbb{G}} \le c\omega \quad in\ \Omega$$

for a suitable constant $c > 0$. In particular, chosen $t = 1/2$, we obtain:

$$\Gamma_{\mathbb{G}}(x, \frac{1}{2}) \le c\ exp(-\sigma d^2_{\mathbb{G}}(x)/2), \quad if \;\; d_{\mathbb{G}}(x) \ge 1$$

By the homogeneity of $\Gamma_{\mathbb{G}}$, we then deduce:

$$\Gamma_{\mathbb{G}}(x,t) = (2t)^{-Q/2} \Gamma_{\mathbb{G}}(\sigma_{1/\sqrt{2t}} x, \frac{1}{2}) \le c t^{-Q/2} exp(-\frac{\sigma d^2_{\mathbb{G}}(x)}{4t}), \;\; if\ 0 < 2t \le d^2_{\mathbb{G}}(x)$$

On the other hand, if $d^2_{\mathbb{G}}(x) < 2t$, then

$$\Gamma_{\mathbb{G}}(x,t;\xi,\tau) \le c(d(x,\xi) + |t-\tau|^{1/2})^{-Q}$$

directly yields:

$$\Gamma_{\mathbb{G}}(x,t) \leq ct^{-Q/2} \leq ct^{-Q/2} exp(-\frac{d^2_{\mathbb{G}}(x)}{ct})$$

The second inequality is an easy consequence of the above inequality. This end complete the proof. □

Finally, it can be demonstrated that there exists a positive constant $c > 0$ such that:

$$\Gamma_{\mathbb{G}}(x,t) \geq c^{-1}t^{-Q/2} exp(-c\frac{d^2_{\mathbb{G}}(x)}{t}) \tag{5.18}$$

Before illustrating the parametrix method we need some continuity properties of the functions that we will use. Therefore in the next section we will prove that the series and the convolutional operator defined in the parametrix method are continuous.

5.4 Parametrix method

Now that we have introduced the main results for the foundamental solution on an homogeneous Carnot group, we want to prove the continuity of the functions that will be defined for the parametrix method.

First of all notice that the estimates found in section (5.3) hold for the foundamental solution Γ_{ξ_0} of the following heat operator:

$$L_{\xi_0} = X^2_{1,\xi_0} + X^2_{2,\xi_0} - \partial_t \tag{5.19}$$

Infact, as we've already proved, the operator is defined on the Heisenberg group, which is a homogeneous Carnot group.
In particular in this case we have that the homogeneous dimension is:

$$Q = \sum_i deg(X_i) = 1 + 1 + 2 = 4$$

Infact, in our case $deg(X_{1,\xi_0}) = deg(X_{2,\xi_0}) = 1$ and $deg(X_{3,\xi_0}) = 2$.
As theorem (5.3.6) shows, the following inequalities hold for the foundamental solution Γ_{ξ_0} of the operator L_{ξ_0}:

$$\begin{aligned}
|\Gamma_{\mathbb{G}}(x,t)| &\leq C_0 t^{-2} exp(\frac{-d^2_{\mathbb{G}}(x,0)}{C_0 t}) = C_1 E(x, C_0 t) \\
|X_{i,\xi_0}\Gamma_{\mathbb{G}}(x,t)| &\leq C_1 t^{-\frac{5}{2}} exp(\frac{-d^2_{\mathbb{G}}(x,0)}{C_1 t}) = \frac{C_1}{\sqrt{t}} E(x, C_1 t) \quad i = 1,2 \\
|X_{i,\xi_0}X_{j,\xi_0}\Gamma_{\mathbb{G}}(x,t)| &\leq C_2 t^{-3} exp(\frac{-d^2_{\mathbb{G}}(x,0)}{C_2 t}) = \frac{C_2}{t} E(x, C_2 t) \quad i,j = 1,2
\end{aligned} \tag{5.20}$$

where we have defined:

$$E(x,t) = t^{-2} exp(\frac{-d^2_{\mathbb{G}}(x,0)}{t})$$

As we have already seen in chapter (4), if the parametrix method is applied, the foundamental solution of the heat operator:

$$L = X_1^2 + X_2^2 - \partial_t$$

will be found as $\lim_{j\to+\infty} K_j$, where:

$$K_j(\cdot) = K_{j-1}(\cdot) - \Gamma_{\xi_0,\tau} * R_{j-1}(\xi_0, \tau; \cdot) \tag{5.21}$$

and

$$\begin{cases} R_1(\xi_0, \tau; \cdot) = (L - L_{\xi_0})\Gamma_{\xi_0,\tau}(\cdot) \\ R_j(\xi_0, \tau; \cdot) = -R_1 * R_{j-1}(\xi_0, \tau; \cdot) \quad for j > 1 \end{cases} \tag{5.22}$$

In an analogous way we obtain:

$$LK_j(\cdot) = \delta(\cdot) + R_j(\xi, \tau; \cdot)$$

and

$$\Gamma(\cdot) = \lim_{j\to+\infty} K_j(\cdot) = \Gamma_{\xi_0,\tau}(\cdot) - \Gamma_{\xi_0,\tau} * \phi(\cdot) \tag{5.23}$$

where

$$\phi(\cdot) = \sum_{j=1}^{+\infty} R_j(\xi_0, \tau, \cdot)$$

In the next section we will see these calculations in a more formal way, but for the purpose of this section is sufficient to recall the definitions of the functions used in the parametrix method. Therefore we notice that, also in this case, we need to prove some regularity properties of the functions $R_j(\xi_0, \tau, \cdot)$. In order to prove the continuity of the function $R_j(\xi_0, \tau, \cdot)$ we need to prove the following theorems:

Lemma 5.4.1. *There exist constants C_0, C_1 such that the following estimates hold:*

- $|(X_i - X_{i,\xi_0})\Gamma_{\xi_0,\tau}(\xi, t)| \leq C_0 E(\xi, t)$;
- $|X_i \Gamma_{\xi_0,\tau}(\xi, t)| \leq \frac{C_1}{\sqrt{t}} E(\xi, t)$

for $i = 1, 2$

Proof: First of all notice that, from definition 5.3 we obtain:

$$\begin{aligned} \partial_x &= X_{1,\xi_0} cos\theta_0 - X_{3,\xi_0}((\theta - \theta_0)cos\theta_0 - sin\theta_0) \\ \partial_y &= X_{3,\xi_0}(cos\theta_0 - (\theta - \theta_0)sin\theta_0) + X_{1,\xi_0} sin\theta_0 \end{aligned} \tag{5.24}$$

For $i = 2$ the first statement is trivial. Consider $i = 1$:

$$(X_1 - X_{1,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t) = (cos\theta - cos\theta_0 + (\theta-\theta_0)sin\theta_0)\partial_x\Gamma_{\xi_0,\tau}(\xi,t)+$$
$$(sin\theta - sin\theta_0 - (\theta-\theta_0)cos\theta_0)\partial_y\Gamma_{\xi_0,\tau}(\xi,t) =$$
$$(cos\theta - T_{1,\theta_0}cos\theta)(X_{1,\xi_0}cos\theta_0 - X_{3,\xi_0}((\theta-\theta_0)cos\theta_0 - sin\theta_0))\Gamma_{\xi_0,\tau}(\xi,t)+$$
$$(sin\theta - T_{1,\theta_0}sin\theta)(X_{3,\xi_0}(cos\theta_0 - (\theta-\theta_0)sin\theta_0) + X_{1,\xi_0}sin\theta_0)\Gamma_{\xi_0,\tau}(\xi,t) \tag{5.25}$$

where $T_{1,\theta_0}f(\theta) = f(\theta_0) + f'(\theta_0)(\theta-\theta_0)$ is the first order truncated Taylor series of f at θ_0. Then we obtain:

$$|(X_1 - X_{1,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t)| \leq C\,|\Gamma_{\xi_0,\tau}(\xi,t)| \leq C_0E(\xi,t)$$

where the last inequality follows from (5.20). The second inequality follows directly from the second inequality of (5.20), infact:

$$\begin{aligned}|X_i\Gamma_{\xi_0,\tau}(\xi,t)| &\leq |(X_i - X_{i,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t)| + |X_{i,\xi_0}\Gamma_{\xi_0,\tau}(\xi,t)| \leq \\ &\leq C_0E(\xi,t) + \frac{C_1}{t}E(\xi,t) \leq \frac{C_1}{\sqrt{t}}E(\xi,t)\end{aligned} \tag{5.26}$$

□

Lemma 5.4.2. *There exists constant C_1, C_2 such that the following estimates hold:*

- $\left|X^2_{i,\xi_0}\Gamma_{\xi_0}(\xi,t)\right| \leq \frac{C_1}{t}E(\xi,t)$
- $|(L - L_{\xi_0})\Gamma_{\xi_0}(\xi,t)| \leq \frac{C_2}{\sqrt{t}}E(\xi,t)$

Proof: The first inequality follows directly from (5.20). For the second statement consider:

$$\begin{aligned}(X_1^2 - X^2_{1,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t) &= (X_1^2 - X_1X_{1,\xi_0} + X_1X_{1,\xi_0} - X^2_{1,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t) = \\ &= X_1(X_1 - X_{1,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t) + (X_1 - X_{1,\xi_0})X_{1,\xi_0}\Gamma_{\xi_0,\tau}(\xi,t) = \\ &= (X_1 - X_{1,\xi_0} + X_{1,\xi_0})(X_1 - X_{1,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t) + (X_1 - X_{1,\xi_0})X_{1,\xi_0}\Gamma_{\xi_0,\tau}(\xi,t) = \\ &= (X_1 - X_{1,\xi_0})^2\Gamma_{\xi_0,\tau}(\xi,t) + X_{1,\xi_0}\Gamma_{\xi_0,\tau}(\xi,t) + (X_1 - X_{1,\xi_0})X_{1,\xi_0}\Gamma_{\xi_0,\tau}(\xi,t)\end{aligned} \tag{5.27}$$

Then we obtain:

$$|(L - L_{\xi_0})\Gamma_{\xi_0,\tau}(\xi,t)| = \left|(X_1^2 - X^2_{1,\xi_0})\Gamma_{\xi_0,\tau}(\xi,t)\right| \leq C_0E(\xi,t) + \frac{C_1}{\sqrt{t}}E(\xi,t) \leq \frac{C_1}{\sqrt{t}}E(\xi,t) \tag{5.28}$$

□

Theorem 5.4.3. *For every $j \in \mathbb{N}$ the following estimates hold:*

$$|R_j(\xi_0,\tau;x,t)| \leq c_1^j b_j(1)(t-\tau)^{-1+j/2}\Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x, c_2(t-\tau)), \tag{5.29}$$

where $b_j(1) = \gamma^j(\frac{1}{2})/\gamma(\frac{j}{2})$(here γ denotes the Euler gamma function).

Proof: We prove the thesis by induction. For $j = 1$ we have:

$$\begin{aligned}|R_1(\xi_0, \tau; x, t)| &= \left|(L - L_{\xi_0})\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ x, t - \tau)\right| \le \\ &\le \frac{C_1^1}{\sqrt{t-\tau}}\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ x, t - \tau)\end{aligned} \tag{5.30}$$

and we can conclude because $b_1(1) = \gamma(1/2)/\gamma(1/2) = 1$.
Suppose now the thesis hold for j and consider:

$$\begin{aligned}&|R_{j+1}(\xi_0, \tau; x, t)| = |R_1 * R_j(\xi_0, \tau; x, t)| = \left|\int_{\mathbb{R}^N\times[\tau,t]} R_1(x, t; \eta, s)R_j(\eta, s, \xi_0, \tau)d\eta ds\right| \le \\ &\le \int_{\mathbb{R}^N\times[\tau,t]} |R_1(x, t; \eta, s)R_j(\eta, s, \xi_0, \tau)|\, d\eta ds \le \\ &\le \int_{\mathbb{R}^N\times[\tau,t]} \frac{C_1^1}{\sqrt{t-\tau}}\Gamma_{\eta,s}(\eta^{-1} \circ x, t - s)C_1^j(s-\tau)^{-1+j/2}b_j(1)\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ \eta, s - \tau)d\eta ds = \\ &= C_1^{j+1}b_j(1)\int_{\mathbb{R}^N} \Gamma_{\eta,s}(\eta^{-1} \circ x, t - s)\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ \eta, s - \tau)d\eta \times \int_\tau^t (t-s)^{-1/2}(s-\tau)^{-1+j/2}ds = \\ &= C_1^{j+1}b_j(1)\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ x, t - \tau)\int_\tau^t (t-s)^{-1/2}(s-\tau)^{-1+j/2}ds\end{aligned} \tag{5.31}$$

where the last equality for $\Gamma_{\xi_0,\tau}$ follows from (5.15). If we apply the following change of variables:

$$\rho = \frac{s-\tau}{t-\tau}$$

and consider $a = 1/2$, $b = j/2$, a simple computation yields:

$$\int_\tau^t (t-s)^{-1/2}(s-\tau)^{-1+j/2}ds = (t-\tau)^{-1+\frac{j+1}{2}}\int_0^1 \rho^{a-1}\rho^{b-1}d\rho = (t-\tau)^{-1+\frac{j+1}{2}}\frac{\gamma(1/2)\gamma(j/2)}{\gamma((j+1)/2})$$

where the last equality follows from the property of the Gamma Euler function:

$$\int_0^1 \rho^{a-1}\rho^{b-1}d\rho = \frac{\gamma(a)\gamma(b)}{\gamma(a+b)}$$

This ends the proof, infact:

$$\begin{aligned}&|R_{j+1}(\xi_0, \tau; x, t)| \le C_1^{j+1}b_j(1)\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ x, t - \tau)\int_\tau^t (t-s)^{-1/2}(s-\tau)^{-1+j/2}ds = \\ &= C_1^{j+1}\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ x, t - \tau)(t-\tau)^{-1+\frac{j+1}{2}}\frac{\gamma(1/2)\gamma(j/2)}{\gamma((j+1)/2)}b_j(1) = \\ &= C_1^{j+1}\Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ x, t - \tau)(t-\tau)^{-1+\frac{j+1}{2}}b_{j+1}(1)\end{aligned} \tag{5.32}$$

□We will use this result in order to prove the convergence of the series but, first of all, we want to prove that the functions $R_j(\xi_0, \tau; \cdot)$ are continuous.

Theorem 5.4.4. *For every $x, x' \in \mathbb{R}^N$ and $t > \tau$ we have that*

$$R_j(\xi_0, \tau; x, t)$$

is continous for every $j \geq 1$.

Proof: Let $x, x' \in \mathbb{R}^N$ and consider:

$$\begin{aligned}
&\left|R_1(\xi_0,\tau;x,t) - R_1(\xi_0,\tau;x',t)\right| = \left|(L - L_{\xi_o,\tau})(\Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x, t-\tau) - \Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x', t-\tau))\right| = \\
&= \left|(X_1^2 - X_{1,\xi_0}^2)(\Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x, t-\tau) - \Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x', t-\tau))\right| \leq \\
&\leq \frac{C_1}{\sqrt{t-\tau}}(E(\xi_0^{-1}\circ x, t-\tau) - E(\xi_0^{-1}\circ x', t-\tau))
\end{aligned} \tag{5.33}$$

and the continuity follows from the continuity of the function $E(\xi, t)$.
Consider now $j > 1$ and:

$$\begin{aligned}
&\left|R_{j+1}(\xi_0,\tau;x,t) - R_{j+1}(\xi_0,\tau;x',t)\right| \leq \\
&\int_{\mathbb{R}^N\times[\tau,t]} |R_j(\xi_0,\tau;y,s)| \left|R_1(y,s;x,t) - R_1(y,s;x',t)\right| dyds \leq \\
&\leq c^{j+1} b_j(1) \int_{\mathbb{R}^N\times[\tau,t]} (s-\tau)^{-1+j/2}(t-s)^{-1/2} \Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ y, s-\tau)(\Gamma_{y,s}(y^{-1}\circ x, t-s) + \\
&- \Gamma_{y,s}(y^{-1}\circ x', t-s))dyds = c^{j+1} b_j(1)(\Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x, t-\tau) + \\
&- \Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x', t-\tau)) \int_\tau^t (s-\tau)^{-1+j/2}(t-s)^{-1/2} ds \\
&= C(\Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x, t-\tau) - \Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x', t-\tau))
\end{aligned} \tag{5.34}$$

and the continuity follows from the continuity of the foundamental solution $\Gamma_{\xi_0,\tau}$. Consider now $t > t_0 > \tau$. Analogously we obtain:

$$|R_1(\xi_0,\tau;x,t_0) - R_1(\xi_0,\tau;x,t)| \leq \frac{C_1}{\sqrt{t_0-\tau}}(E(\xi_0^{-1}\circ x, t-\tau) - E(\xi_0^{-1}\circ x, t_0-\tau)) \underset{t_0\to t}{\longrightarrow} 0$$

If we consider now a generic $j > 1$ we have:

$$\begin{aligned}
&|R_{j+1}(\xi_0,\tau;x,t_0) - R_{j+1}(\xi_0,\tau;x,t)| \leq \\
&\leq \int_\tau^{t_0-\delta} |R_j(\xi_0,\tau;y,s)| \, |R_1(y,s;x,t_0) - R_1(y,s;x,t)| \, dyds + \\
&+ \int_{t_0-\delta}^t |R_j(\xi_0,\tau;y,s)| \, |R_1(y,s;x,t)| \, dyds + \\
&+ \int_{t_0-\delta}^t |R_j(\xi_0,\tau;y,s)| \, |R_1(y,s;x,t_0)| \, dyds
\end{aligned} \tag{5.35}$$

It is easy to see that the last two integrals in the right-hand side are small if $|t - t_0| < \delta$ and δ is small enough. On the other hand, for $|t - t_0| < \delta/2$ we get:

$$\begin{aligned} &\int_{\tau}^{t_0-\delta} |R_j(\xi_0, \tau; y, s)| \left|R_1(y, s; x, t) - R_1(y, s; x', t)\right| dyds \leq \\ &\leq c(\delta, j)\, |t - t_0|^{1/2} \int_{\tau}^{t_0} (s - \tau)^{-1+j/2} ds \leq c(\delta, j, t_0, \tau)\, |t - t_0|^{1/2} \end{aligned} \tag{5.36}$$

In this way the continuity at t_0 of $R_j(\xi_0, \tau; x, \cdot)$ is proved. □

We are now ready to illustrate the parametrix method and prove that the function defined in (5.23) is the foundamental solution of the heat operator (5.19).

5.5 Construction of the parametrix

In this last section we will prove that the series resulting from the iterative method of the parametrix is convergent.Moreover the results obtained in the previous section will permit us to prove that Γ defined as limit of K_j is the foundamental solution of the heat operator L.

We recall for completeness the calculations made in the previous chapter. Consider:

$$\begin{cases} K_1(\cdot) = \Gamma_{\xi_0,\tau}(\cdot) \\ R_1(\xi_0, \tau; \cdot) = (L - L_{\xi_0,\tau})\Gamma_{\xi_0,\tau}(\cdot) \end{cases} \tag{5.37}$$

We obtain:

$$LK_1(\cdot) = L\Gamma_{\xi_0,\tau}(\cdot) = (L - L_{\xi_0,\tau})\Gamma_{\xi_0,\tau}(\cdot) + L_{\xi_0,\tau}\Gamma_{\xi,\tau}(\cdot) = R_1(\xi_0, \tau; \cdot) + \delta(\cdot) \tag{5.38}$$

Now we define:

$$\begin{cases} K_2(\cdot) = K_1(\cdot) - \Gamma_{\xi_0,\tau} * R_1(\xi_0, \tau; \cdot) \\ R_2(\xi, \tau; \cdot) = -R_1 * R_1(\xi, \tau; \cdot) \end{cases} \tag{5.39}$$

If we now consider LK_2, we have:

$$\begin{aligned} LK_2(\cdot) &= LK_1(\cdot) - L\Gamma_{\xi_0,\tau} * R_1(\xi_0, \tau; \cdot) = \\ &= R_1(\xi_0, \tau; \cdot) + \delta - (L - L_{\xi_0,\tau})\Gamma_{\xi_0,\tau} * R_1(\xi_0, \tau; \cdot) - L_{\xi_0,\tau}\Gamma_{\xi_0,\tau} * R_1(\xi_0, \tau; \cdot) = \\ &= R_1(\xi_0, \tau; \cdot) + \delta(\cdot) - R_1 * R_1(\xi_0, \tau; \cdot) - R_1(\xi_0, \tau; \cdot) = \delta(\cdot) + R_2(\xi_0, \tau; \cdot) \end{aligned} \tag{5.40}$$

Iterating, we have:

$$\begin{cases} K_j(\cdot) = K_{j-1}(\cdot) - \Gamma_{\xi_0,\tau} * R_{j-1}(\xi_0, \tau; \cdot) \\ R_j(\xi_0, \tau; \cdot) = -R_1 * R_{j-1}(\xi, \tau; \cdot) \end{cases} \tag{5.41}$$

and:

$$LK_j(\cdot) = \delta(\cdot) + R_j(\xi_0, \tau; \cdot) \tag{5.42}$$

From equation (5.42) we can clearly understand that, if we consider the limit for $j \to +\infty$ and we demonstrate that $\lim_{j\to+\infty} R_j(\xi, \tau; \cdot) = 0$, then $\lim_{j\to+\infty} K_j$ is the foundamental solution of the parabolic operator L.

We rewrite K_j in the following way:

$$\begin{aligned} K_j(\cdot) &= K_{j-1}(\cdot) - \Gamma_{\xi_0,\tau} * R_{j-1}(\xi_0, \tau; \cdot) = \\ &= K_{j-2}(\cdot) - \Gamma_{\xi_0,\tau} * R_{j-2}(\xi_0, \tau; \cdot) - \Gamma_{\xi_0,\tau} * R_{j-1}(\xi_0, \tau; \cdot) = \\ &= K_1(\cdot) - \Gamma_{\xi_0,\tau} * R_1(\xi_0, \tau; \cdot) - \Gamma_{\xi_0,\tau} * R_2(\xi_0, \tau; \cdot) ... - \Gamma_{\xi_0,\tau} * R_{j-1}(\xi_0, \tau; \cdot) = \\ &= \Gamma_{\xi_0,\tau}(\cdot) - \Gamma_{\xi_0,\tau} * \phi(\cdot) \end{aligned} \tag{5.43}$$

with:

$$\phi(\cdot) = \sum_{k=1}^{j-1} R_k(\xi_0, \tau; \cdot) \tag{5.44}$$

$$\begin{cases} R_1(\xi_0, \tau; \cdot) = & (L - L_{\xi_0,\tau})\Gamma_{\xi_0,\tau}(\cdot) \\ R_k(\xi_0, \tau; \cdot) = & -R_1 * R_{k-1}(\xi_0, \tau; \cdot) \quad for\ k \geq 2 \end{cases} \tag{5.45}$$

It's clear now that the foundamental solution Γ of L will be:

$$\lim_{j\to+\infty} K_j(\cdot) = \Gamma_{\xi_0,\tau}(\cdot) - \Gamma_{\xi_0,\tau} * \sum_{j=1}^{+\infty} R_j(\xi, \tau; \cdot)$$

Then now the problem is studying the convergence of the series $\phi(\cdot) = \sum_{j=1}^{+\infty} R_j(\xi, \tau; \cdot)$.

Theorem 5.5.1. *The series*

$$\phi(x, t) = \sum_{j=1}^{+\infty} R_j(\xi, \tau; x, t) \tag{5.46}$$

totally converges on the set $\{0 < t-\tau \leq T, d(x, \xi) + |t - \tau|^{\frac{1}{2}} \geq \frac{1}{T}\}$ *(for every* $T > 0$*), and satisfies the estimate*

$$|\phi(x, t)| \leq c(T)(t - \tau)^{-1/2} E(\xi_0^{-1} \circ x, c(t - \tau)), \quad 0 < t - \tau \leq T. \tag{5.47}$$

Proof: The proof follows directly from theorem (5.4.3), observing that the power series $\sum_{j=1}^{+\infty} b_j(1)\omega^j$ has infinite radius of convergence. □

Theorem 5.5.2. *We have:*

$$\phi(x, t) = R_1(\xi, \tau; x, t) + \int_{\mathbb{R}^N\times[\tau,t]} R_1(\xi, \tau; x, t)\phi(x, t)dxdt \tag{5.48}$$

for every $(x, t), (\xi_0, \tau) \in \mathbb{R}^{N+1}$, $t > \tau$.

Proof: From the theorem above, it follows that the series:

$$\sum_{j=1}^{+\infty} \int_{\mathbb{R}^N \times [\tau,t]} |R_1(\xi_0,\tau,x,t)R_j(\xi_0,\tau,x,t)|\, dxdt$$

is convergent.Hence:

$$\int_{\mathbb{R}^N \times [\tau,t]} R_1(\xi_0,\tau;x,t)\phi(x,t)dxdt = \sum_{j=1}^{+\infty} \int_{\mathbb{R}^N \times [\tau,t]} R_1(\xi_0,\tau;x,t)R_j(\xi_0,\tau;x,t)dxdt =$$
$$= \sum_{j=1}^{+\infty} R_{j+1}(\xi_0,\tau;x,t) = \phi(x,t) - R_1(\xi_0,\tau;x,t) \tag{5.49}$$

This ends the proof. □

Theorem 5.5.3. *Let $T > 0$. We have*

$$\left|\phi(x,t) - \phi(x',t)\right| \le c(T)d(x,x')^{1/2}(t-\tau)^{-3/4}(E(\xi^{-1}\circ x, c(t-\tau)) + E(\xi^{-1}\circ x', c(t-\tau))) \tag{5.50}$$

for every $x, x', \xi \in \mathbb{R}^N$, $0 < t-\tau \le T$. Moreover,$\phi(\cdot)$ is a continous fucntion in its domain of definition.

Proof:The proof follows directly from theorems (5.4.3) and (5.4.4). □
Consider now the following convolutional operator:

$$J^{\xi_0,\tau}(x,t) = (\Gamma_{\xi_0,\tau} * \phi)(x,t)$$

Then it holds the following:

Theorem 5.5.4. *For every fixed $(\xi_0,\tau) \in \mathbb{R}^{N+1}$, the function $J^{\xi_0,\tau}(\cdot)$ is in $C^2(\{z = (x,t) \in \mathbb{R}^{N+1} : t > \tau\})$, and we have:*

$$X_j(J^{\xi_0,\tau})(x,t) = \int_{\mathbb{R}^{N+1}\times[\tau,t]} X_j\Gamma_{\xi_0,\tau}(y^{-1}\circ x, t-s)\phi(y,s)dyds \tag{5.51}$$

$$X_iX_j(J^{\xi_0,\tau})(x,t) = \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^{N+1}\times[\tau,t-\varepsilon]} X_jX_i\Gamma_{\xi_0,\tau}(y^{-1}\circ x, t-s)\phi(y,s)dyds \tag{5.52}$$

$$\partial_t(J^{\xi_0,\tau})(x,t) = \phi(x,t) + \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^{N+1}\times[\tau,t-\varepsilon]} \partial_t\Gamma_{\xi_0,\tau}(y^{-1}\circ x, t-s)\phi(y,s)dyds \tag{5.53}$$

Proof:First of all we want to show that J is continuous in $\{t > \tau\}$. This can be done by showing that the functions

$$J_\sigma^{\xi_0,\tau}(x,t) = \int_{\mathbb{R}^N\times[\tau+\sigma,t-\sigma]} \Gamma_{\eta,s}(x,t)\phi(\eta,s)d\eta ds$$

are continuous and converge uniformly to J, as $\sigma \to 0^+$, on the compact subsets of $\{t > \tau\}$. The continuity of J_σ follows from the continuity of $\Gamma_{\eta,s}$ and of $\phi(\eta, s)$ by dominated convergence, using the estimates (5.20). On the other hand, for every $K \Subset \{t > \tau\}$,

$$\begin{aligned}
&\sup_K |J_\sigma - J| \le \sup_{(x,t),(\xi_0,\tau)\in K} \left(\int_\tau^{\tau+\sigma} + \int_{t-\sigma}^t\right) \int_{\mathbb{R}^N} |\Gamma_{\eta,s}(x,t)\phi(\eta,s)|\, d\eta ds \\
&\le c(K) \sup_{(x,t),(\xi_0,\tau)\in K} \left(\int_\tau^{\tau+\sigma} + \int_{t-\sigma}^t\right)(s-\tau)^{-1/2} \int_{\mathbb{R}^N} \Gamma_{\eta,s}(\eta^{-1}\circ x, t-s) \\
&\times \Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ \eta, s-\tau) d\eta ds = \\
&= c(K) \sup_{(x,t),(\xi_0,\tau)\in K} \Gamma_{\xi_0,\tau}(\xi_0^{-1}\circ x, t-\tau)\left(\int_\tau^{\tau+\sigma} + \int_{t-\sigma}^t\right)(s-\tau)^{-1/2} ds \le \\
&\le c(K)\sigma^{1/2} \to 0
\end{aligned} \tag{5.54}$$

Here we used the reproduction property and the estimates (5.20) and (5.47). □

In order to prove that the Lie derivatives in (5.51) and (5.52) exist, we shall use the following:

Lemma 5.5.5. *Let $X \in \mathfrak{g}$ and let $\{u_j\}_j$ be a sequence of continuous functions, defined on an open set $A \subseteq \mathbb{R}^N$, with continuous Lie derivative along X. Suppose that u_j converges pointwise in A to some function u and that Xu_j converges to some function ω uniformly on the compact subsets of A. Then there exists the Lie derivative of u along X, $Xu(x) = \omega(x)$, for every $x \in A$.*

Proof:see [7].

Then, let us set:

$$J_\varepsilon^{\xi_0,\tau}(x,t) = \int_{\mathbb{R}^N\times[\tau,t-\varepsilon]} \Gamma_{\eta,s}(x,t)\phi(\eta,s) d\eta ds \tag{5.55}$$

so that $J_\varepsilon^{\xi_0,\tau}$ converges pointwise to $J^{\xi_0,\tau}$, as $\varepsilon \to 0^+$. It is not difficult to see that $J_\varepsilon^{\xi_0,\tau}(\cdot, t)$ has continuous Lie derivatives up to the second order along the vector fields X_1, X_2, obtained deriving (5.55) under the integral sign. In order to prove (5.51) it is then sufficient to show that

$$\sup_{x\in\mathbb{R}^N} \int_{\mathbb{R}^N\times[t-\varepsilon,t]} \left|X_j \Gamma_{\eta,s}(\eta^{-1}\circ x, t-s)\phi(\eta,s)\right| d\eta ds \to 0 \quad as \quad \varepsilon \to 0^+,$$

which is an easy consequence of the estimates (5.47)and (5.20).
In order to prove (5.52), we only have to show that the limit in (5.52) exists,

uniformly in $x \in \mathbb{R}^N$. To this end, let us consider the integral

$$I = \int_{\mathbb{R}^N} X_i X_j \Gamma_{\eta,s}(\eta^{-1} \circ x, t-s) d\eta, \quad \tau < s < t.$$

Using (5.20), (5.47) and the reproduction property , it is easy to see that

$$|I| \le c(T)(t-s)^{-1}(s-\tau)^{-1/2} E(\xi_0^{-1} \circ x, t-\tau).$$

Moreover, for every fixed $y_0 \in \mathbb{R}^N$, we have $I = I_1 + I_2 + I_3$, where

$$\begin{aligned} I_1 &= \int_{\mathbb{R}^N} X_i X_j \Gamma_{\eta,s}(\eta^{-1} \circ x, t-s)(\phi(\eta,s) - \phi(\eta_0,s)) d\eta \\ I_2 &= \phi(\eta_0,s) \int_{\mathbb{R}^N} X_i X_j (\Gamma_{\eta,s} - \Gamma_{\eta_0,s})(\eta^{-1} \circ x, t-s) d\eta \\ I_3 &= \phi(\eta_0 s) \int_{\mathbb{R}^N} X_i X_j \Gamma_{\eta_0,s}(\eta^{-1} \circ x, t-s) d\eta \end{aligned} \tag{5.56}$$

Since $\int_{\mathbb{R}^N} \Gamma_{\eta_0,s}(\eta^{-1} \circ x, t-s) d\eta = 1$, deriving under the integral sign, recalling the estimates (5.20), we obtain $I_3 \equiv 0$. We now chose $\eta_0 = x$ and we estimate I_1, I_2. Making use of the reproduction property, (5.20) and 5.47, we get

$$\begin{aligned} |I_1| &\le C(T)(t-s)^{-1}(s-\tau)^{-3/4} \int_{\mathbb{R}^N} d(x,\eta)^{1/2} E(\eta^{-1} \circ x, t-s) \\ &\times (E(\xi_0^{-1} \circ \eta, s-\tau) + E(\xi_0^{-1} \circ x, s-\tau)) d\eta \\ &\le c(T)((t-s)(s-\tau))^{-3/4} (E(\xi_0^{-1} \circ x, t-s) \int_{\mathbb{R}^N} \Gamma_{\eta,s}(\eta^{-1} \circ x, t-s) d\eta \\ &+ \int_{\mathbb{R}^N} \Gamma_{\xi_0,\tau}(\xi_0^{-1} \circ \eta, s-\tau) \Gamma_{\eta,s}(\eta^{-1} \circ x, t-s) d\eta) \\ &\le c(T)((t-s)(s-\tau))^{-3/4} (E(\xi_0^{-1} \circ x, t-\tau) + E(\xi_0^{-1} \circ x, s-\tau)). \end{aligned} \tag{5.57}$$

Using (5.20), the definition of the function $R_1(\xi_0, \tau; \cdot)$ and (5.47), we obtain:

$$\begin{aligned} |I_2| &\le c(T)(s-\tau)^{-1/2} E(\xi_0^{-1} \circ x, s-\tau) \\ &\times \int_{\mathbb{R}^N} d(x,\eta)(t-s)^{-1} E(\eta^{-1} \circ x, s-\tau) d\eta \\ &\le c(T)(t-s)^{-1/2}(s-\tau)^{-1/2} E(\xi_0^{-1} \circ x, s-\tau) \end{aligned} \tag{5.58}$$

Collecting the above estimates, it is now immediate to recognize that the limit in (5.52) exists, and it is uniform in $x \in \mathbb{R}^N$.

In order to conclude the proof of the statement, it remains only to prove that $J^{\xi_0,\tau}(x,\cdot)$ has continuous derivatives given by (5.53). To this end, it is sufficient to show that $J_\varepsilon^{\xi_0,\tau}(x,\cdot)$ has continuous derivative given by

$$\begin{aligned} \partial_t J_\varepsilon^{\xi_0,\tau}(x,t) &= \int_{\mathbb{R}^N} \Gamma_{y,t-\varepsilon}(y^{-1} \circ x, \varepsilon) \phi(y, t-\varepsilon) dy \\ &+ \int_{\mathbb{R}^N \times [\tau, t-\varepsilon]} \partial_t \Gamma_{y,s}(y^{-1} \circ x, t-s) \phi(y,s) dy ds \end{aligned} \tag{5.59}$$

and that

$$\sup_{t\in K}\left|\phi(x,t)-\int_{\mathbb{R}^N}\Gamma_{y,t-\varepsilon}(y^{-1}\circ x,t-\varepsilon)dy\right|\to 0 \tag{5.60}$$

$$\sup_{t\in K}\int_{t-\varepsilon}^{t}\left|\int_{\mathbb{R}^N}\partial_t\Gamma_{y,s}(y^{-1}\circ x,t-s)\phi(y,s)dy\right|ds\to 0 \tag{5.61}$$

as $\varepsilon\to 0^+$, for every $K\Subset(\tau,+\infty)$.We have

$$\begin{aligned}&\frac{1}{h}(J_\varepsilon^{\xi_0,\tau}(x,t+h)-J_\varepsilon^{\xi_0,\tau}(x,t))\\&=\int_{t-\varepsilon}^{t-\varepsilon+h}\frac{1}{h}\int_{\mathbb{R}^N}\Gamma_{y,s}(y^{-1}\circ x,t+h-s)\phi(y,s)dyds\\&+\int_{\mathbb{R}^N\times[\tau,t-\varepsilon]}\frac{1}{h}(\Gamma_{y,s}(y^{-1}\circ x,t+h-s)-\Gamma_{y,s}(y^{-1}\circ x,t-s))\phi(y,s)dyds\end{aligned} \tag{5.62}$$

The second integral in (5.62) converges (as $h\to 0$) to the second integral in (5.59), by dominated convergence. The first integral in (5.62) is equal to

$$\int_0^1\int_{\mathbb{R}^N}\Gamma_{y,t-\varepsilon+rh}(y^{-1}\circ x,\varepsilon+h-rh)\phi(y,t-\varepsilon+rh)dydr$$

which converges to the first integral in (5.59) by dominated convergence, by means of (5.20), (5.47) and Theorem (5.5.3).This proves (5.59). Using the properties just recalled, it is also easy to see that $\partial_t J_\varepsilon^{\xi_0,\tau}(x,\cdot)$ is continuous, again by dominated convergence. We now prove (5.60) and (5.61). The supremum in (5.60) is lower than $S_1+S_2+S_3$, where

$$\begin{aligned}S_1&=\sup_{t\in K}\int_{\mathbb{R}^N}\left|(\Gamma_{y,t-\varepsilon}(y^{-1}\circ x,\varepsilon)-\Gamma_{x,t-\varepsilon}(y^{-1}\circ x,\varepsilon))\phi(y,t-\varepsilon)\right|dy\\S_2&=\sup_{t\in K}\int_{\mathbb{R}^N}\Gamma_{x,t-\varepsilon}(y^{-1}\circ x,\varepsilon)\left|\phi(y,t-\varepsilon)-\phi(x,t-\varepsilon)\right|dy\\S_3&=\sup_{t\in K}\left|\phi(x,t-\varepsilon)-\phi(x,t)\right|\end{aligned} \tag{5.63}$$

From the continuity of $\phi(x,\cdot)$, we infer that $S_3\to 0$ as $\varepsilon\to 0^+$. Using (5.20) and (5.47), we obtain

$$\begin{aligned}S_1&\le c(K,\tau)(\int_{d(x,y)\le\delta}d(x,y)E(y^{-1}\circ x,\varepsilon)dy+2\int_{d(x.y)>\delta}E(y^{-1}\circ x,\varepsilon)dy)\\&\le c(K,\tau)(\delta\int_{\mathbb{R}^N}\Gamma_{y,s}(y^{-1}\circ x,\varepsilon)dy+\int_{d(0,y')>\frac{\delta}{\sqrt{\varepsilon}}}exp(-cd(o,y')^2)dy')\\&\le c(K,\tau)(\delta+\int_{d(0,y')>\frac{\delta}{\sqrt{\varepsilon}}}exp(-cd(o,y')^2)dy')\end{aligned} \tag{5.64}$$

Hence $S_1 \to 0$ as $\varepsilon \to 0^+$. In a similar manner one can see that aso S_2 vanishes as ε goes to zero. This proves (5.60). The proof of (5.61) closely follows the lines of the proof of (5.60), and therefore it is omitted.This completes the proof. □

We are finally ready to show that the function Γ is the foundamental solution of the operator L.

Theorem 5.5.6. *For every fixed* $(\xi_0, \tau) \in \mathbb{R}^{N+1}$, *we have*

$$\Gamma(\cdot) \in C^2(\mathbb{R}^{N+1}/\{(\xi_0,\tau)\}), \quad L(\Gamma(\cdot)) = \delta(\cdot) \ \ in \ \ \mathbb{R}^{N+1}/\{(\xi_0,\tau)\} \tag{5.65}$$

Proof: Recalling the construction of Γ, from the theorems above, we only have to prove that $L(\Gamma(\cdot)) = \delta(\cdot)$. We have, making use of (5.52) and theorem (5.5.2):

$$\begin{aligned} &L(\Gamma(x,t)) = \\ &= L(\Gamma_{\xi_0,\tau}(x,t)) - \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^{N+1}\times[\tau,t-\varepsilon]} L(\Gamma_{y,s}(y^{-1} \circ x, t-s))\phi(y,s)dyds = \\ &= R_1(\xi_0,\tau;x,t) + \delta(x,t) + \int_{\mathbb{R}^N\times[\tau,t]} R_1(y,s;x,t)\phi(y,s)dyds - \phi(x,t) = \\ &= \delta(x,t) \end{aligned} \tag{5.66}$$

□Now that we have analyzed the parametrix method in our case of interest $\mathbb{R}^2 \times S^1$ we can go back to the main aim of this thesis: find the approximations for the $\alpha-$diffusion kernel and dilation/erosion kernel defined in chapter (3).

Chapter 6

Kernel approximation

In this chapter we will analyze the kernels found in chapter (2). In particular we want to show that the kernel of the fractional diffusion step can be found applying the parametrix method in case $\alpha = 1$. For $\alpha \in (0, 1)$ we will use a key relation between the $\alpha-$kernel and the diffusion kernel and we will show an approximation of the $\alpha-$kernel.
The dilation and erosion kernel will be found starting from the $\alpha-$kernel because we will show that a particular transform, the Cramér Fourier transform, relates linear convolution with morphological convolution.

The chapter is organized as follows. In the first section we will analyze the fractional diffusion part of the PDE. In particular we will apply the parametrix method in the case $\alpha = 1$ and we will see an approximation in case $\alpha \in]0, 1[$. In the second section, we will analyze the dilation/erosion part of the PDE; in particular we will introduce the notion of Cramér transform and see that the kernel in this case can be seen as the Cramér transform of the fractional diffusion kernel.

6.1 Fractional diffusion

The fractional diffusion step solves the PDE

$$\begin{cases} \frac{\partial W^2}{\partial t} = -(-\Delta_{\mathcal{G}_2})^\alpha W^2(p,t) & for\ p \in G/H, t \geq 0 \\ W^2(p,0) = U^2(p) & for\ p \in G/H \end{cases} \tag{6.1}$$

As with fractional diffusion on $\mathbb{R}^n$ there exists a smooth function:

$$K_\cdot^\alpha : (0,\infty) \times (G/H) \to [0,\infty)$$

called the foundamental solution of the $\alpha-$diffusion equation, such that for every initial condition U^2, the solution to the PDE (6.1) is given by the convolution of the function U^2 with the foundamental solution K_t^α

$$W^2(p,t) = (K_t^\alpha *_G U^2(\cdot,t))(p)$$

Our approximations will make use of the following norm on the Lie algebra of the group G that is induced by a metric tensor field on a homogeneous space.

Definition 6.1.1. Let $\mathcal{G}$ be a left-invariant metric tensor field on the homogeneous space G/H with reference element p_0 then

$$\forall v \in T_eG : \| v \|_{\mathcal{G}} := \| \frac{\partial}{\partial t} exp_G(tv).p_0 \, |_{t=0} \|_{\mathcal{G}|_{p_0}} \tag{6.2}$$

is its norm on the Lie algebra $T_e(G)$ where $exp_G : T_e(G) \to G$ is the exponential map of G.

Note that due to the left invariance of $\mathcal{G}$ we can pick any point $p \in G/H$ instead of p_0 in the definition above and get the same result.

For small enough distances from p_0 this norm then approximates the metric $d_{\mathcal{G}}$ of the homogeneous space as

$$d_{\mathcal{G}}(p_0, p) \approx \inf_{g \in G_p} \| log_G g \|_{\mathcal{G}}, \tag{6.3}$$

where log_G is the logarithmic map of G. The distance between p_0 and p is found as follows: we select all exponential curves in G whose actions on the homogeneous space connect p_0 to p. Then, from this set we choose the exponential curve that has the lowest velocity according to the norm in Def. 6.1.1 and use its velocity as the distance estimate.

Definition 6.1.2. Let $\mathcal{G}$ be a left-invariant metric tensor field on the homogeneous space G/H then we define the logaritmic metric estimate as

$$\rho_{\mathcal{G}}(p) := \inf_{g \in G_p} \| log_G g \|_{\mathcal{G}} \tag{6.4}$$

In our case of interest $\mathbb{M}_2$ the metric estimate simplifies to

$$\rho_{\mathcal{G}}(g) = \| log_G g \|_{\mathcal{G}|_e} \tag{6.5}$$

Infact, using the coordinates (x, y, θ) for $\mathbb{M}_2$ and a left invariant metric tensor field of the form (3.17) we formulate the metric estimate in terms of the following auxiliary functions called the exponential coordinates of the first kind:

$$\begin{aligned} c^1(x, y, \theta) &:= \begin{cases} \frac{\theta}{2}(y + x cot\frac{\theta}{2}) & if \ \theta \neq 0, \\ x & \theta = 0, \end{cases} \\ c^3(x, y, \theta) &:= \begin{cases} \frac{\theta}{2}(-x + y cot\frac{\theta}{2}) & if \ \theta \neq 0, \\ y & \theta = 0, \end{cases} \\ c^2(x, y, \theta) &:= \theta. \end{aligned} \tag{6.6}$$

Then the logarithmic metric estimate for $SE(2)$ is given by

$$\rho_{\mathcal{G}}(x,y,\theta) = \sqrt{D_M c^1(x,y,\theta)^2 + D_L c^3(x,y,\theta)^2 + D_A c^2(x,y,\theta)^2}, \qquad (6.7)$$

where the expression (6.7) is a consequence of theorem (3.1.14). We can see that the metric estimate $\rho_{\mathcal{G}}$ has the necessary compatibility property to be a kernel used in convolutions per Def (3.4.6).

Lemma 6.1.3. *Let $\mathcal{G}$ be a left-invariant metric tensor field on the homogeneous space G/H with reference element p_0 and consider $\rho_{\mathcal{G}}$ the logaritmic metric estimate. Then*

$$\forall h \in H : \rho_{\mathcal{G}}(h.p) = \rho_{\mathcal{G}}(p). \qquad (6.8)$$

Proof: We apply Def 6.1.2 and find

$$\rho_{\mathcal{G}}(p) = \inf_{g \in G_p} \| \frac{\partial}{\partial t} exp_G(t\ log_G g).p_0 \ |_{t=0} \|_{\mathcal{G}|_{p_0}}$$

Due to the left-invariance of $\mathcal{G}$ and the fact that $h.p_0 = p_0$ the following equality holds for all $h \in H$:

$$= \inf_{g \in G_p} \| (L_h)_* \frac{\partial}{\partial t}(t\ log_G g).p_0 \ |_{t=0} \|_{\mathcal{G}|_{p_0}}$$

This can be rewritten as:

$$= \inf_{g \in G_p} \| \frac{\partial}{\partial t} h\ exp_G(t\ log_G g).p_0 \ |_{t=0} \|_{\mathcal{G}|_{p_0}}$$

Now we see that we are optimizing over a set of left-invariant curves whose actions connect p_0 (at $t = 0$) to $h.p$ (at $t = 1$), then we have:

$$= \inf_{g \in G_{h.p}} \| \frac{\partial}{\partial t} exp_G(t\ log_G g).p_0 \ |_{t=0} \|_{\mathcal{G}|_{p_0}} = \rho_{\mathcal{G}}(h.p)$$

This completes the proof. □

We can now make good approximations using the tools that we just developed. First of all consider the case $\alpha = 1$: the diffusion equation in this case becomes

$$\begin{cases} \frac{\partial W^2}{\partial t} = \Delta_{\mathcal{G}_2} W^2(p,t) & for\ p \in G/H, t \geq 0 \\ W^2(p,0) = U^2(p) & for\ p \in G/H \end{cases} \qquad (6.9)$$

In our particular case we have $G/H = \mathbb{M}_2 = \mathbb{R}^2 \times S^1$ which is the space we had considered in the previous chapter.Recall that in this case we have:

$$\Delta_{\mathcal{G}} = X_1^2 + X_2^2$$

We can then apply the parametrix method seen before.

If we indicate with $K^1_{\xi_0}$ the foundamental solution of the approximate subelliptic operator:

$$L_{\xi_0} = X^2_{\xi_0,1} + X^2_{\xi_0,2} - \partial t = \Delta^{\xi_0}_{\mathcal{G}_2} - \partial_t$$

then it follows from the parametrix method that the foundamental solution of (6.9) is:

$$K^1_t(p,t) = K^1_{\xi_0}(p,t) - K^1_{\xi_0} * \phi(p,t)$$

where

$$\phi = \sum_{j=1}^{+\infty} R_j(\xi_0;\cdot)$$

and $R_j(\xi_0;\cdot)$ is defined in (5.45). Then the kernel can be approximated by the foundamental solution of an operator frozen at ξ_0 minus the convolution between this foundamental solution and a convergent series. The parametrix method in this case is really helpful because it finds out a local approximation for each point of the group.

We want now to extend the result found for $\alpha = 1$ for a generic value of $\alpha \in (0,1)$. From semi-group theory [9] it follows that semi-groups generated by taking fractional powers of the generator (in our case $\Delta_{\mathcal{G}} \mapsto -|\Delta_{\mathcal{G}}|^{\alpha}$) amount to the following key relation between the $\alpha-$kernel and the diffusion kernel:

$$K^{\alpha}_t(p) := \int_0^{+\infty} q_{t,\alpha}(\tau) K^1_{\tau}(p) d\tau, \quad for\ \alpha \in (0,1], t > 0 \tag{6.10}$$

where $q_{t,\alpha} \geq 0$ is the inverse Laplace transform of $r \mapsto e^{-tr^{\alpha}}$[9]. It follows that:

$$\begin{aligned} K^{\alpha}_t(p) &:= \int_0^{+\infty} q_{t,\alpha}(\tau) K^1_{\tau}(p) d\tau = \\ &= \int_0^{+\infty} q_{t,\alpha}(\tau)(K^1_{\xi_0}(p,\tau) - K^1_{\xi_0} * \phi(p,\tau))d\tau = \\ &= \int_0^{+\infty} q_{t,\alpha}(\tau) K^1_{\xi_0}(p,\tau) d\tau - \int_0^{+\infty} q_{t,\alpha}(\tau) K^1_{\xi_0} * \phi(p,\tau))d\tau = \\ &= K^{\alpha}_{\xi_0}(p,t) - K^{\alpha}_{\xi_0} * \phi(p,t) \end{aligned} \tag{6.11}$$

where $K^{\alpha}_{\xi_0}$ is the foundamental solution of the operator:

$$L^{\alpha}_{\xi_0} = -(-\Delta^{\xi_0}_{\mathcal{G}_2})^{\alpha} - \partial t$$

frozen at ξ_0.

It follows that the foundamental solution of the fractional diffusion equation for $\alpha \in (0,1]$ can be approximated as:

$$K^{\alpha}_t(p) = K^{\alpha}_{\xi_0}(p,t) - K^{\alpha}_{\xi_0} * \phi(p,t)$$

where ϕ is a convergent series.

In the next section we will see the final step of the PDE unit: the dilation/erosion substep. In particular we will see how we can find an approximation of the dilation/erosion kernel starting from the approximate diffusion kernel that we just found.

6.2 Dilation/Erosion

The dilation/erosion step solves the PDE

$$\begin{cases} \frac{\partial W^3}{\partial t} = \pm \parallel \nabla_{\mathcal{G}_3} W^3(p,t) \parallel_{\mathcal{G}_3}^{2\alpha} & for\ p \in G/H, t \geq 0 \\ W^3(p,0) = U^3(p) & for\ p \in G/H \end{cases} \tag{6.12}$$

By a generalization of the Hopf-Lax formula [21], the solution is given by morphological convolution defined in Def. 3.5.2:

$$W^3(p,t) = -(k_t^\alpha \square_G - U^3)(p) \tag{6.13}$$

for the +(dilation) variant and

$$W^3(p,t) = (k_t^\alpha \square_G U^3)(p) \tag{6.14}$$

for the -(erosion) variant, where the kernel k_t^α, also called the structuring element in the context of morphology, is a proper (then not everywhere equal to ∞) lower semi-continous function of the type

$$k_\cdot^\alpha : (0,\infty) \times (G/H) \to \mathbb{R} \cup \{\infty\} \tag{6.15}$$

The morphological convolution is the one that has been defined in (3.50). As with fractional diffusion we do not have a general analytic expression for the foundamental solution to the dilation/erosion problem but we can make the following analytic estimates.

Theorem 6.2.1. *The morphological convolution kernel k_t^α is for small times t and $\alpha \in (1/2, 1]$ well-approximated by*

$$k_t^\alpha(p) \approx k_t^{\alpha,approx}(p) = \left(\frac{2\alpha - 1}{(2\alpha)^{2\alpha/(2\alpha-1)}}\right) t^{-\frac{1}{2\alpha-1}} \rho_{\mathcal{G}_3}(p)^{\frac{2\alpha}{2\alpha-1}}, \tag{6.16}$$

and for $\alpha = 1/2$ by

$$k_t^{1/2}(p) \approx k_t^{1/2,approx}(p) = \begin{cases} 0 & if\ \rho_{\mathcal{G}_3}(p) \leq t \\ \infty & if\ \rho_{\mathcal{G}_3}(p) > t \end{cases} \tag{6.17}$$

where $\rho_{\mathcal{G}_3}$ is the estimate of the Riemannian distance between p and p_0 induced by the left-invariant metric tensor field $\mathcal{G}_3$ given by definition (6.4).

We will prove this result by using a transformation that is able to relate the dilation/erosion PDE with the fractional diffusion PDE. Then we will first introduce this so called Cramér Fourier transform in $\mathbb{R}^d$ but since we do not know a generalization to the group we introduce an approximate Cramér Fourier transform via the Lie algebra. It is this approximate transform that we will apply to the approximate diffusion kernel to get an approximate morphological kernel. Let us first recall the definition of the Fenchel transform

Definition 6.2.2. Let $c \in C_x$, where C_x denotes the set of mappings from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ proper and lower semi continous.Its Fenchel transform is the function from $\mathbb{R}^d$ into $\bar{\mathbb{R}}$ defined by

$$\hat{c}(\theta) \stackrel{def}{=} [\mathfrak{F}(c)](\theta) \stackrel{def}{=} \sup_{x \in \mathbb{R}^d} [\theta \cdot x - c(x)] \tag{6.18}$$

Example 6.2.3. *The Fenchel transform of $\mathcal{M}^p_{m,\sigma}$ defined in example (3.5.3) is*

$$[\mathfrak{F}(\mathcal{M}^p_{m,\sigma})](\theta) = \frac{1}{p'} |\theta\sigma|^{p'} + m\theta,$$

with $1/p + 1/p' = 1$. The particular case $p = 2$ corresponds to the characteristic function of a Gaussian law.

On $\mathbb{R}^d$ the Cramér-Fourier transform is given as follows

Definition 6.2.4. For functions on $\mathbb{R}^d$ that have a real-valued non-negative Fourier transform we define the Cramér-Fourier transform as

$$\mathcal{C}_{\mathcal{F}} := \mathfrak{F} \circ -log \circ \mathcal{F}, \tag{6.19}$$

where $\mathcal{F}$ denotes the Fourier transform, *log* denotes the point-wise logarithm and $\mathfrak{F}$ denotes the Fenchel transform.

Before introducing the notion of approximate Cramér Fourier transform we want to see some important properties of the Cramér Fourier transform. In particular the last theorem we will see is the one that allows us to relate linear convolution with morphological convolution. First we have to state a theorem on the properties of the Fenchel transform that will serve us to prove the main result.

Theorem 6.2.5. *For $f, g \in C_x$ we have:*

1. $\mathfrak{F}(f) \in C_x$,
2. $\mathfrak{F}$ *is an involution, that is* $\mathfrak{F}(\mathfrak{F}(f)) = f$,
3. $\mathfrak{F}(f \square_{\mathbb{R}^d} g) = \mathfrak{F}(f) + \mathfrak{F}(g)$
4. $\mathfrak{F}(f + g) = \mathfrak{F}(f) \square_{\mathbb{R}^d} \mathfrak{F}(g)$.

Lemma 6.2.6. *On $\mathbb{R}^d$ the Cramér Fourier transform relates convolution and morphological convolution in the following manner: let f_1 and f_2 be proper, lower semi-continuous have convex Cramér Fourier transform and real-valued non-negative Fourier transforms. Then*

$$\mathcal{C}_{\mathcal{F}}[f_1 *_{\mathbb{R}^d} f_2] = \mathcal{C}_{\mathcal{F}}[f_1] \Box_{\mathbb{R}^d} \mathcal{C}_{\mathcal{F}}[f_2] \tag{6.20}$$

Proof: Since $\mathcal{C}_{\mathcal{F}}[f]$ and $\mathcal{C}_{\mathcal{F}}[g]$ are lower semi continous, proper and convex, also their conjugates $\mathcal{C}_{\mathcal{F}}[f]^*$ and $\mathcal{C}_{\mathcal{F}}[g]^*$ share these properties. Moreover, it follows that $\mathcal{C}_{\mathcal{F}}^* = -log\mathcal{F}[f]$. Therefore, a direct computation gives

$$\begin{aligned} \mathcal{C}_{\mathcal{F}}[f * g] &= (-log\mathcal{F}[f * g](\frac{\cdot}{2\pi}))^* \\ &= ((-log\mathcal{F}[f](\frac{\cdot}{2\pi})) + (-log\mathcal{F}[g](\frac{\cdot}{2\pi})))^* \\ &= (-log\mathcal{F}[f](\frac{\cdot}{2\pi}))^* \Box_{\mathbb{R}^d} (-log\mathcal{F}[g](\frac{\cdot}{2\pi}))^* \\ &= \mathcal{C}_{\mathcal{F}}[f] \Box_{\mathbb{R}^d} \mathcal{C}_{\mathcal{F}}[g]. \end{aligned} \tag{6.21}$$

where we have applied the convolution theorem

$$\mathcal{F}[f * g] = \mathcal{F}[f]\mathcal{F}[g] \tag{6.22}$$

for the Fourier transform and the well-known property

$$(f + g)^* = f^* \Box_{\mathbb{R}^d} g^* \tag{6.23}$$

of the convex conjugate. This completes the proof. □

Theorem 6.2.7. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be differentiable with compact support and has a real-valued non-negative Fourier transform then for $\alpha \in [1/2, 1]$:*

$$\mathcal{C}_{\mathcal{F}}[-(-\Delta)^\alpha f] = - \parallel \nabla(\mathcal{C}_{\mathcal{F}} f) \parallel^{2\alpha} . \tag{6.24}$$

Proof: See [5].

These equalities allow us to relate the fractional diffusion system in (6.1) to the dilation/erosion system in (6.12) and use the approximate solution we have for the first system to construct an approximate solution to the latter system.

Assuming that in fact we are working on $\mathbb{R}^d$ and we have a solution $W^2 : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ to the system (6.1) of the form $W^2(\cdot, t) = K_t^\alpha *_{\mathbb{R}^d} U^2$. Then:

$$\begin{aligned} &\frac{\partial W^2}{\partial t} = -(-\Delta)^\alpha W^2 \\ \Rightarrow &\frac{\partial}{\partial t}(K_t^\alpha * U^2) = -(-\Delta)^\alpha (K_t^\alpha * U^2) \\ \Rightarrow &\mathcal{C}_{\mathcal{F}}[\frac{\partial}{\partial t}(K_t^\alpha * U^2)] = \mathcal{C}_{\mathcal{F}}[-(-\Delta)^\alpha (K_t^\alpha * U^2)] \\ \Rightarrow &\frac{\partial}{\partial t}\mathcal{C}_{\mathcal{F}}[K_t^\alpha * U^2] = - \parallel \nabla \mathcal{C}_{\mathcal{F}}[K_t^\alpha * U^2] \parallel^{2\alpha} \\ \Rightarrow &\frac{\partial}{\partial t}(\mathcal{C}_{\mathcal{F}}[K_t^\alpha] \Box \mathcal{C}_{\mathcal{F}}[U^2]) = - \parallel \nabla(\mathcal{C}_{\mathcal{F}}[K_t^\alpha] \Box \mathcal{C}_{\mathcal{F}}[U^2] \parallel^{2\alpha} \end{aligned} \tag{6.25}$$

Now choose $U^2 = \mathcal{C}_{\mathcal{F}}^{-1}[U^3]$ and let $W^3 = \mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box U^3$ then we see that W^3 solves the erosion version of the system (6.12) by performing a morphological convolution with the initial condition U^3 using the kernel $\mathcal{C}_{\mathcal{F}}[K_t^\alpha]$. The solution to the dilation version of (6.12) now follows from the last expression with a few extra steps:

$$\begin{aligned}
&\Rightarrow \frac{\partial}{\partial t}(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box U^3) = - \parallel \nabla(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box U^3) \parallel^{2\alpha} \\
&\Rightarrow \frac{\partial}{\partial t}(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3) = - \parallel \nabla(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3) \parallel^{2\alpha} \\
&\Rightarrow - \frac{\partial}{\partial t}(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3) = \parallel \nabla(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3) \parallel^{2\alpha} \\
&\Rightarrow - \frac{\partial}{\partial t}(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3) = \parallel -\nabla(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3) \parallel^{2\alpha} \\
&\Rightarrow \frac{\partial}{\partial t}(-(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3)) = \parallel \nabla(-(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3)) \parallel^{2\alpha}
\end{aligned} \tag{6.26}$$

where we see that letting $W^3 = -(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3)$ solves the dilation PDE. We have then proved that the solution to the erosion PDE is given by:

$$W^3 = \mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box U^3$$

while the solution to the dilation PDE is:

$$W^3 = -(\mathcal{C}_{\mathcal{F}}[K_t^\alpha]\Box - U^3)$$

The Cramér-Fourier transform requires a function on $\mathbb{R}^d$, or at least a d-dimensional vector space that we can then naturally identify with $\mathbb{R}^d$. On a homogeneous space we can use the Lie algebra of its group as that vector space and use the exponential and logarithmic maps to translate between the homogeneous space and the group, this leads to an approximate Cramér-Fourier transform

Definition 6.2.8. For functions on G/H that have the property that the following Fourier transform

$$\mathcal{F}[v \mapsto f(exp_G v.p_0)] \tag{6.27}$$

is real-valued and non-negative for all $v \in T_eG$, we define the approximate Cramér-Fourier transform as

$$\mathcal{C}_{\mathcal{F}}^{approx} f(p) := \inf_{g \in G_p} \mathcal{C}_{\mathcal{F}}[v \mapsto f(exp_G v.p_0)](log_G g), \tag{6.28}$$

where we recall G_p from (3.5).

Note that above we naturally identified the n dimensional vector space T_eG, then the Lie algebra of G, with $\mathbb{R}^n$.

Example 6.2.9. *A function f on $\mathbb{M}_2$ expressed in terms of coordinates (x, y, θ) is first transformed to a function f_1 on the Lie algebra expressed in coordinates relative to the basis (2.47) , then the exponential coordinates of the first kind, as:*

$$f_1(c^1, c^2, c^3) := \begin{cases} f(\frac{c^1 sinc^2 - c^3(1-cosc^2)}{c^2}, \frac{c^1(1-cosc^2)+c^3 cosc^2}{c^2}, c^2) & if\ c^2 \neq 0 \\ f(c^1, c^3, 0) & if\ c^2 = 0 \end{cases} \tag{6.29}$$

The function now lives on $\mathbb{R}^3$ and we can apply the Cramér-Fourier transform on it to obtain the function f_2:

$$f_2 := \mathcal{C}_{\mathcal{F}} f_1, \tag{6.30}$$

the function f_2 again lives on $\mathbb{R}^3$ and via the mapping $(c^1, c^2, c^3) \mapsto c^1 X_1 \mid_e + c^2 X_2 \mid_e + c^3 X_3 \mid_e$ on the Lie algebra. We can use the exponential mapping to bring this function back to $\mathbb{M}_2$ again to complete the approximate transform:

$$(\mathcal{C}_{\mathcal{F}}^{approx} f)(x, y, \theta) = \begin{cases} f_2(\frac{\theta}{2}(y + xcot\frac{\theta}{2}), \frac{\theta}{2}(-x + ycot\frac{\theta}{2}), \theta) & if\ \theta \neq 0 \\ f_2(x, y, 0) & if\ \theta = 0 \end{cases} \tag{6.31}$$

Now that we have developed the approximate Cramér-Fourier transform on homogeneous spaces we can obtain an approximation of the morphological kernel k_t^α.

Theorem 6.2.10. *Let $\alpha \in (1/2, 1]$ and let $\mathcal{G}$ be a left-invariant metric tensor field on G/H, then:*

$$k_t^\alpha(p) \approx k_t^{\alpha,approx}(p) = (\mathcal{C}_{\mathcal{F}}^{approx} K_t^{\alpha,approx})(p) = \frac{2\alpha - 1}{(2\alpha)^{2\alpha/(2\alpha-1)}} \frac{\rho_{\mathcal{G}}(p)^{2\alpha/(2\alpha-1)}}{t^{1/(2\alpha-1)}} \tag{6.32}$$

for $\alpha \to 1/2$ this converges to:

$$k_t^{1/2,approx}(p) = \begin{cases} 0 & if\ \rho_{\mathcal{G}}(p) \leq 1 \\ \infty & elsewhere \end{cases} \tag{6.33}$$

These estimates hold for all $p \in G/H$ for sufficiently small $t > 0$.

Proof: From the construction of $K_t^{\alpha,approx}$ we have that

$$\mathcal{F}[v \mapsto K_t^{\alpha,approx}(exp_G v.p_0)] = e^{-t\|\cdot\|_{\mathcal{G}}^{2\alpha}}, \tag{6.34}$$

which is real valued and positive, taking -log yields

$$v \mapsto t \parallel v \parallel_{\mathcal{G}}^{2\alpha} .$$

Applying the Fenchel transform $\mathfrak{F}$ yields

$$\mathfrak{F}[t \parallel \cdot \parallel_{\mathcal{G}}^{2\alpha}](w) = \sup_{v \in T_e G} \{(w, v)_{\mathcal{G}} - t \parallel v \parallel_{\mathcal{G}}^{2\alpha}\}.$$

Obviously to maximize this v must be chosen aligned with w(assume $\neq 0$) and pointing in the same direction, then $v = \lambda w / \parallel w \parallel_{\mathcal{G}}^2$ for some $\lambda > 0$:

$$\begin{aligned} &= \sup_{\lambda > 0} \{(w, \lambda w / \parallel w \parallel_{\mathcal{G}}^2)_{\mathcal{G}} - t \parallel \lambda w / \parallel w \parallel_{\mathcal{G}}^2 \parallel_{\mathcal{G}}^{2\alpha}\} \\ &= \sup_{\lambda > 0} \{\lambda - t \frac{\lambda^{2\alpha}}{\parallel w \parallel_{\mathcal{G}}^{2\alpha}}\}. \end{aligned} \tag{6.35}$$

Now we consider two cases.
Case 1: $\alpha = 1/2$

$$\begin{aligned} \mathfrak{F}[t \parallel \cdot \parallel_{\mathcal{G}}^{2\alpha}](w) &= \sup_{\lambda > 0} \{\lambda - t \frac{\lambda}{\parallel w \parallel_{\mathcal{G}}}\} \\ &= \sup_{\lambda > 0} \{\lambda (1 - \frac{t}{\parallel w \parallel_{\mathcal{G}}})\} \\ &= \begin{cases} 0 & if \ \parallel w \parallel_{\mathcal{G}} \leq t, \\ \infty & if \ \parallel w \parallel_{\mathcal{G}} > t. \end{cases} \end{aligned} \tag{6.36}$$

Next taking the mapping to the homogeneous space we get

$$\begin{aligned} (\mathcal{C}_{\mathcal{F}}^{approx} K_t^{\alpha,approx})(p) &= \inf_{g \in G_p} \mathfrak{F}[t \parallel \cdot \parallel_{\mathcal{G}}^{2\alpha}](log_G g) \\ &= \inf_{g \in G_p} \begin{cases} 0 & if \ \parallel log_G g \parallel_{\mathcal{G}} \leq t, \\ \infty & if \ \parallel log_G g \parallel_{\mathcal{G}} > t, \end{cases} \\ &= \begin{cases} 0 & if \ \inf_{g \in G_p} \parallel log_G g \parallel_{\mathcal{G}} \leq t, \\ \infty & if \ \inf_{g \in G_p} \parallel log_G g \parallel_{\mathcal{G}} > t, \end{cases} \\ &= \begin{cases} 0 & if \ \rho_{\mathcal{G}}(p) \leq t, \\ \infty & if \ \rho_{\mathcal{G}}(p) > t, \end{cases} \end{aligned} \tag{6.37}$$

which proves (6.33).
Case 2: $\alpha \in (1/2, 1]$ Seeing that the objective function is concave we apply the first order test to find the supremum is reached forall

$$\lambda = \frac{\parallel w \parallel_{\mathcal{G}}^{\frac{2\alpha}{2\alpha-1}}}{(2\alpha t)^{\frac{1}{2\alpha-1}}},$$

after substituting and simplifying this yields

$$\mathfrak{F}[t \parallel \cdot \parallel_{\mathcal{G}}^{2\alpha}](w) = \frac{2\alpha - 1}{(2\alpha)^{2\alpha/(2\alpha-1)}} \frac{\rho_{\mathcal{G}}(p)^{2\alpha/(2\alpha-1)}}{t^{1/(2\alpha-1)}}$$

finally taking the mapping to the homogeneous space just like in the first case we confirm 6.32. This completes the proof. $\square$

We end this section proving two theorems that show how, with particular values of α, we can find the ordinary max pooling operation and the ReLu operation. Then morphological convolution will be a generalization of these well known operations.

Theorem 6.2.11. *Let $f \in L^\infty(G/H)$, let $S \subset G/H$ be non empty and define $k_S : G/H \to \mathbb{R} \cup \{\infty\}$ as:*

$$k_S(p) := \begin{cases} 0 & if\ p \in S \\ \infty & else \end{cases} \tag{6.38}$$

Then:

$$-(k_S \square -f)(p) = \sup_{g \in G : g^{-1}.p \in S} f(g.p_0). \tag{6.39}$$

This is the morphological kernel with $\alpha = 1/2$ and we can recognize the morphological convolution as a generalized form of max pooling of the function f with stencil S.
Proof: Filling in (6.38) into definition 3.5.2 yields:

$$\begin{aligned} -(k_S \square - f)(p) &= -inf\{\inf_{g \in G : g^{-1}.p \in S} -f(g.p_0), \inf_{g \in G : g^{-1}.p \notin S} -f(g.p_0) + \infty\} \\ &= -\inf_{g \in G : g^{-1}.p \in S} -f(g.p_0) \\ &= \sup_{g \in G : g^{-1}.p \in S} f(g.p_0). \end{aligned} \tag{6.40}$$

$\square$In the particular case $G = G/H = \mathbb{R}^n$ we recover a more familiar form of max pooling:

Theorem 6.2.12. *Let $G = G/H = \mathbb{R}^n$ and let $f \in C^0(\mathbb{R}^n)$ with $S \subset \mathbb{R}^n$ compact then:*

$$-(k_S \square_{\mathbb{R}^n} - f)(x) = \max_{y \in S} f(x - y) \tag{6.41}$$

ReLU is the other common CNN operation that can be generalized by morphological convolution:

Theorem 6.2.13. *Let f be a compactly supported continuous function on G/H. Then dilation with the kernel*

$$k_{ReLU,f}(p) := \begin{cases} 0 & if\ p = p_0 \\ \sup_{y \in G/H} f(y) & else \end{cases} \tag{6.42}$$

equates to applying a Rectified Linear Unit (ReLu) to the function f:

$$-(k_{ReLU} \square_G - f)(p) = max\{0, f(p)\} \tag{6.43}$$

Proof: Filling in k into the definition of morphological convolution:

$$\begin{aligned} -(k_{ReLU}\square_G - f)(p) &= -\inf_{g\in G} k_{ReLU}(g^{-1}.p) - f(g.p_0) \\ &= -\inf_{g\in G}\{\inf_{g^{-1}.p=p_0} -f(g.p_0), \inf_{g^{-1}.p\neq p_0} -f(g.p_0) + \sup_{y\in G/H} f(y)\} \\ &= \sup_{g\in G}\{f(p), \sup_{z\in G/H: z\neq p} f(z) - \sup_{y\in G/H} f(y)\}, \end{aligned} \tag{6.44}$$

due to the continuity and compact support of f its supremum exists and moreover we have $\sup_{z\in G/H:z\neq p_0} f(z) = \sup_{y\in G/H} f(y)$ and thereby we obtain the required result

$$= max\{f(p), 0\}$$

□

Chapter 7

An application to a group-invariant CNN for image processing

In this chapter we will analyze the PDE-based convolutional neural network defined by Duits,M.N.Smets, Portegies and Bekkers in [1]. In this framework a network layer is seen as a set of PDE-solvers where the equation's geometrically meaningful coefficients become the layer's trainable weights. Precisely, each layer will be expressed by mean of equation (3.1) and the parameters are the left-invariant vector field c and the left-invariant metric tensor fields $\mathcal{G}_1, \mathcal{G}_2$.

The chapter is organized as follows. In the first section we will decsribe the overall architecture of the network, while in the second section we will clarify how to use PDE acting on $\mathbb{R}^2 \times S^1$ to process an image defined on $\mathbb{R}^2$ and we will focus on the lifting layer which transforms the input to the desired homogeneous space and on the projection layer that transforms the output back to the required output space. In the third section we will focus on the design of a PDE layer,in particular we will focus on the design of a PDE unit. Finally, in the last section we will focus on the trainable parameters of the network and on the loss function.

7.1 Overall achitecture

The main scope of the net can be of impainting or denoising, or general image processing. For this reason the network N will be considered as an operator that will take in input a corrupted mage X and return a reconstructed image $N(X)$.The Network operator N is obtained by composing elementary operators, called the layers of the network, which will be represented in terms of the PDE in (3.1).Since this PDE only depends on the unknown coefficients of convection, diffusion and erosion/dilation, finding the best operator N, simply reduces to find these parameters. These parameters are found by minimizing a functional called the loss function.

One of the main peculiarities of this network is that the PDE layers are defined on a homogeneous space. In our case of interest the data lives on $\mathbb{R}^2$ and not on $\mathbb{M}_2$ where we propose to do processing. We then need the addition of lifting and projection layers. We will require the network to be equivariant, which means that it will give the same result to first applying the operator and then a transformation or vice versa. Of course for the entire network to be equivariant we require these transformation layers to be equivariant as well.Hence the network will have the following overall architecture:

- a preprocessing layer, with no unknown
- a PDE layer expressed in terms of left invariant vector fields. Each of them will be expressed through equation (3.1 and the unknown are the coefficients of the equation
- the loss function to be minimized, in order to find the coefficients
- a postporcessing layer, with no unknown

We refer to [10],[11] and [1] for details.

7.2 Preprocessing and post processing layer

The preprocessing operator maps $L^2(\mathbb{R}^2)$ to $L^2(\mathbb{R}^2 \times S^1)$. In order to describe it,we recall that the structure of the group can be mapped to other mathematical objects (such as 2D images) via representations. Group representations describe groups in terms of bijective linear transformations of vector spaces; in particular, they can be used to represent group elements as invertible matrices. In our case of interest the left-regular $SE(2)$ representation on 2D images $f \in L_2(\mathbb{R}^2)$ is given by

$$(\mathcal{U}_g f)(x') = f(R_\theta^{-1}(x' - x)) \tag{7.1}$$

with $g = (x, \theta) \in SE(2)$, $x' \in \mathbb{R}^2$. It corresponds to a roto-translation of the image. Now, on $\mathbb{R}^2$ we define convolution via inner products of translated kernels. Let τ_x be the translation operator, the left-regular representation of the translation group $(\mathbb{R}^2, +)$. Then:

$$(k \bigstar_{\mathbb{R}^2} f)(x) = (\tau_x k, f)_{L_2(\mathbb{R}^2)} := \int_{\mathbb{R}^2} k(x' - x) f(x') dx' \tag{7.2}$$

In the $SE(2)$ lifting layer we now simply replace translations of k by roto-translations via the $SE(2)$ representation $\mathcal{U}_g$ defined in (7.1).

Let $f, k : \mathbb{R}^2 \to \mathbb{R}^{N_c}$ be a vector valued 2D image and kernel (with N_c channels), with $f = (f_1, ..., f_{N_c})$ and $k = (k_1, ..., k_{N_c})$, then the group lifting correlations for vector valued images are defined by:

$$(k\bar{\bigstar}f)(g) := \sum_{c=1}^{N_c} (\mathcal{U}_g k_c, f_c)_{L_2(\mathbb{R}^2)} = \sum_{c=1}^{N_c} \int_{\mathbb{R}^2} k_c(R_\theta^{-1}(y-x)) f_c(y) dy \tag{7.3}$$

These correlations lift 2D image data to data that lives on the 3D position orientation space $\mathbb{R}^2 \times S^1 \equiv SE(2)$.
Projecting from $\mathbb{M}_2$ back down to $\mathbb{R}^2$ can be performed by a simple maximum projection. Then the projection layer projects a multi-channel $SE(2)$ image back to $\mathbb{R}^2$ in the following way: let $f : \mathbb{M}_2 \to \mathbb{R}$ then

$$(x, y) \mapsto \max_{\theta \in [0, 2\pi)} f(x, y, \theta) \tag{7.4}$$

is a roto-translation equivariant projection as used in [11]. In these pre and post processing layers we do not learn parameters.

7.3 Design of a PDE layer

Let us now describe the structure of the generic layer. It will be composed by one PDE formally expressed by equation (3.1),for each feature. In this way it will have 9 unknown parameters for each layer:

- 3 parameters to specify the convection vector field
- 3 parameters to specify the fractional diffusion metric tensor field $\mathcal{G}_1$
- 3 parameters to specify the dilation/erosion metric tensor field $\mathcal{G}_2$.

Each PDE takes as input the output of the previous layer and it works on the input up to a fixed time T. Afterwards, we take all the M_l solutions at time T and take M_{l+1} linear combinations of them. These will be the inputs for the next PDE layer. This design is illustrated in figure (7.1) Each PDE unit will be an N-fold repetition of a timestep-unit which is a composition of convection, diffusion and dilation/erosion substeps, where N is some natural number.

Let's focus on the PDE unit: the default PDE unit that we are considering computes the approximate solution to the PDE

$$\begin{cases} \frac{\partial W}{\partial t}(p,t) = \overbrace{-cW(p,t)}^{Convection} \overbrace{-(-\Delta_{\mathcal{G}_1})^\alpha W(p,t)}^{Fractional\ diffusion} \overbrace{\pm \parallel \nabla_{\mathcal{G}_2} W(p,t) \parallel_{\mathcal{G}_2}^{2\alpha}}^{Dilation/Erosion} & for\ p \in G/H, t \geq 0 \\ W(p,0) = U(p) & for\ p \in G/H \end{cases} \tag{7.5}$$

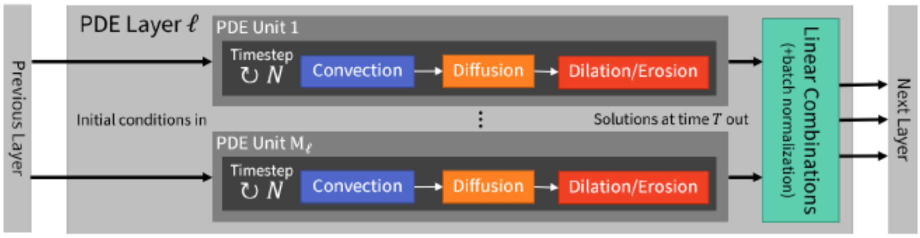

Figure 7.1: Illustrating the architecture of a PDE layer. The M_l PDE units take as inputs the outputs of the previous layer. Then the solutions at a fixed time T are combined in order to form the inputs for the next layer.

Here, c is a left invariant vector field on G/H, $\alpha \in [1/2, 1]$, $\mathcal{G}_1$ and $\mathcal{G}_2$ are left invariant metric tensor fields on G/H, U is the initial condition and $\Delta_{\mathcal{G}}$ and $\| \cdot \|_{\mathcal{G}}$ denote the Laplace operator and norm induced by the metric tensor field $\mathcal{G}$.

A PDE unit will be an N-fold repetition of a timestep-unit which is a composition of convection, diffusion and dilation/erosion substeps, where N is a natural number. These units all take their input as an initial condition of a PDE, and produce as output the solution of a PDE at time $t = T$. The output of a previous timestep-unit is taken as the input for the next timestep-unit. The convection, diffusion and dilation/erosion steps are implemented with respectively a shift, convolution and morphological convolution, as illustrated in the diagram (7.2). The composition of the substeps does not solve

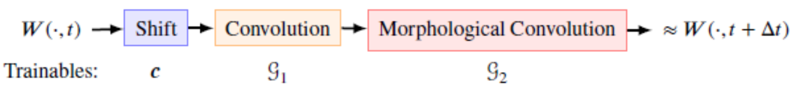

Figure 7.2: Convection, diffusion and dilation/erosion steps

(7.5) exactly, but for small Δt, it approximates the solution by a principle called *operator splitting*. Of course in order to implement the three substeps we need to make explicit approximations of the kernels found in chapter (3).

As we have seen in section (3.3) the solution of the convection PDE does not require the definition of a kernel but is just a translation of the initial condition, then we don't need to do any approximation. The approximate kernels for the diffusion equation and the dilation/erosion equation are the one that have been found in chapter (6) through the parametrix method and whose approximations have been discussed within the thesis.

7.4 Trainable parameters and loss function

Now we want to focus on the trainable parameters of the network: in particular we will analyze our case of interest $\mathbb{M}_2$. Training the PDE unit comes down to adapting the parameters of the PDE, such as the convection vector field c and the metric tensor fields $\mathcal{G}_1, \mathcal{G}_2$, in order to minimize a given loss function. In this sense the vector field and the metric tensor fields are analogous to the weights of this layer. Moreover, since we required the convection vector field and the metric tensor fields to be left-invariant, the parameter space is finite-dimensional as a consequence of theorem (3.1.11) and (3.1.13).

As we've already said $\mathbb{M}_2$ can be identified with $\mathbb{R}^2 \times S^1$. Then the left-invariant vector fields are spanned by the basis (2.47), while the metric tensor fields are of the form (3.17). Therefore each PDE unit would have the following 9 trainable parameters:

- 3 parameters to specify the convection vector field
- 3 parameters to specify the fractional diffusion metric tensor field $\mathcal{G}_1$
- 3 parameters to specify the dilation/erosion metric tensor field $\mathcal{G}_2$.

The choice of the loss function to be minimized depends on the application in which the network is used. For example, in [3] R.Duits,B.Smets,E.Bekkers and J.Portegies tested the viability of PDE-based Group Convolutional Neural Networks (PDE-G-CNNs) for automatic segmentation of vasculature. The loss function used in this case is the continuous DICE loss function, which is implemented as follows:

$$D = \frac{2\sum_{i=1}^{N} p_i g_i}{\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} g_i^2} \tag{7.6}$$

where p_i and g_i represent pairs of corresponding pixel values of prediction and ground truth, respectively.

Bibliography

[1] Smets, B., Portegies, J., Bekkers, E.,Duits, R. (2020). PDE-based Group Equivariant Convolutional Neural Networks. ArXiv, abs/2001.09046.

[2] Stinga, Pablo Raúl. User's guide to the fractional Laplacian and the method of semigroups. Volume 2 Fractional Differential Equations, edited by Anatoly Kochubei and Yuri Luchko, Berlin, Boston: De Gruyter, 2019, pp. 235-266.

[3] Duits, R., Smets, B. M. N., Bekkers, E. J., Portegies, J. W. (2021). Equivariant Deep Learning via Morphological and Linear Scale Space PDEs on the Space of Positions and Orientations. 27-39. Paper presented at 8th International Conference on Scale Space and Variational Methods in Computer Vision.

[4] Akian M., Quadrat JP., Viot M. (1994) Bellman processes. In: Cohen G., Quadrat JP. (eds) 11th International Conference on Analysis and Optimization of Systems Discrete Event Systems. Lecture Notes in Control and Information Sciences, vol 199. Springer, Berlin, Heidelberg.

[5] Schmidt, M., Weickert, J. Morphological Counterparts of Linear Shift-Invariant Scale-Spaces.J Math Imaging Vis 56, 352-366 (2016)

[6] Bucur C., Valdinoci E. (2016) An Introduction to the Fractional Laplacian. In: Nonlocal Diffusion and Applications. Lecture Notes of the Unione Matematica Italiana, vol 20. Springer, Cham.

[7] Bonfiglioli, Andrea, Ermanno Lanconelli, and Francesco Uguzzoni. Fundamental Solutions for Non-Divergence Form Operators on Stratified Groups. Transactions of the American Mathematical Society 356, no. 7 (2004): 2709-737.

[8] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni Uniform Gaussian estimates for the fundamental solutions for heat operators on Carnot groups, Advances in Differential Equations, Adv. Differential Equations 7(10), 1153-1192, (2002)

[9] Yosida, K. : Functional Analysis (1974)

[10] Erik J Bekkers, Maxime W Lafarge, Mitko Veta, Koen AJ Eppenhof, Josien PW Pluim, and Remco Duits.“Roto-translation covariant convolutional networks for medical image analysis”. In: International Conference on Medical Image Computing and Computer-Assisted Intervention.Springer.2018, pp. 440-448.

[11] B.M.N. Smets. “Geometric Image Denoising and Machine Learning”. Master thesis. TU Eindhoven, 2019.

[12] P.Olver. Applications of Lie Groups to Differential Equations.Springer-Verlag,2000.

[13] A. Sarti and G. Citti. A cortical based model of perceptual completion in the roto-translation space. In Proceeding of the Workshop on Second Order Subelliptic Equations and Applications. Cortona, June,15-21 2003.

[14] G. Citti and A. Sarti. A cortical based model of perceptual completion in the roto-translation space. Journal of Mathematical Imaging and Vision,24(3):307-326, may 2006.

[15] G. Sanguinetti. “Invariant models of vision between phenomenology, image statistics and neurosciences. ”. Doctoral thesis. University of Bologna.

[16] C. Nardoni.“Diffusion maps in the subriemannian motion group and perceptual grouping ”. Master thesis. University of Bologna, 2013.

[17] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni, Stratified Lie groups and potential theory for their sub-Laplacians. Springer Monographs in Mathematics,Springer, Berlin, 2007.

[18] J. M. Bony, Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés Ann. Inst.Fourier (Grenoble) 19 (1969), 277-304

[19] L. Hörmander, “Hypoelliptic second-order differential equations”, Acta Math., Vol. 119, pp. 147-171, 1967.

[20] A. Friedman, Partial Differential Equations of Parabolic Type, Dover,2008.

[21] L. C. Evans, Partial Differential Equations, Graduate Studies in Mathematics, Vol 19, 2010.

[22] J. Jost, Riemannian geometry and geometric analysis, fifth edition, Springer, 2017

[23] Remco Duits, Erik Bekkers, and Alexey Mashtakov. "Fourier transform on the homogeneous space of 3D positions and orientations for exact solutions to linear PDEs". In: Entropy 21.1 (2019), p. 38.