Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE

Corso di Laurea in Matematica

Maximum Propagation Principle for some Hörmander Operators

Tesi di Laurea in Analisi Matematica

Relatore:
Chiar.mo Prof.
Andrea Bonfiglioli

Presentata da:
Giacomo Budini

Anno Accademico 2024/2025

"Quando uno legge uno scritto di cui vuole conoscere il senso, non ne disprezza i segni e le lettere, né li chiama illusione, accidente e involucro senza valore, bensì li decifra, li studia e li ama, lettera per lettera."

- Hermann Hesse

Introduction

The aim of this thesis is to prove the so-called *Maximum Propagation Principle* for semielliptic operators L, to derive from it the *Strong Maximum Principle* (SMP, for short) for selected classes of partial differential operators and to deepen the Maximum Propagation Principle further analyzing a version of it called "*propagation along the drift*".

The following presentation will be useful to get an idea of the principal concepts and of the demonstration paths of the thesis.

First of all, we clarify what kind of differential operator we handle (during all the thesis): given $\varnothing \neq D \subseteq \mathbb{R}^N$ an open set, for each $i, j \in \{1, \dots, N\}$, we assume that $a_{i,j} = a_{j,i}$ and b_i are fixed real-valued continuous functions on D and we consider the second order linear homogeneous (that is $L(1) \equiv 0$) partial differential operators

$$L := \sum_{i,j=1}^{N} a_{i,j}(x) \frac{\partial^2}{\partial x_i \, \partial x_j} + \sum_{i=1}^{N} b_i(x) \frac{\partial}{\partial \, x_i},$$

for which we denote

$$A(x) := \big(a_{i,j}(x)\big)_{i,j \leq N}, \qquad x \in D,$$

that is the *matrix of the principal part* of L, which is supposed to be positive semidefinite and because of that L is said *semielliptic*.

Then, it is essential to understand how we see vector fields, in general: a vector field $X : D \to \mathbb{R}^N$ is a continuous function identified with the first order differential operator

$$X \equiv \sum_{i=1}^{N} X^i \frac{\partial}{\partial x_i},$$

where X^i is the i-th component-function of X. Consistently with this vision, we can multiplicate two vector fields, viewing it as a composition, through the distributive property and Leibniz rule. Practically, operators L as above are often created composing vector fields in this way; for example:

1. given $a_{i,j} = a_{j,i}$ C^1 functions on D $(0 \leq i, j \leq N)$, any *divergence form* operator

$$L = \sum_{i=1}^{N} \frac{\partial}{\partial x_i} \left(\sum_{j=1}^{N} a_{i,j}(x) \frac{\partial}{\partial x_j} \right) = \operatorname{div}\left(A(x) \cdot \nabla^T\right)$$

   satisfies the above assumptions;

2. or also $L = \sum_{j=1}^{m} X_j^2 + X_0$ is a *sum of squares* of C^1 vector fields $X_1, \ldots, X_m$ plus a C^0 *drift* X_0 on D, indeed even in this case all desired assumptions are true.

This last example of L is called *Hörmander sum of squares with drift* if
$S := \{X_0, X_1, \ldots, X_m\}$ is a *Hörmander vector-field system*, that is $S \subseteq C^\infty(D)$ and

$$\text{Span}(\{X(x) \in \mathbb{R}^N : X \in \text{Lie(S)}\}) = \mathbb{R}^N, \quad \text{for each } x \in D,$$

where $\text{Lie}(S)$ is the *Lie algebra generated* by S inside $\mathcal{X}(D) := \{C^\infty \text{ vector fields over } D\}$, which is a Lie algebra with the commutator.

For these operators L (with the extra hypothesis $X_0 \equiv 0$) we will prove the SMP that is, in general, what follows:

we say that L satisfies the *Strong Maximum Principle* on the connected open set Ω if it satisfies the following condition: for every function $u \in C^2(\Omega)$ such that

$$Lu \geq 0 \quad \text{and} \quad u \leq 0 \quad \text{on } \Omega,$$

the existence of $x_0 \in \Omega$ such that $u(x_0) = 0$ implies that $u \equiv 0$ on the whole of Ω.

In order to attain this result, we firstly need the *Maximum Propagation Principle* (MPP, for short), that is the principal result of this thesis:

let $\Omega \subseteq D$ be an open set; for every function $u \in C^2(\Omega)$ satisfying

$$Lu \geq 0 \quad \text{and} \quad \leq 0 \quad \text{on } \Omega,$$

the set $F(u) = \{x \in \Omega : u(x) = 0\}$ contains the trajectories, starting at points of $F(u)$, of the integral curves of any C^1 principal vector field for L over Ω.

Just to get the idea, an *integral curve* of a vector field X starting at the point x_0 is a solution γ (on an interval containing zero) for the Cauchy problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = X(\gamma(t)) \\ \gamma(0) = x_0. \end{cases}$$

And there is also a precise definition for *L-principality* of a vector field:

if X is a vector field on Ω, we say that X is a *principal vector field for* L (on Ω) if for every $x \in \Omega$ there exists a real number $\lambda(x) > 0$ such that

$$\langle X(x), \xi \rangle^2 \leq \lambda(x) \left\langle A(x)\, \xi, \xi \right\rangle, \qquad \text{for every } \xi \in \mathbb{R}^N.$$

The MPP can be rephrased by the concept of *invariant set* with respect to a vector field:

let X be a vector field on Ω, and let F be a subset of Ω; we say that F is *positively X-invariant* (or *positively invariant with respect to X*) if, for every integral curve γ of X, $\gamma : [0, T] \to \Omega$ satisfying $\gamma(0) \in F$, we have $\gamma(t) \in F$ for every $t \in [0, T]$; we say that F is *X-invariant* (or *invariant with respect to X*) if it is positively invariant with respect to X and to $-X$.

To proceed we introduce a definition of tangentiality, that is in a certain sense milder than the one from differential geometry, and it can be seen in the concept of *external orthogonality*:

let F be a closed subset of Ω and let $y \in \Omega \cap \partial F$; we say that a non-null vector $\nu \in \mathbb{R}^N$ is *externally orthogonal to F at y* if

$$\overline{B\big(y + \nu, \|\nu\|\big)} \subseteq (\Omega \setminus F) \cup \{y\};$$

in this case we shall write $\nu \perp F$ at y and we also let

$$F^* := \Big\{ y \in \Omega \cap \partial F \,\Big|\, \text{there exists } \nu \perp F \text{ at } y \Big\}.$$

Provided that F is a closed proper subset in Ω, we prove that $F^* \neq \varnothing$.

Always as for tangentiality, we have a notion of *vector field tangent to a closed set*:

let F be a closed subset in Ω; we suppose that X is a vector field on Ω; we say that *X is tangent to F* if

$$\langle X(y), \nu \rangle = 0, \quad \text{for each } \; y \in F^* \quad \text{and} \quad \text{for each } \; \nu \perp F \text{ at } y.$$

Then we show the *Nagumo-Bony Theorem* which establishes the equivalence between invariancy and tangentiality. Indeed one of its immediate consequences is:

let X be a C^1 vector field on Ω; suppose that F is a relatively closed subset of Ω; then, F is X-invariant if and only if X is tangent to F.

This result, once we have introduced suitable notation, is summed up by $\mathrm{Tg}(F) = \mathrm{Inv}(F)$.

Going deeper, a whole section is entirely devoted to the demonstration of the so-called *Hopf Lemma*. Here, we give only its statement and its direct corollary:

let $\Omega \subseteq D$ be a connected open set; we consider $u \in C^2(\Omega)$ such that $Lu \geq 0$ and $u \leq 0$ over Ω; we set $F(u) := \{x \in \Omega : u(x) = 0\}$ and we suppose that $F(u)$ is a proper subset of Ω;

then, for each $y \in F(u)^*$ and for each $\nu \perp F$ at y, we have

$$\langle A(y)\nu, \nu\rangle = 0, \text{ that is } \nu \in \mathrm{Isotr}(A(y)).$$

It easily follows that, if X is a principal vector field for L, then X is tangent to $F(u)$.

Actually, at this point, the proof of the MPP is quite easy: one needs only to put together all the pieces briefly described above. Furthermore, as a plus, we derive the following result:

given F a non-empty closed set of $\mathbb{R}^N$ contained in Ω, we find that

$$\{X \in \mathcal{X}(\Omega) : F \text{ is } X\text{-invariant}\} = \mathrm{Inv}(F) = \mathrm{Tg}(F) = \{X \in \mathcal{X}(\Omega) : X \text{ is tangent to } F\}$$

is Lie sub-algebra of $\mathcal{X}(\Omega)$ (i.e., it is closed by the commutator $[X, Y] = XY - YX$).

(This last proposition requires a preliminary study of *vector-field flows.*)

And finally we derive our first goal:

if $\{X_1, \ldots, X_m\}$ are a Hörmander vector fields over an open set $D \subseteq \mathbb{R}^N$, then the associated Hörmander sum of squares $\sum_{i=1}^m X_i^2$ satisfies the SMP over each connected open set $\Omega \subseteq D$.

In conclusion, in the last chapter, we analyze a propagation extra-result that, from a certain point of view, completes the previous analysis. Indeed, until the last result, nothing has been said about the propagation along the integral curves of the drift X_0, provided that one is also interested in the operators of the form $\sum_{j=1}^m X_j^2 + X_0$. More precisely, rewriting the initial differential operator L in the following form

$$\begin{aligned} L &= \sum_{i=1}^N \frac{\partial}{\partial x_i}\left(\sum_{j=1}^N a_{i,j}(x)\frac{\partial}{\partial x_j}\right) + \sum_{j=1}^N \left(b_j(x) - \sum_{i=1}^N \frac{\partial a_{i,j}}{\partial x_i}(x)\right)\frac{\partial}{\partial x_j} \\ &= \sum_{i=1}^N \frac{\partial}{\partial x_i}\Big(X_i\Big) + X_0, \end{aligned}$$

where X_0 is called *the drift* of L, the result has the following proposition:

given $u \in C^2(\Omega)$ an L-subharmonic function over a non-empty open set $\Omega \subseteq D$, if we assume the set of maximum points $F(u)$ of u is non-empty, then $F(u)$ is positively X_0-invariant.

It is not a result of X_0-invariance, but only of positive invariance. And there is the example of the known *Heat Operator*

$$L_{\text{Heat}} = (\partial x_1)^2 + \cdots + (\partial x_N)^2 - \partial x_{N+1},$$

to testify that this is the strongest propagation result that we can reach about the drift.

Contents

Chapter 1

Preliminaries

1.1 Assumptions and examples

During all this thesis, we shall be dealing with the following type of partial differential operators.

Let $\varnothing \neq D \subseteq \mathbb{R}^N$ be an open set. For every $i, j \in \{1, \dots, N\}$, we assume that $a_{i,j} = a_{j,i}$ and b_i are fixed real-valued continuous functions on D. We consider the second order linear homogeneous (that is $L(1) \equiv 0$) partial differential operator on D

$$L := \sum_{i,j=1}^{N} a_{i,j}(x) \frac{\partial^2}{\partial x_i \, \partial x_j} + \sum_{i=1}^{N} b_i(x) \frac{\partial}{\partial \, x_i}. \tag{1.1}$$

We introduce the notation

$$A(x) := \big(a_{i,j}(x)\big)_{i,j \leq N}, \qquad x \in D,$$

for the *matrix of the principal part* of L. The map $\xi \mapsto q_L(x, \xi) := \langle A(x)\xi, \xi \rangle$ denotes the *characteristic form* of L (at $x \in D$).

In the sequel, we make the following assumption:

$$L \text{ is } \textit{semielliptic},$$

that is, $A(x)$ is positive semidefinite for every $x \in D$.

Throughout the chapter, the above notations and assumptions on L are tacitly assumed.

Remark 1.1. We say once and for all that, when we consider a C^1 vector field $X = (X_1, \dots, X_N)$ on D, we identify X with the associated first order differential operator $X \equiv \sum_{i=1}^{N} X_i(x) \frac{\partial}{\partial x_i}$. As a consequence, when we "multiply" vector fields with each other, we understand the composition of differential operators.

Example 1.2. We have the following examples.

1. Any sum of squares of C^1 vector fields $L = \sum_{j=1}^m X_j^2$ on D is of the form required above. Indeed L is semielliptic, since a brief calculation gives

$$q_L(x,\xi) = \sum\nolimits_{j=1}^m \left\langle X_j(x), \xi \right\rangle^2, \qquad x \in D, \quad \xi \in \mathbb{R}^N. \tag{1.2}$$

2. Suppose $L = \sum_{j=1}^m X_j^2 + X_0$ is a *sum of squares* of C^1 vector fields $X_1, \ldots, X_m$ plus a C^0 *drift* X_0 on D. We claim that L is semielliptic. Indeed this property only depends on the principal part of L, which is completely determined by $\sum_{j=1}^m X_j^2$; hence our claim follows from (1.2).

3. If $a_{i,j} = a_{j,i}$ are C^1 functions on D $(0 \le i,j \le N)$, any *divergence form* operator

$$L = \sum_{i=1}^N \frac{\partial}{\partial x_i} \left(\sum_{j=1}^N a_{i,j}(x) \frac{\partial}{\partial x_j} \right) = \operatorname{div}\left(A(x) \cdot \nabla^T \right)$$

satisfies the above assumptions, provided that $A(x)$ is positive semidefinite at any $x \in D$. Under this assumption on A, the same is true of

$$L = \frac{1}{V(x)} \sum_{i=1}^N \frac{\partial}{\partial x_i} \left(V(x) \sum_{j=1}^N a_{i,j}(x) \frac{\partial}{\partial x_j} \right),$$

with $V \in C^1$ and $V > 0$ on D. Out of curiosity, the latter is the typical form of the *Laplace-Beltrami operator* written in coordinates, and of many meaningful partial differential operators on Lie groups.

1.2 A kit of several tools

To deepen the principal topic, we need some definitions and results that will be useful in the following chapters.

Now, we can analyze a linear algebra preliminary proposition, which will be of assistance in demonstrating some principality characterizations in the next chapter start.

Proposition 1.3. *We consider two $(N \times N)$-dimensional real matrices B and C, and a $(N \times m)$-dimensional real matrix S. Then we have:*

1. $\ker B \subseteq \operatorname{Isotr}(B)$ *and* $(\ker B)^\perp = \operatorname{Im}(B^\intercal)$ *(in particular if B is symmetric we gain immediately* $(\ker B)^\perp = \operatorname{Im}(B)$*);*

2. *if B is positive semidefinite, then* $\ker B = \operatorname{Isotr}(B)$*;*

3. if B is positive semidefinite and C is negative semidefinite, then $\operatorname{trace}(BC) \leq 0$;

4. the $(N \times N)$-dimensional real matrix $D := SS^\intercal$ is symmetric and positive semidefinite, $\ker D = \ker S^\intercal$ *and* $\operatorname{Im}(D) = \operatorname{Im}(S)$.

We have set $\operatorname{Isotr}(B) := \{v \in \mathbb{R}^N : v^\intercal B v = 0\}$ *and the orthogonal vector space is performed referring to the standard scalar product in* $\mathbb{R}^N$.

Proof. Regarding the first point: $\ker B \subseteq \operatorname{Isotr}(B)$ is trivial; furthermore $(\ker B)^\perp = \operatorname{Im}(B^\intercal)$ is equivalent to $\ker B = (\operatorname{Im}(B^\intercal))^\perp$, which is easy to prove.

As for the second point, from the fact that B is positive semidefinite it follows that B is orthogonally diagonalizable: $B = Q^\intercal M Q$, where M is a diagonal matrix that can be supposed to have the first r diagonal terms strictly positive and the following diagonal terms equal to zero. Now, if x is an isotropic vector of B, we find that

$$0 = \sum_{i=1}^{r} M_i((Qx)_i)^2, \text{ which gives } (Qx)_1 = \cdots = (Qx)_r = 0.$$

This implies the thesis, due to the form of the matrix

$$M = \begin{pmatrix} M_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & M_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

About the third point, we can say that if B is symmetric, B is orthogonally similar to a diagonal matrix with non-negative diagonal elements, which can be supposed to be equal to 1 or 0. If E is such a diagonal matrix, then $B = PEP^\intercal$. Then we also have $E = EE^\intercal$. Hence, by placing $R := PE$, we can write

$$B = PEE^\intercal P^\intercal = RR^\intercal \text{ and } \operatorname{trace}(BC) = \operatorname{trace}(RR^\intercal C) = \operatorname{trace}(R^\intercal CR),$$

where in the last identity we used the invariance of the trace by commutation of matrices. In conclusion, if C is symmetric and negative semidefinite, the same is true of $R^\intercal CR$, thus in particular this last matrix is similar to a diagonal matrix with non-positive diagonal elements, which gives what we desired, by remembering that the trace is invariant by similarity.

Referring to the last point, one has

$$D^\intercal = (SS^\intercal)^\intercal = (S^\intercal)^\intercal S^\intercal = SS^\intercal = D,$$

and, for each $\nu \in \mathbb{R}^N$

$$\nu^\intercal D\nu = \nu^\intercal SS^\intercal \nu = (S^\intercal \nu)^\intercal (S^\intercal \nu) = \|S^\intercal \nu\|^2 \geq 0.$$

Hence, from the second point, we derive $\ker D = \mathrm{Isotr}(D)$. From this we can prove $\ker D \subseteq \ker S^\intercal$: if $\nu \in \ker D = \mathrm{Isotr}(D)$, then $\|S^\intercal \nu\|^2 = 0$, that implies $S^\intercal \nu = 0$, that is $\nu \in \ker S^\intercal$. The other inclusion is trivial, so $\ker D = \ker S^\intercal$.

Finally, we prove that $\mathrm{Im}(D) = \mathrm{Im}(S)$: if $\ker D = \ker S^\intercal$, then

$$\dim(\mathrm{Im}(D)) = N - \dim(\ker D) = N - \dim(\ker S^\intercal) = \dim(\mathrm{Im}(S^\intercal)) = \dim(\mathrm{Im}(S)),$$

and trivially $\mathrm{Im}(D) \subseteq \mathrm{Im}(S)$ also holds. □

Then, we can introduce the strict subharmonicity referring to L, a concept that will be useful in Hopf Lemma and, first of all, it appears in every maximum propagation result of this thesis.

Definition 1.4 (L-subharmonicity). *A function $u \in C^2(\Omega)$ is called strictly L-subharmonic over Ω if*

$$Lu(x) > 0 \text{ for each } x \in \Omega,$$

and it is simply called L-subharmonic over Ω if

$$Lu(x) \geq 0 \text{ for each } x \in \Omega.$$

Proposition 1.5. *If $u \in C^2(\Omega)$ is strictly L-subharmonic over Ω, then u has no local maximum points inside Ω.*

Proof. Proceeding by contradiction, we suppose that such a u has a local maximum point $\xi_0 \in \Omega$. Applying the operator L to u, we find

$$Lu(\xi_0) = \sum_{i,j=1}^{N} a_{i,j}(\xi_0)\, \frac{\partial^2 u}{\partial x_i\, \partial x_j}(\xi_0) + \sum_{i=1}^{N} b_i(\xi_0)\, \frac{\partial u}{\partial\, x_i}(\xi_0).$$

By the contradiction hypothesis, if $u \in C^2(\Omega)$, the gradient of u at ξ_0 equals zero, that is $\nabla_{\xi_0} u = 0$, and the Hessian matrix of u at ξ_0, denoted by $H_{\xi_0} u$, is symmetric and negative semidefinite. This being established, the previous identity becomes

$$Lu(\xi_0) = \mathrm{trace}(A(\xi_0) \cdot H_{\xi_0} u) \leq 0,$$

where the last inequality comes from the third point of the Proposition 1.3. This is in contradiction with the strict L-subharmonicity of u over Ω, that would give $Lu(\xi_0) > 0$. □

We also have to introduce what integral curves are. In the following chapter, they will be the ones along which the maximum propagates in the following chapters.

Definition 1.6 (**Integral curve**). *Given a sufficiently regular vector field X over D and fixed a point $x_0 \in D$, we call the integral curve of X with initial point x_0 the maximal solution $\gamma(t)$ (as a function of t) on its maximal domain $I(X, x)$ (that is known to be an open interval of $\mathbb{R}$) of the following Cauchy problem:*

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = X(\gamma(t)) \\ \gamma(0) = x_0. \end{cases}$$

Remark 1.7. In the following chapter we will talk about integral curves of a vector field simply by referring to functions $\gamma : [a, b] \to D$ which satisfies the differential equation $\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = X(\gamma(t))$ for each $t \in [a, b]$, where $a \leq 0 \leq b$. This latter γ will be understood as a restriction of an integral curve that respects the previous definition.

Definition 1.8 (**Flow**). *Let X be a C^1 vector field over D and $x \in D$. We denote by $\Psi_t^X(x)$ the flow of the vector field X at the time t, which is the map from D_t^X to D that, fixing time t, associates $x \mapsto \gamma_x(t) \equiv \gamma(t)$, where γ is the integral curve of X starting at the point x and $D_t^X := \{x \in \Omega : t \in \ I(X, x)\}$.*

Remark 1.9. Let $f \in C^2(D, \mathbb{R})$ and the previous definition be understood. Fixing $x \in \Omega$, we can consider the composition

$$D(X, x) \to \mathbb{R}, \quad t \mapsto f(\Psi_t^X(x)) \equiv f(\gamma(t)).$$

It is easy to recognize that the second order Taylor expansion of this function in 0 is

$$f(\gamma(t)) = f(x) + t(Xf)(x) + \frac{t^2}{2}(X^2 f)(x) + o(t^2), \text{ if } t \to 0. \tag{1.3}$$

Iteratively choosing f as the i-th projection $D \to \mathbb{R}$ and placing $I : D \to D$ the identity over D, we attain the N-dimensional development

$$\Psi_t^X(x) \equiv \gamma(t) = f(x) + t(XI)(x) + \frac{t^2}{2}(X^2 I)(x) + o(t), \text{ if } t \to 0.$$

If f is as above and (t, s) is in a suitable neighborhood of $(0, 0) \in \mathbb{R}^2$, giving two C^2 vector fields X and Y over Ω, the following composition is well posed:

$$\Gamma(t, s) = f(\Psi_s^Y \circ \Psi_t^X).$$

(After the following proposition, we will dwell on the well-posed definition and the regularity of this function Γ.)

Proposition 1.10 (Flow composition). *Let X, Y be C^2 vector fields over D and $x \in D$. Then*

$$\Psi_t^{-Y} \circ \Psi_t^{-X} \circ \Psi_t^{Y} \circ \Psi_t^{X}(x) = x + t^2([X,Y]I)(x) + o(t^2), \text{ if } t \to 0.$$

Here and in the sequel $[X,Y] = XY - YX$ is the commutator of X and Y.

Proof. Given $f \in C^2(D)$, we firstly define

$$\Lambda(t_1, t_2, t_3, t_4) := f(\Psi_{t_4}^{Y} \circ \Psi_{t_3}^{X} \circ \Psi_{t_2}^{Y} \circ \Psi_{t_1}^{X}(x)),$$

that is well posed and C^2 in a suitable symmetric neighborhood of $(t_1, t_2, t_3, t_4) = (0,0,0,0) \in \mathbb{R}^4$ (see Remark 1.11). Applying several times (1.3), we find the second order Taylor polynomial of Λ at $(t_1, t_2, t_3, t_4) = (0,0,0,0)$ is:

$$\sum_{0 \le k_1+k_2+k_3+k_4 \le 2} \left(\left[\frac{(t_1X)^{k_1}}{k_1!} \frac{(t_2X)^{k_2}}{k_2!} \frac{(t_3X)^{k_3}}{k_3!} \frac{(t_4X)^{k_4}}{k_4!} \right] f \right)(x).$$

Choosing $t_1 = t_2 \equiv t$ and $t_3 = t_4 \equiv -t$, we have the second order Taylor polynomial of the function $\Lambda(t,t,-t,-t) := f(\Psi_{-t}^{Y} \circ \Psi_{-t}^{X} \circ \Psi_t^{Y} \circ \Psi_t^{X}(x))$ that, after some cancellations, can be seen to be $f(x) + t^2([X,Y]f)(x)$.

Replacing f with I, we obtain the thesis:

$$\Psi_{-t}^{Y} \circ \Psi_{-t}^{X} \circ \Psi_t^{Y} \circ \Psi_t^{X}(x) = x + t^2([X,Y]I)(x) + o(t^2), \text{ if } t \to 0.$$

We have concluded proof (also by the next Remark). □

Remark 1.11. We provide three clarifications:

1. The C^2-regularity of the flow $\Psi_t^X(x)$ with respect to time t and position $x \in D$ used in the above demonstration derives from the C^2-regularity of the associated vector field X, through some general ODE theory results; this is the reason why Λ is C^2.

2. If X is a C^1 vector field over D, $x \in D$ and $r, \alpha \in \mathbb{R}$ are such that $\alpha r \in D(X,x)$, then $r \in D(\alpha X, x)$ and $\Psi_{\alpha r}^X(x) = \Psi_r^{\alpha X}(x)$; we can give a proof to this fact:

   let $\gamma : I(X,x) \to D$ be the unique maximal solution of the Cauchy problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = X(\gamma(t)) \\ \gamma(0) = x_0. \end{cases}$$

we initially suppose $\alpha \neq 0$ and we set

$$J := \frac{I(X,x)}{\alpha} := \{r \in \mathbb{R} : \alpha r \in I(X,x)\},$$

which is an interval containing 0; we can well-pose the function

$$u : J \to \Omega, \quad u(r) := \gamma(\alpha r);$$

we find $u(0) = \gamma(0) = x$ and

$$\frac{\mathrm{d}}{\mathrm{d}r}u(r) = \alpha\dot{\gamma}(\alpha r) = \alpha X(\gamma(\alpha r)) = (\alpha X)(u(r)), \text{ for each } r \in J,$$

which means that u is a solution of the Cauchy problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}r}u(r) = (\alpha X)(u(r)) \\ u(0) = x_0. \end{cases}$$

This implies $J \subseteq I(\alpha X, x)$ and, on J, u coincides with the maximal solution of this last problem; in particular $\alpha r \in I(X,x)$ (being $\alpha \neq 0$) implies $r \in J \subseteq I(\alpha X, x)$, hence

$$\Psi^X_{\alpha r}(x) = \gamma(\alpha r) = u(r) = \Psi^{\alpha X}_r(x);$$

the case $\alpha = 0$ is trivial because a null vector field has as integral curve starting at x which is a constant function equal to x (whose domain obviously contains $0 = \alpha r$).

3. Again via some general ODE theory results, given a starting point $x \in D$ and a C^2 vector field X on D, there exists $\epsilon > 0$ and $r > 0$ such that $B(x,r) \subseteq D$ and the function

$$[-\epsilon, \epsilon] \times B(x,r) \to D, \quad (s,y) \mapsto \Psi^X_s(y)$$

is well-posed (that is $s \in I(X,y)$ for each $(s,y) \in [-\epsilon,\epsilon] \times B(x,r)$); then, given another C^2 vector field Y, the function

$$I(Y,x) \to D, \quad t \mapsto \Psi^Y_t(x)$$

is continuous at 0, hence there is a $\delta > 0$ such that for each $t \in [-\delta, \delta]$ we have $\Psi^Y_t(x) \in B(x,r)$; all this being said, we infer that the following composition of flows is well-defined in a symmetric neighborhood of $(0,0)$:

$$[-\delta,\delta] \times [-\epsilon,\epsilon] \to D, \quad (t,s) \mapsto \Psi^X_s(\Psi^Y_t(x));$$

and the same can be done with a finite number of flows of different vector fields.

With the last section of this thesis in mind, we have to deepen our knowledge of the flows, looking for the second-order MacLaurin expansion of a flow in both time and space.

Remark 1.12. We consider a C^1-vector field X over the open set D. Then we define

$$\mathcal{D}(X) := \{(t,x) \in \mathbb{R} \times \mathbb{R}^N : x \in D,\ t \in I(X,x)\}.$$

Hence we can see the flow of X as the map

$$\mathcal{D}(X) \to D, \quad (t,x) \mapsto \Psi_t^X(x). \tag{1.4}$$

From general ODE theory results, $\mathcal{D}(X)$ is an open set in $\mathbb{R} \times \mathbb{R}^N$ and, for each $k \in \{1,2,3,\dots,\infty\}$, if X is C^k, the function (1.4) is C^k as well, and it admits all $(k+1)$-th order derivatives of the type

$$\frac{\partial^{k+1}}{\partial_t \partial x_{i_1} \dots \partial x_{i_k}} \Psi_t^X(x),$$

with $i_1, \dots, i_k \in \{1,\dots,N\}$, where differentiation can be interchanged without modification of the outcome and these derivatives are continuous in $(t,x) \in \mathcal{D}(X)$. Thus, if X is C^1, for any $(t,x) \in \mathcal{D}(X)$ and for any $j \in \{1,\dots,N\}$, we have

$$\frac{\partial^2}{\partial_t \partial x_j} \Psi_t^X(x) = \frac{\partial^2}{\partial x_j \partial_t} \Psi_t^X(x). \tag{1.5}$$

We also remember that, from Definitions 1.6 and 1.8, we have, for each $(t,x) \in \mathcal{D}(X)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} \Psi_t^X(x) = X(\Psi_t^X(x)), \quad \Psi_0^X(x) = x. \tag{1.6}$$

The equalities (1.5) and (1.6) give

$$\mathrm{J}_{\Psi_t^X(x)}(XI) \cdot \mathrm{J}_x(\Psi_t^X) = \mathrm{J}_x(XI \ \circ \ \Psi_t^X) = \mathrm{J}_x\left(\frac{\mathrm{d}}{\mathrm{d}t}\Psi_t^X\right) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{J}_x(\Psi_t^X), \quad \mathrm{J}_x(\Psi_0^X) = \mathrm{Id}_{N\times N}.$$

This means that the matrix-valued map

$$\mathcal{D} \to \mathbb{R}^{N\times N}, \quad (t,x) \mapsto \mathrm{W(t,x)} := \mathrm{J}_x(\Psi_t^X)$$

solves the linear ODE system

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{W(t,x)} = \mathrm{C}(t,x) \cdot \mathrm{W(t,x)}, & (t,x) \in \mathcal{D}(X) \\ \mathrm{W(0,x)} = \mathrm{Id}_{N\times N}, & x \in D, \end{cases} \tag{1.7}$$

where $\mathrm{C}(t,x) := \mathrm{J}_{\Psi_t^X(x)}(XI)$ for each $(t,x) \in \mathcal{D}(X)$.

The system (1.7), by the Fundamental Theorem of Calculus, is equivalent to

$$\mathrm{J}_x(\Psi_t^X) = \mathrm{Id}_{N\times N} + \int_0^t \mathrm{J}_{\Psi_\tau^X(x)}(XI) \cdot \mathrm{J}_x(\Psi_\tau^X)\ \mathrm{d}\tau, \quad \text{for each } (t,x) \in \mathcal{D}(X). \tag{1.8}$$

Now, we are finally ready to show, supposing X is C^2 on D and $0 \in D$, denoting $X = \sum_{i=1}^N \sigma_i(x)\frac{\partial}{\partial x_i}$ and $\gamma(t,x) := \Psi_t^X(x)$ (for each $(t,x) \in \mathcal{D}(X)$), for every $i \in \{1,\dots,N\}$, the second-order expansion

$$\gamma_i(t,x) = \sigma_i(0)t + x_i + \frac{1}{2}\Big((X\sigma_i)(0)t^2 + 2t\langle\nabla_0\sigma_i, x\rangle\Big) + \mathrm{o}(\|(t,x)\|^2), \tag{1.9}$$

as $(t,x) \to (0,0)$ in $\mathbb{R}\times\mathbb{R}^N$.

By the definition of flow

$$\gamma(0,x) = x \quad \text{and} \quad \frac{\partial \gamma_i}{\partial t}(t,x) = \sigma_i(x) \quad \text{for each } (t,x) \in \mathcal{D}(X), \tag{1.10}$$

hence

$$\gamma(0,0) = 0, \quad \frac{\partial \gamma_i}{\partial t}(0,0) = \sigma_i(0), \quad \nabla_0\gamma_i(0,\cdot) = e_i;$$

then, from (1.8), for every $i \in \{1,\dots,N\}$, we have

$$\nabla_x\gamma_i(t,\cdot) = e_i + \int_0^t \nabla_{\gamma(x,\tau)}\sigma_i \cdot \mathrm{J}_x(\gamma(\tau,\cdot))\,\mathrm{d}\tau, \quad \text{for each } (t,x) \in \mathcal{D}(X). \tag{1.11}$$

Thus, using (1.10) and (1.11), for every $i \in \{1,\dots,N\}$, we gain

$$\frac{\partial^2\gamma_i}{\partial t^2}(0,0) = \frac{\partial i}{\partial t}\Big|_{t=0}\sigma_i(\gamma(t,0)) = \nabla_0\sigma_i \cdot X(0) = (X\sigma_i)(0),$$

$$(\mathrm{Hess}_{(0,0)}\gamma_i(\cdot,\cdot))_{h,k} = \frac{\partial^2\gamma_i}{\partial x_h\partial x_k}(0,0) = \frac{\partial^2}{\partial x_h\partial x_k}\Big|_{x=0}\gamma_i(0,x) = 0 \quad \text{for each } h,k \in \{1,\dots,N\},$$

$$\nabla_x\Big|_{x=0}\frac{\partial}{\partial t}\Big|_{t=0}(\gamma_i(t,x)) = \frac{\partial}{\partial t}\Big|_{t=0}\nabla_x\Big|_{x=0}(\gamma_i(t,x)) = \nabla_0\sigma_i \cdot \mathrm{J}_0(\gamma(0,\cdot)) = \nabla_0\sigma_i,$$

where in the first equality of the third row we have used (1.5).

If we insert what we found in

$$\gamma_i(t,x) = \gamma(0,0) + \langle\nabla_{(0,0)}\gamma_i(\cdot,\cdot),(t,x)\rangle + \frac{1}{2}\langle\mathrm{Hess}_{(0,0)}\gamma_i(\cdot,\cdot)\cdot(t,x),(t,x)\rangle + \mathrm{o}(\|(t,x)\|^2),$$

as $(t,x) \to (0,0)$, we get (1.9) (for every $i \in \{1,\dots,N\}$).

Now, we turn to a completely different issue. It is quite easy to see that the minimal distance between a point and a closet set (not containing that point) is equal to the distance between the point and the boundary of the set; but we also have to prove it formally.

Proposition 1.13 (**Distance and boundary**). *If F is a closed set in $\mathbb{R}^N$, $x \in \mathbb{R}^N \setminus F$, $v \in F$ is such that $\|v - x\| = \mathrm{dist}(x,F) > 0$, then $v \in \partial F$ and $\mathrm{dist}(x,F) = \mathrm{dist}(x,\partial F)$.*

Proof. We suppose, by contradiction, that $v \notin \partial F$ and we consider the continuous function

$$H : \mathbb{R}^N \to \mathbb{R}, \quad H(y) := \begin{cases} \mathrm{dist}(y, \partial F), & y \in F \\ -\mathrm{dist}(y, \partial F), & y \notin F. \end{cases}$$

If $\mathrm{dist}(v, \partial F) > 0$ and $-\mathrm{dist}(x, \partial F) < 0$, then, thanks to the Zero Theorem, we gain that there is a $q \in \{tx + (1-t)v \in \mathbb{R}^N : t \in [0,1]\} \setminus \{x, v\}$ such that $\mathrm{dist}(q, \partial F) = 0$. ∂F is closed, thus $q \in \partial F$. And $\|q - x\| < \|v - x\|$, which contradicts $\|v - x\| = \mathrm{dist}(x, F)$. □

As we shall see, triangle inequality will not be enough to prove the Nagumo-Bony Theorem. We have to generalize it.

Lemma 1.14 (Generalization of the triangle inequality). *For each $a, b, c \in \mathbb{R}^N$, whenever $a \neq c, a \neq b$ the next inequality holds:*

$$\|a - b\| \leq \left\langle \frac{c-a}{\|c-a\|}, \frac{b-a}{\|b-a\|} \right\rangle \cdot \|c - a\| + \|c - b\|.$$

Proof. We distinguish two cases.

In the first case we suppose $\langle c - a, b - a \rangle \geq 0$. From a simple direct calculation we gain

$$b - a = \left\langle c - a, \frac{b-a}{\|b-a\|} \right\rangle \cdot \frac{b-a}{\|b-a\|} - \left\langle c - b, \frac{b-a}{\|b-a\|} \right\rangle \cdot \frac{b-a}{\|b-a\|}.$$

Thanks to triangle inequality and, for the second summand, the Cauchy-Schwarz inequality, we have

$$\|b - a\| \leq \left| \left\langle c - a, \frac{b-a}{\|b-a\|} \right\rangle \right| + \|c - b\| = \left| \left\langle \frac{c-a}{\|c-a\|}, \frac{b-a}{\|b-a\|} \right\rangle \right| \cdot \|c - a\| + \|c - b\|.$$

Due to $\langle c - a, b - a \rangle \geq 0$ we have the desired result.

Now instead we instead suppose $\langle c - a, b - a \rangle < 0$. Even in this case, through a direct calculation, we obtain

$$\|b - a\| = \left\langle c - a, \frac{b-a}{\|b-a\|} \right\rangle - \left\langle c - b, \frac{b-a}{\|b-a\|} \right\rangle.$$

In our hypothesis, collecting factors from the scalar products, the previous implies:

$$\|b - a\| + \left| \left\langle \frac{c-a}{\|c-a\|}, \frac{b-a}{\|b-a\|} \right\rangle \right| \cdot \|c - a\| = \left| \left\langle \frac{c-b}{\|c-b\|}, \frac{b-a}{\|b-a\|} \right\rangle \right| \cdot \|c - b\|.$$

By the Cauchy-Schwarz inequality applied to the second part of the equality, we set

$$\|b - a\| + \left| \left\langle \frac{c-a}{\|c-a\|}, \frac{b-a}{\|b-a\|} \right\rangle \right| \cdot \|c - a\| \leq \|c - b\|.$$

That is

$$\|c - b\| - \left| \left\langle \frac{c-a}{\|c-a\|}, \frac{b-a}{\|b-a\|} \right\rangle \right| \cdot \|c - a\| \geq \|b - a\|.$$

Remembering the assumption of this second case, this is the end of the proof. □

Again functional to the proof of the Nagumo-Bony Theorem we will require a real Analysis lemma.

Lemma 1.15. *Let $g : [0, T] \to \mathbb{R}$ be a continuous function such that*

$$\limsup_{h\to 0^-} \frac{g(t+h) - g(t)}{h} \leq M \quad \textit{for every } t \in (0, T],$$

for a constant $M \in \mathbb{R}$. Then $g(t) \leq g(0) + M\,t$, for all $t \in [0, T]$.

Proof. Let $\epsilon > 0$. We define the following function

$$G : [0, T] \to \mathbb{R}, \quad t \mapsto G(t) := g(t) - g(0) - (M + \epsilon)t.$$

Thanks to Weierstrass Theorem, as G continuous on a compact interval, there exists $t_0 \in [0, T]$ such that $G(t_0) = \max_{[0,T]} G$. Now we prove that $t_0 = 0$.

By contradiction, we suppose $t_0 \in (0, T]$. Hence for each $t \in [0, T]$ we have

$$g(t) - g(0) - (M + \epsilon)t \leq g(t_0) - g(0) - (M + \epsilon)t_0.$$

We can set $t = t_0 + h$ with $h < 0$ such that $t_0 + h \in (0, t_0)$. From this we gain

$$M + \epsilon \leq \frac{g(t_0 + h) - g(t_0)}{h}.$$

Passing to $\limsup_{h\to 0^-}$ we have $M + \epsilon \leq M$, that cannot be true if $\epsilon > 0$. So $t_0 = 0$.

Hence the first inequality of this proof becomes

$$g(t) - g(0) - (M + \epsilon)t \leq 0, \quad \text{for each } t \in [0, T].$$

Letting ϵ go to 0, the thesis follows. □

Lemma 1.16. *Let $g : [0, T] \to \mathbb{R}$ be a continuous and non-negative function such that $g(0) = 0$ and*

$$\limsup_{h\to 0^-} \frac{g(t+h) - g(t)}{h} \leq L\,g(t) \quad \textit{for every } t \in (0, T], \tag{1.12}$$

for a constant $L \geq 0$. Then $g \equiv 0$ on $[0, T]$.

Proof. Let $\varepsilon > 0$ be so small that $\varepsilon < T$ and $L\varepsilon < 1$. We show that $g \equiv 0$ on $[0, \varepsilon]$; by repeating the same argument finitely-many times we derive that $g \equiv 0$ on $[0, T]$. g satisfies the hypothesis of the Lemma 1.15 on $[0, \varepsilon]$, with the constant $M = L \sup_{[0,\varepsilon]} g$. Thus, from the Lemma 1.15 (and $g(0) = 0$), we get

$$g(t) \leq M\,t \leq M\,\varepsilon = L\,\varepsilon \sup_{[0,\varepsilon]} g, \qquad \forall\ t \in (0, \varepsilon].$$

By taking the supremum over $[0, \varepsilon]$, we get

$$\sup_{[0,\varepsilon]} g \leq L\,\varepsilon \sup_{[0,\varepsilon]} g, \quad \text{which implies} \quad (1 - L\varepsilon) \sup_{[0,\varepsilon]} g \leq 0.$$

Since $L\varepsilon < 1$ and $g \geq 0$, this is possible only if $g \equiv 0$ on $[0, \varepsilon]$. □

The last tool that we need (in particular, for the Maximum Propagation Principle along the drift) is summerized in the following lemma.

Lemma 1.17. *For every $i, j \in \{1, \dots, N\}$, let $m_{i,j}$ be differentiable functions on an open set $U \subseteq \mathbb{R}^N$. Let $M(x) := (m_{i,j}(x))_{i,j}$ be symmetric and positive semidefinite for each $x \in U$. Let us assume that there exists a point $x_0 \in U$ and two vectors $u, v \in \mathbb{R}^N$ such that $\langle M(x_0)u, u\rangle = \langle M(x_0)v, v\rangle = 0$. Then, for every $k \in \{1, \dots, N\}$,*

$$\langle \partial_{x_k} M(x_0)u, v\rangle = 0,$$

where $\partial_{x_k} M(x) := \left(\frac{\partial}{\partial x_k} m_{i,j}(x)\right)_{i,j}$, for each $x \in U$.

Proof. $M(x_0)$ is symmetric and semidefinite, hence it is orthogonally diagonalizable, let we say $M(x_0) = Q^\intercal D Q$, where diagonal coefficients of D are all non-negative. So there exists a diagonal real matrix $\widetilde{D}$ (with non-negative diagonal terms as well) such that $D = \widetilde{D}\widetilde{D} = (\widetilde{D})^\intercal \widetilde{D}$. Thus, placing $L := \widetilde{D}Q$, one easily derives $M(x_0) = L^\intercal L$.

By this fact (using the Cauchy-Schwarz inequality with respect to the standard scalar product in $\mathbb{R}^N$) we can write

$$0 \le \langle M(x_0)u, v\rangle^2 = \langle Lu, Lv\rangle^2 \le \langle Lu, Lu\rangle\langle Lv, Lv\rangle = \langle M(x_0)u, u\rangle\langle M(x_0)v, v\rangle = 0.$$

Thus $\langle M(x_0)u, v\rangle = 0$, and exchanging the roles of u and v we also have $\langle M(x_0)v, u\rangle = 0$. Furthermore,

$$\langle M(x_0)(u+v), (u+v)\rangle = \langle M(x_0)u, u\rangle + \langle M(x_0)u, v\rangle + \langle M(x_0)v, u\rangle + \langle M(x_0)v, v\rangle = 0.$$

Now, we let $\xi \in \mathbb{R}^N$ be a vector such that $\langle M(x_0)\xi, \xi\rangle = 0$ and let us consider the differentiable function

$$f : U \to \mathbb{R}, \quad x \mapsto \langle M(x)\xi, \xi\rangle.$$

In this way ξ is a minimum point for f, which gives $\nabla_{x_0} f = 0$, that is, for each $k \in \{1, \dots, N\}$,

$$0 = \frac{\partial}{\partial x_k} f(x_0) = \frac{\partial}{\partial x_k}\langle M(x_0)\xi, \xi\rangle = \langle \partial_{x_k} M(x_0)\xi, \xi\rangle. \tag{1.13}$$

In conclusion,

$$\begin{aligned}
0 &= \langle \partial_{x_k} M(x_0)(u+v), (u+v)\rangle \\
&= \langle \partial_{x_k} M(x_0)u, u\rangle + \langle \partial_{x_k} M(x_0)u, v\rangle + \langle \partial_{x_k} M(x_0)v, u\rangle + \langle \partial_{x_k} M(x_0)v, v\rangle \\
&= 0 + \langle \partial_{x_k} M(x_0)u, v\rangle + \langle \partial_{x_k} M(x_0)v, u\rangle + 0 \\
&= 2\langle \partial_{x_k} M(x_0)u, v\rangle,
\end{aligned}$$

where we have used (1.13) three times (with $\xi = u + v$, $\xi = u$ and $\xi = v$), and we have also used the symmetry of $\partial_{x_k} M(x_0)$. $\square$

1.3 Hörmander vector fields

In this section, we want provide the definition of Hörmander systems. But firstly see the following forewords.

It is easy to recognize that $\mathcal{X}(D)$, the $\mathbb{R}$-vector space of the C^∞ vector fields on D, is a Lie $\mathbb{R}$-algebra with the product defined by the commutator

$$[X, Y] := XY - YX \quad \text{with } X, Y \in \mathcal{X}(D).$$

(It is easy to prove that this operation is $\mathbb{R}$-bilinear, it gives $[X, X] = 0$ and it satisfies the Jacobi identity and, first of all, it is well-posed by the Schwarz Theorem).

This being said, given $m \in \mathbb{N}$ and

$$S := \{X_1, \dots, X_m\} \subseteq \mathcal{X}(D),$$

we denote by Lie(S) the *Lie algebra generated* by S inside $\mathcal{X}(D)$, that is the intersection of all the sub-algebras of $\mathcal{X}(D)$ containing S. (Lie(S) can be considered also with respect to a generic $S \subseteq \mathcal{X}(D)$, with the same definition.)

Remark 1.18. Some vector fields $X_1, \dots, X_m$ are linear dependent in $\mathcal{X}(D)$, by definition, if there exist $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ (not simultaneously zero) such that $\sum_{i=1}^m \lambda_i X_i$ is the null vector field, which means that its components are all identically zero.

We have not to confuse two different linear-independence concepts: the first is the opposite of the previous clause, that is the vector fields linear independence in $\mathcal{X}(D)$; the second, fixed $x \in D$, is the vectors $X_1(x), \dots, X_m(x)$ being linear independent in the vector space $\mathbb{R}^N$.

On the other hand, we can easily see that the second independence (once it is verified only at $x \in D$) implies the first one, showing that the second dependence (with respect to every $x \in D$) is implied by the first dependence:

$$\sum_{i=1}^m \lambda_i X_i \equiv 0 \text{ over } \Omega \quad \text{implies} \quad \sum_{i=1}^m \lambda_i X_i(x) = 0 \in \mathbb{R}^N \text{ for each } x \in \Omega.$$

The following example ensures that, in general, the vice-versa s false.

Example 1.19. Taken the 1-dimensional vector fields $X_1 := \partial_1$ and $X_2 := x_1\partial_1$, X_1 and X_2 are linearly independent in $\mathcal{X}(\mathbb{R}^1)$ but, for each $x \in \mathbb{R}$, $X_1(x), X_2(x)$ are linearly dependent in $\mathbb{R}$.

Proposition 1.20. *We consider $S \subseteq \mathcal{X}(D)$ and $x \in D$.*

Then we find the dimensional inequalities

$$\dim(\mathrm{Span}(\{Y(x) \in \mathbb{R}^N : Y \in S\})) \le \min\{\dim(\mathrm{Span}(S)), N\} \le \min\{\dim(\mathrm{Lie}(S)), N\},$$

where the first Span *is a vector subspace in* $\mathbb{R}^N$*, the second one and third one meant as dimensions in* $\mathcal{X}(D)$*.*

Proof. Given $x \in D$, $V_x := \mathrm{Span}(\{Y(x) \in \mathbb{R}^N : Y \in S\})$ is a vector subspace in $\mathbb{R}^N$, hence it has finite dimension $m \leq N$ and we can consider a basis $\{v_1, \dots, v_m\}$ in V_x.

By definition of V_x, for every $i \in \{1, \dots, m\}$, v_i is a linear combination of a finite number of vector fields in S evaluated in x. In total, we found a finite number of vector fields in S which, evaluated in x, generate V_x as a vector space. We can extract a basis $\{X_1(x), \dots, X_m(x)\}$, where $X_1, \dots, X_m \in S$.

From Remark 1.18 we gain that $X_1, \dots, X_m$ are linear independent in $\mathcal{X}(D)$, which implies $m \leq \dim(\mathrm{Span}(S))$.

Finally, $S \subseteq \mathrm{Lie}(S)$ gives us $\dim(\mathrm{Span}(S)) \leq \dim(\mathrm{Lie}(S))$. □

Now we are ready for the next definitions.

Definition 1.21 (**Hörmander vector fields system**). *We are given* $S \subseteq \mathcal{X}(D)$*.*

We say that S *is a Hörmander system of vector fields on* D *if*

$$\dim(\{Y(x) \in \mathbb{R}^N : Y \in \mathrm{Lie}(S)\}) = N \text{ for each } x \in D.$$

Obviously $\{Y(x) \in \mathbb{R}^N : Y \in \mathrm{Lie}(S)\}$ *is a vector subspace of* $\mathbb{R}^N$*, hence this previous condition is equivalent to*

$$\{Y(x) \in \mathbb{R}^N : Y \in \mathrm{Lie}(S)\} = \mathbb{R}^N \text{ for each } x \in D.$$

Definition 1.22 (**Hörmander sum of squares**). *If* $S := \{X_0, X_1, \dots, X_m\} \subseteq \mathcal{X}(D)$ *is a Hörmander system on* D*, the differential operator defined by*

$$L_S := X_1^2 + \cdots + X_m^2 + X_0$$

is called a Hörmander sum of squares with drift X_0 *associated to* S*.*

If $X_0 \equiv 0$ *over* D *we refer to Hörmander sum of squares associated to* S*.*

Example 1.23. The Laplace operator, the *Laplacian*, over $\mathbb{R}^N$ is defined by

$$L_\Delta = \left(\frac{\partial}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial}{\partial x_N}\right)^2$$

where the vector fields

$$\Delta := \left\{X_0 = 0, X_1 = \frac{\partial}{\partial x_1}, \dots, X_N = \frac{\partial}{\partial x_N}\right\}$$

form a Hörmander system on $\mathbb{R}^N$. Indeed, denoting $\{e_1, \dots, e_N\}$ the canonical basis in $\mathbb{R}^N$, $X_i \equiv e_i \in \mathrm{Lie}(\Delta)$ for each $i = 1, \dots, N$, hence for every $x \in \mathbb{R}^N$ we have $X_1(x) = e_1, \dots, X_N(x) = e_N \in \{Y(x) \in \mathbb{R}^N : Y \in \mathrm{Lie}(\Delta)\}$.

In other words L_Δ is the Hörmander sum of squares associated to Δ.

Example 1.24. The *Heat operator* in R^{N+1}

$$L_{\mathrm{Heat}} = \left(\frac{\partial}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial}{\partial x_N}\right)^2 - \frac{\partial}{\partial x_{N+1}}$$

is a Hörmander sum of squares with drift, the one related to

$$\mathrm{Heat} := \left\{X_0 = -\frac{\partial}{\partial x_{N+1}}, X_1 = \frac{\partial}{\partial x_1}, \ldots, X_N = \frac{\partial}{\partial x_N}\right\},$$

that is clearly a Hörmander system on R^{N+1}.

We will discuss again about Hörmander vector fields at the end of the following chapter.

Chapter 2

Maximum propagation principle

2.1 Principal vector fields

The primary aim of this chapter is to prove the so called Maximum Propagation Principle, which we now introduce. We first need the concept of principal vector field with respect to the partial differential operator L, the latter being assumed to be as in Section 1.1 (more precisely in Definition 1.1).

Definition 2.1 (Principal vector field). *Let X be a vector field on D. We say that X is a principal vector field for L (on D) if for every $x \in D$ there exists a real number $\lambda(x) > 0$ such that*

$$\langle X(x), \xi \rangle^2 \le \lambda(x) \left\langle A(x)\, \xi, \xi \right\rangle, \qquad \text{for every } \xi \in \mathbb{R}^N. \tag{2.1}$$

If, as usual, $\mathcal{X}(D)$ denotes the set of the smooth vector fields on D, we define

$$\mathrm{Pr}(L) := \{X \in \mathcal{X}(D) \,:\, X \text{ is a principal vector field for } L\}.$$

Observe that (2.1) can be rewritten as

$$\langle X(x), \xi \rangle^2 \le \lambda(x)\, q_L(x, \xi), \qquad \text{for every } \xi \in \mathbb{R}^N.$$

Note that $\xi \mapsto \langle X(x), \xi \rangle^2$ is the quadratic form associated with the matrix $X(x)\,(X(x))^T$; if $X = \sum_{j=1}^N \alpha_j(x) \partial_j$, this matrix is simply $\big(\alpha_i(x)\, \alpha_j(x)\big)_{i,j}$.

With the aid of some elementary linear algebra we can provide a simple characterization of the principality of a vector field.

Proposition 2.2 (Characterizations of Principality). *The vector field X is principal for L (on D) if and only if one of the following equivalent conditions is satisfied for every $x \in D$:*

1. *the vector $X(x)$ belongs to $(\ker A(x))^{\perp}$;*

2. *the vector $X(x)$ is a linear combination of the columns of $A(x)$;*

3. *every isotropic vector for $A(x)$ is orthogonal to $X(x)$:*

$$\mathrm{Isotr}(A(x)) \subseteq (X(x))^{\perp}, \quad \forall\, x \in D.$$

Proof. We refer to the proposition "the vector field X is principal for L on D" as the phrase "0". Let us see the following implications.

$$[\, 1 \leftrightarrow 2 \,]$$

For each x in D, $A(x)$ is symmetric, hence, due to Proposition 1.3, $(\ker A(x))^{\perp} = \mathrm{Im}\, A(x)$.

$$[\, 1 \leftrightarrow 3 \,]$$

For every x in D, due to $A(x)$ being positive semidefinite, again from Proposition 1.3, we find

$$\ker A(x) = \mathrm{Isotr}(A(x)).$$

If $X(x)$ belongs to $(\ker A(x))^{\perp}$ and we take an element in $\mathrm{Isotr}(A(x))$, y, then y belongs to $\ker A(x)$; hence $X(x)$ and y are orthogonal to each other.

As for the opposite direction, for what has been said, it is sufficient to take the orthogonal.

$$[\, 0 \mapsto 1 \,]$$

From the definition of the clause 0, for every $x \in D$ there exists a real number $\lambda(x) > 0$ such that $\langle X(x), \xi\rangle^2 \leq \lambda(x) \left\langle A(x)\, \xi, \xi \right\rangle$, for every $\xi \in \mathbb{R}^N$.

Thus for every $\xi \in \mathbb{R}^N$ if $A(x) = 0$ we have $\langle X(x), \xi\rangle = 0$.

$$[\, 1 \mapsto 0 \,]$$

This is the most complex implication.

Set x in D. If $X(x) = 0$, we find the thesis. We suppose $X(x) \neq 0$. Thus $A(x) \neq 0$, because $X(x) \in (\ker A(x))^{\perp}$. We consider the set

$$S := (\ker A(x))^{\perp} \cap \{v \in \mathbb{R}^N : \|v\| = 1\}$$

and the function

$$f : S \to \mathbb{R}, \quad f(w) = \frac{\langle X(x), w\rangle^2}{\langle A(x)w, w\rangle}.$$

This function is well-posed, indeed for each w in S the denominator of $f(w)$ is different from 0, since $(\ker A(x))^\perp = \operatorname{Im} A(x)$.

From the compactness of S and the continuity of f there exists

$$\lambda(x) := \max_{w \in S} f(w) \geq 0.$$

From this, for each $w \in D$, $\langle X(x), w\rangle^2 \leq \lambda(x)\langle A(x)w, w\rangle$. Then, for every $\xi_1 \in (\ker A(x))^\perp$, there is a $w \in S$ such that $\xi_1 = \|\xi_1\| w$, so that

$$\langle X(x), \xi_1\rangle^2 \leq \lambda(x)\langle A(x)\xi_1, \xi_1\rangle.$$

In conclusion, for each $\xi \in \mathbb{R}^N$ we have $\xi_1 \in (\ker A(x))^\perp$ and $\xi_2 \in \ker A(x)$ such that $\xi = \xi_1 + \xi_2$, that gives

$$\begin{aligned}(\langle X(x), \xi\rangle)^2 &= (\langle X(x), \xi_1\rangle + \langle X(x), \xi_2\rangle)^2 = (\langle X(x), \xi_1\rangle + 0)^2 \\ &\leq \lambda(x)\langle A(x)\xi_1, \xi_1\rangle \leq \lambda(x)\langle A(x)\xi, \xi\rangle,\end{aligned}$$

where the last inequality is obtained from $A(x)$ being positive definite, by which

$$0 \leq \lambda(x)\langle A(x)(\xi - \xi_1), (\xi - \xi_1)\rangle.$$

We have found the desired result with $\lambda(x) \geq 0$, and if the thesis is true with $\lambda(x) = 0$ then it is true with another $\lambda(x)$ strictly larger than zero as well. □

We now see some examples in order to become accustomed to thinking about principality with respect to L.

Example 2.3. The vector field $X = \frac{\partial}{\partial x_{N+1}}$ is not principal for the Heat operator

$$L = \left(\frac{\partial}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial}{\partial x_N}\right)^2 - \frac{\partial}{\partial x_{N+1}},$$

already seen in $\mathbb{R}^{N+1}$ in Example 1.24. Indeed, for each $x \in \mathbb{R}^{N+1}$, $X(x)$ is the $(N+1)$-th vector in the canonical basis in $\mathbb{R}^{N+1}$, whereas the columns in the associated matrix $A(x)$ are the other first N vectors of basis; thus $X(x)$ cannot be a linear combination of the latter, violating the second condition in the previous result.

Example 2.4. We reconsider, from Example 1.2, the operator $L = \sum_{j=1}^m X_j^2 + X_0$, where $X_0, \ldots, X_m$ are C^1 vector fields on D. Then, any vector field of the form

$$X = \sum_{j=1}^m g_j(x)\, X_j,$$

where $g_1, \ldots, g_m$ are real-valued functions on D, is principal for L.

Indeed, owing to the Cauchy-Schwarz inequality in $\mathbb{R}^m$, we have

$$\langle X(x),\xi\rangle^2 = \Big(\sum_{j=1}^m g_j(x)\cdot\langle X_j(x),\xi\rangle\Big)^2 \le \sum_{j=1}^m |g_j(x)|^2\cdot\sum_{j=1}^m\langle X_j(x),\xi\rangle^2 = \widetilde{\lambda}(x)\,q_L(x,\xi),$$

where $\widetilde{\lambda}(x) := \sum_{j=1}^m |g_j(x)|^2$. This gives (2.1) by taking $\lambda(x) := \max\{\widetilde{\lambda}(x),1\}$.

By means of the characterization in Proposition 2.2, this means that, for every $x\in D$, $\sum_{j=1}^m g_j(x)\,X_j(x)$ is a linear combination of the column-vectors of

$$A(x) = S(x)\cdot S(x)^T, \quad \text{where } S(x) := (X_1(x),\dots,X_m(x))^T.$$

Indeed this $A(x)$ is the matrix of the principal part of $\sum_{j=1}^m X_j^2 + X_0$. This fact is not surprising, since, for any $N\times m$ matrix S one has (from Proposition 1.3)

$$\mathrm{Im}(S) = \mathrm{Im}(S\,S^T),$$

where $\mathrm{Im}(C)$ denotes, in general, the span of the column vectors of the matrix C. As a consequence

$$\mathrm{Im}(A(x)) = \mathrm{Im}(S(x)) = \mathrm{span}\{X_1(x),\dots,X_m(x)\}, \quad \text{for each } x\in D.$$

Hence a vector field X is principal for $L=\sum_{j=1}^m X_j^2 + X_0$ if and only if

$$X(x)\in\mathrm{span}\{X_1(x),\dots,X_m(x)\} \quad \text{for each } x\in D.$$

In particular, if the functions $g_1,\dots,g_m$ are all chosen to be 0 except for one of them, which is identically $+1$ or -1, we see that

$$\pm X_1,\dots,\pm X_m \text{ are principal vector fields for } L=\textstyle\sum_{j=1}^m X_j^2 + X_0.$$

Remark 2.5. By arguing via the second condition in Proposition 2.2, one recognizes that $\mathrm{Pr}(L)$ is a module over $C^\infty(D)$, that is

$$X,Y\in\mathrm{Pr}(L), \quad f,g\in C^\infty(D) \quad \text{implies} \quad f\,X+g\,Y\in\mathrm{Pr}(L).$$

The next example shows that $\mathrm{Pr}(L)$ could not be a Lie-subalgebra of $\mathcal{X}(D)$.

Example 2.6. Consider the *Kohn-Laplacian on the Heisenberg group* in $\mathbb{R}^3$:

$$L = X_1^2 + X_2^2, \qquad \text{where } X_1 = \partial_1 + 2\,x_2\,\partial_3,\ \ X_2 = \partial_2 - 2\,x_1\,\partial_3.$$

Since L is a sum of squares, the associated second order matrix $A(x)$ is

$$A(x) = S(x)\cdot S(x)^T, \qquad \text{where } S(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2\,x_2 & -2\,x_1 \end{pmatrix}.$$

Taking into account Example 2.4, we have $\mathrm{Im}(A(x)) = \mathrm{Im}(S(x))$ and X is principal for L if and only if it is of the form

$$X = g_1(x)\,(\partial_1 + 2\,x_2\,\partial_3) + g_2(x)\,(\partial_2 - 2\,x_1\,\partial_3),$$

for some real-valued functions g_1, g_2. Note that, even if X_1, X_2 are principal for L, their commutator $[X_1, X_2] = -4\,\partial_3$ is not principal for L, because $(0, 0, -4)$ does not belong to $\mathrm{Im}(S(x))$, for any $x \in \mathbb{R}^3$.

Thus, $\mathrm{Pr}(L)$ may fail to be a Lie algebra of vector fields.

2.2 Propagation and Strong Maximum Principle

We are ready to state the main result of this chapter. The notations and hypotheses of Section 1.1 are always understood.

Theorem 2.7 (**Maximum Propagation Principle**). *Let L be semielliptic on D. Let $\Omega \subseteq D$ be an open set. For every function $u \in C^2(\Omega)$ satisfying $Lu \geq 0$ and $u \leq 0$ on Ω, the set $F(u) = \{x \in \Omega : u(x) = 0\}$ contains the trajectories, starting at the points of $F(u)$, of the integral curves of any C^1 principal vector field for L.*

In this case we say that the set $F(u)$, which is the set of the maximum points of u (when non-void), propagates along the trajectories of the integral curves of the C^1 principal vector fields for L, whence the name of the theorem, the Maximum Propagation Principle.

We shall prove Theorem 2.7 in the next sections. We observe that, as it will be clear in our proof, we can also consider locally Lipschitz continuous principal vector fields for L instead of C^1 ones, which is assumed for simplicity.

Remark 2.8. Since L is homogeneous, that is $L(1) = 0$, the hypothesis $u \leq 0$ in Theorem 2.27 is irrelevant: it suffices to remove it and to replace $F(u)$ by $F = \{x \in \Omega : u(x) = M\}$, where $M = \sup_\Omega u$. If $M = +\infty$ or, more generally, if F is empty, then there is nothing to prove (since no integral curve can start at a point of an empty set). Otherwise, if there exists $x \in \Omega$ such that $u(x) = \sup_\Omega u$, then x is a maximum point of u and (since L is homogeneous) we can apply Theorem 2.7 to $v = u - M$ (which satisfies $Lv = Lu \geq 0$ and $v \leq 0$).

We therefore obtain that, if it exists, the maximum of an L-subharmonic function $u \in C^2(\Omega)$ propagates along the trajectories of the C^1 principal vector fields for L starting at the points where this maximum is attained.

We also give the following definition, that represents the second main topic of this chapter.

Definition 2.9 (Strong Maximum Principle). *We say that L satisfies the Strong Maximum Principle (SMP, for short) on the connected open set Ω if it satisfies the following condition: for every function $u \in C^2(\Omega)$ such that*

$$Lu \geq 0 \quad \textit{and} \quad u \leq 0 \quad \textit{on } \Omega,$$

the existence of $x_0 \in \Omega$ such that $u(x_0) = 0$ implies that $u \equiv 0$ on the whole of Ω.

More generally, if L satisfies the SMP on the connected open set Ω, and if $u \in C^2(\Omega)$ is such that $Lu \geq 0$ and u attains its maximum in Ω, then u is constant (as it is said in Remark 2.8).

The strict relationship between the Maximum Propagation Theorem and the SMP is clear: indeed, roughly put, if L admits sufficiently many principal vector fields running throughout Ω, then the maximum of an L-subharmonic function propagates everywhere and the SMP holds.

2.2.1 Invariant sets, tangentiality and Nagumo-Bony Theorem

The statement of the Maximum Propagation Theorem 2.7 suggests an independent study of the invariance of a set with respect to the trajectories of a vector field; more precisely, given a closed set F, we aim to give a characterization of the vector fields whose integral curves are constrained to remain in F once they touch F at one point at least. We then begin with the relevant definition.

For the rest of the section, $\Omega \subseteq \mathbb{R}^N$ is a non-empty open set.

Definition 2.10 (Invariant set). *Let X be a vector field on Ω, and let F be a subset of Ω.*

We say that F is positively X-invariant (or positively invariant with respect to X) if, for every integral curve γ of X, $\gamma : [0, T] \to \Omega$ satisfying $\gamma(0) \in F$, we have $\gamma(t) \in F$ for every $t \in [0, T]$.

We say that F is X-invariant (or invariant with respect to X) if it is positively invariant with respect to X and to $-X$.

Remark 2.11. It is easy to recognize that F is X-invariant if and only if, for every integral curve $\gamma : [a, b] \to \Omega$ of X (with $a < 0 < b$) such that $\gamma(0) \in F$, one has $\gamma(t) \in F$ for every $t \in [a, b]$.

The role of 0 is immaterial and, by re-parametrization, one can check that F is X-invariant if and only if, for every integral curve $\gamma : [a,b] \to \Omega$ of X such that $\gamma([a,b]) \cap F \neq \varnothing$, one has $\gamma(t) \in F$ for every $t \in [a,b]$.

Remark 2.12. By means of the notion of invariant set with respect to a vector field, we can restate the thesis of the Maximum Propagation Principle in Theorem 2.7 as follows:

If L is semielliptic on Ω and if $u \in C^2(\Omega)$ is L-subharmonic and nonnegative, then the set $F(u) = \{x \in \Omega : u(x) = 0\}$ (when non-void) is X-invariant, for every C^1 principal vector field X for L.

Due to its importance in maxima propagation, it is now of our concern to find an effective characterization of X-invariance: this will be given by the Nagumo-Bony Theorem (where, more generally, positive invariance is studied). Roughly put, if we try to picture a set F which captures the integral curves of a vector field X, we spontaneously pass through the idea, coming from Differential Geometry, that X is somehow "tangent" to the set F.

Unfortunately, since we want to deal with sets $F(u)$ (as in the Maximum Propagation Principle) which are made of maximum points $x \in \Omega$ of u, we have $\nabla u(x) = 0$; thus, we cannot expect $F(u)$ to be a submanifold of Ω. For this reason we have to consider a milder notion of "tangentiality", which we now introduce.

In what follows, we shall denote by $\|\cdot\|$ the usual Euclidean norm and by $B(z,r)$ the Euclidean ball of centre z and radius r:

$$B(z,r) := \{x \in \mathbb{R}^N : \|x - z\| < r\}.$$

Definition 2.13 (**External orthogonality**). *Let F be a closed subset of Ω and let $y \in \Omega \cap \partial F$. We say that a non-null vector $\nu \in \mathbb{R}^N$ is externally orthogonal to F at y if*

$$\overline{B\big(y + \nu, \|\nu\|\big)} \subseteq (\Omega \setminus F) \cup \{y\}.$$

In this case we shall write $\nu \perp F$ at y. We also let

$$F^* := \Big\{ y \in \Omega \cap \partial F \,\Big|\, \text{there exists } \nu \text{ externally orthogonal to } F \text{ at } y \Big\}.$$

Since we are mainly interested in vectors which are externally orthogonal to F, we shall briefly say that ν is orthogonal to F at y, without the reference to 'externality': this is the reason for the brief notation '$\nu \perp F$ at y'.

Now, we are about to see a technical lemma, that proves the non-emptyness of F^*, when F^* is a closed and proper sub-set in Ω.

Lemma 2.14. *Let F be a closed subset in Ω and let $y \in \Omega \cap \partial F$.*

If $x \in \Omega \setminus F$ and $\|y - x\| = \mathrm{dist}(\partial F, x)$, let $\nu := \frac{1}{2}(x - y)$; then we have $\nu \perp F$ at y.

The statement of the thesis is clear and simple to see in the plan and in the space, but it deserves a proof elsewhere.

Proof. Fixing $z \in F \setminus \{y\}$, we have to prove that $z \notin \overline{B(y + \nu, \|\nu\|)}$.

From the hypothesis $\|x - y\| \leq \|x - z\|$. Then, we observe that $\nu \neq 0$ because $x \notin F$ and $y \in F$ (indeed $\Omega \cap \partial F = \partial_\Omega F \subseteq F$ closed in Ω). And we denote

$$c := y + \nu,\ r := \|\nu\|, \text{ hence } \overline{B(y + \nu, \|\nu\|)} \equiv \overline{B(c, r)}.$$

We proceed by contradiction, supposing $z \in \overline{B(c, r)}$.

We can translate by $-x$ the N-dimensional space, without introducing other notations. In this way $x = 0$, y is a vector with norm $2r$, and c of norm r. Now, we can apply an orthogonal transformation by the following instructions: let $\{e_1, \ldots, e_N\}$ be the canonical basis of $\mathbb{R}^N$; we move y into $2re_2$; if y and z are linear independent, we move a unit vector orthogonal to y in $\mathrm{Span}\{y, z\}$ to e_1, while if y and z are linearly dependent, we jump this instruction. By this transformation (that is in particular an isometry of $\mathbb{R}^N$), we move

$$c \mapsto e_2, \quad \overline{B(c, r)} \mapsto \overline{B(re_2, r)}, \quad z \text{ into } \overline{B(re_2, r)} \cap \mathrm{Span}\{e_1, e_2\}.$$

Summing up, we can visualize the situation in the plan generated by the first two coordinates of the arrival N-dimensional space, where there is a circumference of center re_2, radius r, south-pole the origin, north-pole $2re_2$ and inside the circle there is z.

Euclidean distances are not changed, hence, as initially we had $\|x - y\| \leq \|x - z\|$, now we have $\|z\| = 2r$. At this point it is trivial to see that $z = 2re_2$ (where $2re_2$ is the "old" y), and this is absurd.

If we want to formalize the final step, we can parametrize the circumference and the circle by suitable functions. The function

$$f : [0, 2\pi] \to \mathbb{R}, \quad \theta \mapsto (r\cos(\theta))^2 + (r\sin(\theta) + r)^2,$$

from a brief study, has a unique maximum point corresponding to the value $(2r)^2$.

The function

$$g : [0, r] \times [0, 2\pi] \to \mathbb{R}, \quad (\rho, \theta) \mapsto (\rho\cos(\theta))^2 + (\rho\sin(\theta) + r)^2$$

has maximum $(2r)^2$, reached only at points belonging to $[0, 2\pi] \times \{r\}$. Indeed, for each $(\rho, \theta) \in [0, r) \times [0, \pi]$, is easy to prove that $f(\theta) > g(\rho, \theta)$; if $w \equiv (x_w, y_w) \in \mathbb{R}^2$ with $|x_w|, |y_w| \leq r$, it is straight forward that $\|w\| < 2r$ and, for each $\theta \in [0, 2\pi]$, $g(r, \theta) = f(\theta)$.

These observations about functions f and g give $z = y$. $\square$

Proposition 2.15 (**Non-emptyness of** F^*)**.** *With the notations of the above definition, we remark that, provided* F *is a closed proper subset in* Ω *and* Ω *is connected, then* $F^* \neq \varnothing$*.*

Remark 2.16. Arguing on the connected components of Ω, one can remove the connectedness hypothesis for Ω.

Proof. The boundary of a set, the internal part of a set and the complementary of a set are all considered referring to $\mathbb{R}^N$ and its euclidean topology.

If Ω is connected, then $\partial F \cap \Omega \neq \varnothing$ holds. In fact, provided $\partial F \cap \Omega = \varnothing$, we obtain a contradiction with the hypothesis:

$$\Omega = (\Omega \cap \partial F) \sqcup (\Omega \cap \operatorname{int}(F)) \sqcup (\Omega \cap F^c) = (\Omega \cap \operatorname{int}(F)) \sqcup (\Omega \cap F^c),$$

that gives that Ω is disconnected, because $\Omega \cap \operatorname{int}(F)$ and $\Omega \cap F^c$ are two open sets in Ω, both non-empty if F is a proper subset of Ω and a closed set in $\mathbb{R}^N$.

At this point, we can consider an element $z \in \Omega \cap \partial F$ and, given that Ω is a euclidean open set, there exists a ratio r larger than zero such that $B(z, r) \subseteq \Omega$. Then, there is a point $x_0 \in B(z, r/2)$ that does not belong to F (by a known characterization of the boundary of the complementary of the N-dimensional real sets).

Now, from Weierstrass Theorem, we have a point of minimum

$$y \in \overline{B(x_0, r/2)} \cap \partial F \text{ such that } \|x_0 - y\| = \inf\{\|x_0 - x\| : x \in F\}.$$

Clarification of this last clause is due: by Weierstrass Theorem we have the existence of $y \in F$ such that $\|x_0 - y\| = \inf\{\|x_0 - x\| : x \in F\}$; then, by Proposition 1.13, since $x_0 \notin F$, we derive that $y \in \partial F$ and besides

$$\inf\{\|x_0 - x\| : x \in F\} = \inf\{\|x_0 - x\| : x \in F \cap \overline{B(x_0, r/2)}\}$$

is obviously true.

In conclusion, we simply apply the previous technical Lemma 2.14 to gain $y \in F^*$. $\square$

Since "tangentiality" seems a good notion when X-invariance is concerned, it is convenient to give the following definition.

Definition 2.17 (**Tangent vector field**)**.** *Let* F *be a closed subset in* Ω*. We suppose that* X *is a vector field on* Ω*.*

We say that X *is tangent to* F *if*

$$\langle X(y), \nu \rangle = 0, \quad \textit{for each } y \in F^*, \ \textit{for each } \nu \perp F \textit{ at } y. \tag{2.2}$$

We tacitly mean that condition (2.2) *is fulfilled whenever* $F^* = \varnothing$*. We set*

$$\operatorname{Tg}(F) := \{X \in \mathcal{X}(\Omega) \,:\, X \textit{ is tangent to } F\}. \tag{2.3}$$

Remark 2.18. It is a simple exercise to recognize that $\mathrm{Tg}(F)$ is a vector subspace of $\mathcal{X}(\Omega)$ and, more generally, it is a module over $C^\infty(\Omega)$, that is

$$X, Y \in \mathrm{Tg}(F), \quad f, g \in C^\infty(\Omega) \quad \text{implies} \quad f\,X + g\,Y \in \mathrm{Tg}(F).$$

Later, we shall see that $\mathrm{Tg}(F)$ is a Lie-subalgebra of $\mathcal{X}(\Omega)$.

Next we turn to the characterization of positive X-invariance of a set F.

Remark 2.19. If F is relatively closed in Ω, it is not difficult to verify that the condition

$$\langle X(y), \nu \rangle \le 0, \quad \text{for every } \; y \in F^*, \; \text{ for every } \; \nu \perp F \text{ at } y \tag{2.4}$$

is necessary for the positive X-invariance of F.

Indeed, let $y \in F^*$, $\nu \perp F$ at y and let $\gamma : [0, T] \to \Omega$ be an integral curve of X such that $\gamma(0) = y$. By definition of orthogonality of ν at y we have

$$\overline{B(y + \nu, \|\nu\|)} \subseteq (\Omega \setminus F) \cup \{y\}. \tag{2.5}$$

Then, if F is positively X-invariant (that is $\gamma([0, T]) \subseteq F$), from (2.5) we have

$$\|\gamma(t) - (y + \nu)\|^2 \ge \|\nu\|^2 \quad \text{and} \quad \|\gamma(0) - (y + \nu)\|^2 = \|\nu\|^2,$$

for every $t \in [0, T]$. This means that the C^1 real-valued function

$$t \mapsto \|\gamma(t) - (y + \nu)\|^2$$

has a minimum point at $t = 0$. As a consequence

$$0 \le \frac{\mathrm{d}}{\mathrm{d}\,t}\Big|_{t=0} \Big(\|\gamma(t) - (y + \nu)\|^2 \Big) = 2 \left\langle \dot\gamma(0), \gamma(0) - (y + \nu) \right\rangle = \langle X(y), -\nu \rangle.$$

Hence (2.4) is satisfied.

In this way we have proved a direction of the following Nagumo-Bony theorem.

We are about to show that (2.4) is also sufficient for the positive X-invariance of F. Therefore the statement of the next theorem will be completely proven.

Theorem 2.20 (**Nagumo-Bony**). *Let X be a C^1 vector field on Ω, and suppose that F is a relatively closed subset of Ω.*

Then F is positively X-invariant if and only if

$$\langle X(y), \nu \rangle \le 0, \quad \textit{for every } y \in F^* \textit{ and every } \nu \perp F \textit{ at } y. \tag{2.6}$$

Proof. Due to the previous Remark 2.19, we only need to show the "if" part of the assertion. To this end, let $\gamma : [0,T] \to \Omega$ be an integral curve of X such that $x_0 := \gamma(0)$ belongs to F. For $t \in [0,T]$ we define

$$\delta(t) = \mathrm{dist}(\gamma(t), F) := \inf \big\{ \|\gamma(t) - z\| \,:\, z \in F \big\}.$$

Note that $\delta(0) = 0$ and δ is continuous and non-negative. We need to prove that, under condition (2.6), we have $\delta(t) = 0$ for every $t \in [0,T]$. Due to Lemma 1.12, it is enough to prove that

$$\Psi(t) := \limsup_{h \to 0^-} \frac{\delta(t+h) - \delta(t)}{h} \le L\,\delta(t), \quad \forall\, t \in (0,T], \tag{2.7}$$

for some constant $L > 0$. To this aim, let $V \subset \Omega$ be a bounded neighborhood of x_0 containing $\gamma([0,T])$ (such a neighborhood exists due to the compactness of $\gamma([0,T])$), and let

$$L := \sup_{x,z \in V,\ x \neq z} \frac{\|X(x) - X(z)\|}{\|x - z\|} \tag{2.8}$$

be the Lipschitz constant of X on V. By shrinking T if necessary, we can suppose that $V = B(x_0, r)$ with $B(x_0, 2r) \subseteq \Omega$.[1] We shall prove (2.7) with this choice of L.

If $\delta(t) = 0$, inequality (2.7) is trivial, since $h < 0$ and $\delta(t+h) \ge 0$. Suppose $\delta(t) > 0$ and choose a sequence $h_n < 0$, $h_n \to 0$ such that

$$\Psi(t) = \lim_{n \to \infty} \frac{\delta(t+h_n) - \delta(t)}{h_n}.$$

Let us denote $x := \gamma(t)$ and $x_n := \gamma(t+h_n)$. Since $\gamma([0,T]) \subset B(x_0, r)$ and $B(x_0, 2r) \subseteq \Omega$, for every n there exists a point $z_n \in F \cap B(x_0, r)$ such that

$$\|x_n - z_n\| = \mathrm{dist}(x_n, F) = \mathrm{dist}(\gamma(t+h_n), F) = \delta(t+h_n).$$

Obviously, by choosing a subsequence if necessary, we may suppose that z_n converges to some $z \in F \cap \overline{B(x_0, r)}$. As a consequence, since $x_n \to x$, one has

$$\begin{aligned} \|x - z\| &= \lim_{n\to\infty} \|x_n - z_n\| = \lim_{n\to\infty} \mathrm{dist}(x_n, F) \\ &= \mathrm{dist}(x, F) = \mathrm{dist}(\gamma(t), F) = \delta(t). \end{aligned} \tag{2.9}$$

From Proposition 1.13 we have $z \in \partial F$ and from Lemma 2.14, thanks to $\|x - z\| = \mathrm{dist}(x, F)$, we gain

$$\nu := \tfrac{1}{2}\,(x - z) \perp F \ \text{ at } z. \tag{2.10}$$

[1] Otherwise we apply this same argument on a partition of $[0,T]$ into small segments, say $[0,T_1]$, $[T_1,T_2]$, $\ldots, [T_n,T]$, proving that $\delta \equiv 0$ on $[0,T_1]$, then on $[T_1,T_2]$, and so forth.

Then

$$\begin{aligned} \delta(t+h_n)-\delta(t) &= \|x_n-z_n\|-\|x-z\| \geq \|x_n-z_n\|-\|x-z_n\| \\ &\geq -\frac{\langle x-z_n, x-x_n\rangle}{\|x-z_n\|} = \frac{\langle x_n-x, x-z_n\rangle}{\|x-z_n\|}. \end{aligned} \tag{2.11}$$

In the first inequality we used the fact that $\|x-z\| = \mathrm{dist}(x,F)$ and $z_n \in F$; the second inequality, due to $x \neq z_n$ (having $\delta(t) > 0$), if $x \neq x_n$, is a consequence of the following estimate (obtained by taking $a = x$, $b = z_n$, $c = x_n$ in Lemma 1.14):

$$\|x-z_n\| \leq \|x_n-z_n\| + \frac{\langle z_n-x, x_n-x\rangle}{\|z_n-x\|\cdot\|x_n-x\|}\,\|x-x_n\|,$$

while, if $x = x_n$, the inequality is simply $0 \geq 0$. Hence (taking into account that $h_n < 0$ and the continuity of the functions under consideration), we have the following calculation:

$$\begin{aligned} \Psi(t) &= \lim_{n\to\infty} \frac{\delta(t+h_n)-\delta(t)}{h_n} \overset{(2.11)}{\leq} \lim_{n\to\infty} \left\langle \frac{x_n-x}{h_n}, \frac{x-z_n}{\|x-z_n\|} \right\rangle \\ &= \lim_{n\to\infty} \left\langle \frac{\gamma(t+h_n)-\gamma(t)}{h_n}, \frac{x-z_n}{\|x-z_n\|} \right\rangle = \left\langle \dot\gamma(t), \frac{x-z}{\|x-z\|} \right\rangle \\ &\overset{(2.10)}{=} \frac{2}{\|x-z\|}\,\langle X(\gamma(t)),\nu\rangle = \frac{2}{\|x-z\|}\,\langle X(x),\nu\rangle \\ &= \frac{2}{\|x-z\|}\Big\{\langle X(x)-X(z),\nu\rangle + \langle X(z),\nu\rangle\Big\} \leq 2\,\frac{\langle X(x)-X(z),\nu\rangle}{\|x-z\|}. \end{aligned}$$

In the last inequality we used the hypothesis (2.6) applied when y is equal to z (giving $\langle X(z),\nu\rangle \leq 0$). This produces the estimate

$$\Psi(t) \leq 2\,\frac{\langle X(x)-X(z),\nu\rangle}{\|x-z\|}.$$

From the Cauchy-Schwarz inequality, together with $\|\nu\| = \|x-z\|/2$, we get that

$$\Psi(t) \leq \|X(x)-X(z)\|.$$

From the definition (2.8) of L, we finally get

$$\Psi(t) \leq L\,\|x-z\| \overset{(2.9)}{=} L\,\delta(t).$$

This completes the proof of (2.7). □

Remark 2.21. Notice that in the previous demonstration it is sufficient to consider a locally Lipschitz continuous vector field X over Ω instead of a C^1 one, which is assumed for simplicity.

With Definition 2.17 at hand, the Nagumo-Bony Theorem 2.20 provides us with the following crucial result.

Corollary 2.22 (Equivalence of invariance and tangentiality). *Let X be a C^1 vector field on Ω. Suppose that F is a relatively closed subset of Ω.*

Then, F is X-invariant if and only if X is tangent to F.

Proof. By Definition 2.10, F is X-invariant if and only if F is positively invariant with respect to X and $-X$. By the Nagumo-Bony Theorem 2.20, this is equivalent to

$$\langle \pm X(y), \nu \rangle \leq 0, \quad \forall\, y \in F^*, \ \forall\, \nu \perp F \text{ at } y.$$

That is (2.2) holds true, which means that X is tangent to F. □

Remark 2.23. If F is a relatively closed subset of Ω, we let

$$\mathrm{Inv}(F) := \{X \in \mathcal{X}(\Omega) \,:\, F \text{ is } X\text{-invariant}\}.$$

We deduce from Corollary 2.22 that (see the notation in (2.3)) $\mathrm{Tg}(F) = \mathrm{Inv}(F)$. In Section 2.2.3, we shall discover that $\mathrm{Tg}(F) = \mathrm{Inv}(F)$ is not only a module over $C^\infty(\Omega)$, but it is also a Lie sub-algebra of $\mathcal{X}(\Omega)$.

2.2.2 Hopf Lemma

We now see another result that will be required in the proof of the Maximum Propagation Principle Theorem 2.7. It is a classic result about elliptic operators, generalised to an operator L as in the hypothesis of the recalled theorem.

Lemma 2.24 (Hopf Lemma). *Let L be a semielliptic operator over D and $\Omega \subseteq D$ be a connected open set. We consider $u \in C^2(\Omega)$ such that $Lu \geq 0$ and $u \leq 0$ over Ω, we place $F(u) := \{x \in \Omega : u(x) = 0\}$ and we suppose that $F(u)$ is a proper subset of Ω.*

Then, for each $y \in F(u)^*$ and for each $\nu \perp F$ at y

$$\langle A(y)\nu, \nu \rangle = 0, \ \textit{that is } \nu \in \mathrm{Isotr}(A(y)).$$

Remark 2.25. If X is a principal vector field for L (as in Definition 2.1) and the hypothesis of the previous Lemma holds, then X is tangent to $F(u)$.

Indeed, for each $y \in F(u)^*$ and for each $\nu \perp F$

$$0 \leq \langle X(y), \nu \rangle^2 \leq \lambda(y) \langle A(y)\nu, \nu \rangle = 0,$$

by definition of the principality (with a certain $\lambda(y) > 0$), using also the thesis of Hopf Lemma.

Another possible way is as follows:

given $y \in F(u)^*$ and $\nu \perp F$, by Proposition 1.3 with $A(y)$ positive semidefinite, $\mathrm{Isotr}(A(y)) = \ker A(y)$, hence Hopf thesis says $\nu \in \ker A(y)$; on the other hand, if X is principal for L, then Proposition 2.2 gives $X(y) \in (\ker A(y))^\perp$, from which we derive $\langle X(y), \nu \rangle = 0$.

Lemma 2.26 (**The Hopf function**). *Let G be an operator of the form*

$$G := \sum_{i,j=1}^N \alpha_{i,j}(x) \frac{\partial^2}{\partial x_i \, \partial x_j} + \sum_{i=1}^N \beta_i(x) \frac{\partial}{\partial x_i} + \gamma,$$

where $\alpha_{i,j} = \alpha_{j,i}$, β_i and γ are continuous functions on an open set $\Omega \subseteq \mathbb{R}^N$. We suppose there exists $y \in \Omega$ and $\nu \in \mathbb{R}^N \setminus \{0\}$ such that

$$\sum_{i,j=1}^N \alpha_{i,j}(y)\nu_i\nu_j > 0.$$

Then, letting $B := B(y+\nu, \|\nu\|)$, there exists a function $h \in C^\infty(\mathbb{R}^N)$, called "Hopf function for G in y along ν", with the following properties:

(1) $h > 0$ on B, $h = 0$ on ∂B and $h < 0$ on $\mathbb{R}^N \setminus \overline{B}$;

(2) there exists $\delta > 0$ (depending on ν and the coefficients of G in a neighborhood of y) such that $Gh > 0$ on $B(y,\delta)$ and $\overline{B(y,\delta)} \subseteq \Omega$.

Proof. We place $z := y + \nu$ and $r := \|\nu\|$. Now we define the desired smooth function $h : \mathbb{R}^N \to \mathbb{R}$ in this way

$$x \mapsto h(x) := \exp(-\lambda\|x-(y+\nu)\|^2) - \exp(-\lambda\|\nu\|^2),$$

where $\lambda > 0$ will be suitably chosen, depending on ν, $\alpha_{i,j}$, $\beta_{i,j}$, later on.

Notice that the point (1) follows immediately from the definition of h.

We set $H(y) := (\alpha_{i,j}(y))_{i,j}$ and the N-vector $\beta(y) := (\beta_i(y))_i$, by a direct calculation we gain

$$Gh(y) = 4\lambda^2 \exp(-\lambda r^2) \cdot \left\{ \langle H(y)\nu, \nu\rangle + \frac{1}{2\lambda}((\langle \beta(y)\nu, \nu\rangle - \mathrm{trace}(H(y)) \right\}.$$

The hypothesis $\langle H(y)\nu, \nu \rangle = \sum_{i,j=1}^N \alpha_{i,j}(y)\nu_i\nu_j > 0$ ensures that, with a λ sufficiently large, $Gh(y) > 0$. And from this, by continuity, we have (2). □

Therefore, we are ready to prove the Hopf Lemma.

Proof of Hopf lemma. To simplify the notation we let $F := F(u)$. Let $y \in F^*$ and $\nu \perp F$ at y; hence $\overline{B(y+\nu, \|\nu\|)} \subseteq (\Omega \setminus F) \cup \{y\}$. We set $B := B(y+\nu, \|\nu\|)$.

We suppose, by contradiction, the opposite of the thesis, that is, remembering that $A(y)$ is positive semi-definite, $\langle A(y)\nu, \nu\rangle > 0$. With this, by the previous Lemma 2.26 applied to the operator L, there is a Hopf function h for L in y along ν. As in the statement of the precedent Lemma, we have a $\delta > 0$ such that:

(1) $h > 0$ on B, $h = 0$ on ∂B and $h < 0$ on $\mathbb{R}^N \setminus \overline{B}$;

(2) there exists $\delta > 0$ (depending on ν and the coefficients of G in a neighborhood of y) such that $Gh > 0$ on $B(y, \delta)$ and $\overline{B(y,\delta)} \subseteq \Omega$.

We place $z := y + \nu$, $r := \|\nu\|$ and $V := B(y, \delta)$.

We divide the boundary of V in two disjoint subsets:

$$\Gamma_1 := \partial V \setminus \overline{B(z,r)}, \quad \Gamma_2 := \partial V \cap \overline{B(z,r)}.$$

In this way we find $\Gamma_2 \subseteq \Omega \setminus F$ and Γ_2 is a compact set in $\mathbb{R}^N$. Being $u < 0$ over $\Omega \setminus F$, there is a certain $M := \max_{\Gamma_2} u < 0$. In particular there exists $\epsilon > 0$ such that $u + \epsilon h < 0$ over Γ_2, as the continuous function h is bounded over the compact set Γ_2. On the other hand, being $h < 0$ over the complementary set of $\overline{B(z,r)}$ and $u \le 0$ on Ω, we also have $u + \epsilon h < 0$ on Γ_1. Letting $u_\epsilon := u + \epsilon h$, it holds that

$$u_\epsilon < 0 \quad \text{on} \quad \partial V = \Gamma_1 \cup \Gamma_2.$$

Furthermore, $Lu \ge 0$ on Ω by hypothesis and $Lh > 0$ on V by construction; hence

$$L(u_\epsilon) = Lu + \epsilon Lh \ge \epsilon Lh > 0 \quad \text{on} \quad V.$$

In conclusion, due to $u(y) = 0$ (from $y \in F^* \subseteq F$) and $h(y) = 0$ ($y \in \partial B(z,r)$), we gain

$$u_\epsilon(y) = 0.$$

Summing up, we have the three facts

$$u_\epsilon < 0 \text{ on } \partial V = \Gamma_1 \cup \Gamma_2, \quad L(u_\epsilon) > 0 \text{ on } V \quad \text{and} \quad u_\epsilon(y) = 0,$$

which are contradictory with each other. Indeed, we can consider $\alpha := \max_{\overline{V}} u_\epsilon$ and from the third fact $\alpha \ge 0$, and from the first fact this maximum is achieved at an inner point of $\overline{V}$, that is in $V \subseteq \Omega$; but u_ϵ is strictly subharmonic referring to L over V (by the second fact), and this implies, through Proposition 1.5, that u_ϵ can not have a local maximum point in V. This contradiction puts an end to the proof. □

2.2.3 The proof of the Maximum Propagation Principle

The Nagumo-Bony Theorem and the Hopf Lemma are the two fundamental elements for the demonstration of the Maximum Propagation Principle, of which we give again the statement.

Theorem 2.27 (**Maximum Propagation Principle**). *Let L be semielliptic on D. Let $\Omega \subseteq D$ be an open set. For every function $u \in C^2(\Omega)$ satisfying $Lu \geq 0$ and $u \leq 0$ on Ω, the set $F(u) = \{x \in \Omega : u(x) = 0\}$ contains the trajectories, starting at the points of $F(u)$, of the integral curves of any C^1 principal vector field for L over Ω.*

Proof. Let X be a C^1 principal vector field for L. We suppose that $F(u)$ is non-empty, otherwise there is nothing to say. Furthermore, we can also suppose that Ω is path-connected, by possibly arguing over each path connected component of Ω (from which the integral curves of X cannot escape). Due to Definition 2.10, we want to prove that $F(u)$ is an X-invariant set.

If $F(u) = \Omega$ there is nothing to prove; hence we suppose that $F(u)$ a proper set of Ω, so that $F(u)$ is relatively closed in Ω.

Then, we have X tangent to $F(u)$ due to Remark 2.25 (that is a Hopf Lemma consequence), X being a principal vector field for L.

In conclusion, from Nagumo-Bony Theorem, being $F(u)$ is a relatively closed in Ω, we have the equivalence between tangentiality and invariance, hence $F(u)$ is X-invariant. □

Proposition 2.28. *Taking F a non-empty closed set of $\mathbb{R}^N$ contained in Ω and two C^2 vector fields X and Y on Ω, we suppose that F is invariant with respect to both X and Y (or equivalently X and Y are both tangent to F).*

Then F is also $[X, Y]$-invariant (or equivalently $[X, Y]$ is tangent to F).

As a natural consequence we find that

$$\{X \in \mathcal{X}(\Omega) : F \textit{ is } X\textit{-invariant}\} = \mathrm{Inv}(F) = \mathrm{Tg}(F) = \{X \in \mathcal{X}(\Omega) : X \textit{ is tangent to } F\}$$

is Lie sub-algebra of $\mathcal{X}(\Omega)$.

Proof. We remember that, from the Nagumo-Bony Theorem, we have the equivalence between invariance and tangentiality. Furthermore we have already observed that $\mathrm{Tg}(F)$ is a module over $C^\infty(\Omega)$.

It remains to show that $\mathrm{Tg}(F)$ is closed by commutation.

Letting F, X and Y as in hypothesis, $[X, Y]$ is a C^1 vector field on Ω and we can prove that it is tangent to F. We consider $y \in F^*$ and $\nu \perp F$ in y. We know, by Remark 1.11, that there exists $T > 0$ such that we can well-pose the function

$$\Gamma : [0, T] \to \Omega, \quad t \mapsto \Gamma(t) := \Psi^Y_{-\sqrt{t}} \circ \Psi^X_{-\sqrt{t}} \circ \Psi^Y_{\sqrt{t}} \circ \Psi^X_{\sqrt{t}}(y).$$

Through Proposition 1.10, we are able to say

$$\lim_{t\to 0^+}\frac{\Gamma(t)-y}{t}=([X,Y]I)(y).$$

At the same time, being F invariant with respect to both X and Y, F contains all integral curves of X and Y starting at a point in F. Hence $\Gamma([0,T])\subseteq F$.

Then, seen that $\Gamma(0)=y$ and $\overline{\mathrm{B}(y+\nu,\|\nu\|)}\subseteq(\Omega\setminus F)\cup y$, we gain

$$\|\Gamma(t)-(y+\nu)\|^2\geq\|\nu\|^2=\|\Gamma(0)-(y+\nu)\|^2,\quad \text{for each } t\in[0,T].$$

By these two facts

$$0\leq\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}\left(\|\Gamma(t)-(y+\nu)\|^2\right)=2\langle\dot\Gamma(0),\Gamma(0)-(\nu+y)\rangle=-2\langle[X,Y](y),\nu\rangle.$$

This implies

$$\langle[X,Y](y),\nu\rangle\leq 0,\quad \text{for each } y\in F^* \text{ and } \nu\perp F \text{ in } y.$$

Exchanging the roles of X and Y, we also obtain $\langle[Y,X](y),\nu\rangle\leq 0$, and the commutator anti-symmetry gives $\langle[Y,X](y),\nu\rangle=-\langle[X,Y](y),\nu\rangle$.

Hence, summing up

$$\langle[X,Y](y),\nu\rangle=0,\quad \text{for each } y\in F^* \text{ and } \nu\perp F \text{ in } y.$$

Therefore $[X,Y]$ is tangent to F and we have put an end to this proof. □

2.3 Strong Maximum Principle for Hörmander operators

We remember the following definition, already given in a previous section.

Definition 2.29 (Strong Maximum Principle). *We say that L satisfies the Strong Maximum Principle (SMP, for short) on the connected open set Ω if it satisfies the following condition: for every function $u\in C^2(\Omega)$ such that*

$$Lu\geq 0\quad \textit{and}\quad u\leq 0\quad \textit{on } \Omega,$$

the existence of $x_0\in\Omega$ such that $u(x_0)=0$ implies that $u\equiv 0$ on the whole of Ω.

Now, we demonstrate that Hörmander sum of squares operators satisfy the previous definition.

Theorem 2.30. *Let $S := \{X_1, \dots, X_m\}$ be a Hörmander system on an open set $D \subseteq \mathbb{R}^N$.*

Then, the associated Hörmander sum of squares $L_S = \sum_{i=1}^m X_i^2$ satisfies the SMP on each connected open set $\Omega \subseteq D$.

Proof. Let $u \in C^2(\Omega)$ such that $Lu \geq 0$ and $u \leq 0$ on Ω, where Ω is a connected open set inside D. Furthermore, we suppose there is a point $x_0 \in \Omega$ such that $u(x_0) = 0$ and we place $F(u) := \{x \in \Omega : u(x) = 0\}$, so that $x_0 \in F(u)$.

Hence the thesis becomes $F(u) = \Omega$.

Supposing by contradiction that $F(u)$ is a proper subset of Ω and remembering that $X_1, \dots, X_m$ are principle vector fields for such a operator L_S, from the Hopf Lemma, we gain that $X_1, \dots, X_m$ are all tangent to $F(u)$.

Furthermore, by Proposition 2.28, we find that $\mathrm{Tg}(F(u))$ is a Lie algebra.

From these two clauses, we have

$$\mathrm{Lie}(S) \subseteq \mathrm{Tg}(F(u)).$$

If $F(u)$ is a proper subset of Ω we know there exists a $y \in F(u)^*$ and a vector $\nu \perp F(u)$ at y, thanks to Proposition 2.15.

Being S a Hörmander system there exist $Y_1, \dots, Y_N \in \mathrm{Lie}(S)$ such that $\{Y_1(y), \dots, Y_N(y)\}$ is a basis of R^N as a vector space.

From what we said $Y_1, \dots, Y_N$ are tangent to $F(u)$, thus by definition we have

$$\langle Y_k(y), \nu \rangle = 0 \text{ for each } k = 1, \dots, N.$$

By the fact that $Y_1(y), \dots, Y_N(y)$ is a basis of R^N, we obtain $\nu = 0 \in \mathbb{R}^N$, which is in contrast with the definition of external normal (see Definition 2.13).

Once it is obtained this contradiction, $F(u)$ cannot be a proper set of Ω, that is $F(u) = \Omega$. $\square$

Example 2.31. This example serves to remark that the previous result does not work substituting the Hörmander sum of squares with a Hörmander sum of squares with drift.

We take the Heat operator over $\mathbb{R}^{N+1}$

$$L_{\mathrm{Heat}} = \left(\frac{\partial}{\partial x_1}\right)^2 + \cdots + \left(\frac{\partial}{\partial x_N}\right)^2 - \frac{\partial}{\partial x_{N+1}}.$$

The so called "*fundamental solution*" of L_{Heat} (letting $x \equiv (x_1, \dots, x_N)$)

$$\Phi(x, x_{N+1}) := \begin{cases} 0, \text{ if } x_{N+1} \leq 0 \text{ and } (x, x_{N+1}) \neq (0,0) \\ -\frac{1}{(4\pi \ x_{N+1})^{N/2}} \exp\left(-\frac{\|x\|^2}{4 \ x_{N+1}}\right), \text{ if } x_{N+1} > 0 \end{cases}$$

is “L_{Heat}-harmonic” over the connected open set $\Omega := \mathbb{R}^{N+1} \setminus \{0\}$, that is

$$L_{\text{Heat}}\Phi(x) = 0 \text{ for each } x \in \Omega,$$

that obviously implies the L_{Heat}-subharmonicity of Φ on Ω.

Furthermore Φ is ≤ 0 over Ω, but it is not identically zero over Ω.

Hence L_{Heat} does not respect the SMP over Ω.

Notice that in this case

$$\text{Heat} = \left\{ X_0 = -\frac{\partial}{\partial x_{N+1}}, X_1 = \frac{\partial}{\partial x_1}, \ldots, X_N = \frac{\partial}{\partial x_N} \right\}$$

is a Hörmander system in $\mathbb{R}^{N+1}$, but

$$\left\{ X_1 = \frac{\partial}{\partial x_1}, \ldots, X_N = \frac{\partial}{\partial x_N} \right\}$$

is not a Hörmander system in $\mathbb{R}^{N+1}$, indeed:

$$\text{Lie}\left(\left\{ X_1 = \frac{\partial}{\partial x_1}, \ldots, X_N = \frac{\partial}{\partial x_N} \right\} \right) = \left\{ fX_i - gX_j : f, g \in C^\infty(\Omega), i, j = 1, \ldots, N \right\};$$

for each X in this latter algebra and for each $x \in \mathbb{R}^{N+1}$ the $(N+1)$-th component of $X(x) \in \mathbb{R}^{N+1}$ is zero.

Chapter 3

Propagation along the drift

3.1 Maximum Propagation Principle along the drift

We go further in our study of the Maximum Propagation Principle by turning our attention to the role of the so called drift vector field.

The motivation is that, in the previous chapters, in dealing with operators of the form $L = \sum_{i=1}^{N} X_i^2 + X_0$, nothing has been said about the maximum propagation along the integral curves of the drift term X_0. This is not an oversight: the issue is that, in general, X_0 is not a principal vector field for L. The explicit example of the Heat operator $L_{\text{Heat}} = (\partial x_1)^2 + \cdots + (\partial x_N)^2 - \partial x_{N+1}$ and its fundamental solution Φ (see Example 2.31) proves that not only misses to be principal, but neither can we expect (two-sided) propagation of the maximum of an L_{Heat}-subharmonic function along the drift: indeed Φ is null in the half-space $\{x_{N+1} < 0\}$ but elsewhere in $\mathbb{R}^{N+1}$ it is strictly negative.

Nonetheless, a redeeming fact will be proved in this chapter: despite the lack of X_0-invariance, we still have the positive X_0-invariance of the maximum points $F(u)$ of an L-subharmonic function u. The proof of this fact is extremely delicate: we have to proceed through several steps to describe it.

Recovering the differential operator L from the first Section 1.1, we can associate, supposing them to be C^2, the following vector fields over D:

$$X_i := \sum_{j=1}^{N} a_{i,j}(x) \frac{\partial}{\partial x_j}, \quad \text{for } i = 1, \ldots, N,$$
$$X_0 := \sum_{j=1}^{N} \left(b_j(x) - \sum_{i=1}^{N} \frac{\partial a_{i,j}}{\partial x_i}(x) \right) \frac{\partial}{\partial x_j}.$$

By means of these vector fields we can rewrite L as:

$$\begin{aligned} L &= \sum_{i=1}^{N} \frac{\partial}{\partial x_i}\left(\sum_{j=1}^{N} a_{i,j}(x)\frac{\partial}{\partial x_j}\right) + \sum_{j=1}^{N}\left(b_j(x) - \sum_{i=1}^{N}\frac{\partial a_{i,j}}{\partial x_i}(x)\right)\frac{\partial}{\partial x_j} \\ &= \sum_{i=1}^{N}\frac{\partial}{\partial x_i}\Big(X_i\Big) + X_0. \end{aligned} \tag{3.1}$$

X_0 is called *the drift* of L.

Remark 3.1. We observe that $X_1, \dots, X_N$ are principal vector fields over D with respect to L; indeed $X_1(x), \dots, X_N(x)$ are exactly the columns of $A(x)$ for each $x \in D$ (and we have Proposition 2.2).

The main aim of this chapter is to demonstrate the following theorem:

Theorem 3.2 (**Maximum Propagation Principle along the drift**). *Writing L as in* (3.1)*, where X_0 is the drift of L, let $u \in C^2(\Omega)$ be an L-subharmonic function on a non-empty open set $\Omega \subseteq D$. If we know that the set of the maximum points $F(u)$ of u is non-empty, then $F(u)$ is positively X_0-invariant.*

To show this result we need the next proposition.

Proposition 3.3. *Writing the operator L as in* (3.1)*, where X_0 is the drift of L, let $u \in C^2(\Omega)$ be an L-subharmonic function on an open set $\Omega \subseteq D$. If u attains its maximum at a point $x_0 \in \Omega$ and there is a unit vector $\nu \in \mathbb{R}^N$ such that*

$$\langle X_0(x_0), \nu\rangle > 0,$$

then for every $\rho > 0$ the function u attain its maximum not only at x_0 but also at some point of the set $\Omega \cap B(x_0 + \rho\nu, \rho)$.

The latter result has a definitely complex proof, which needs some intermediate results. But in the meantime we can show in which way Proposition 3.3 implies Theorem 3.2.

Proof of Maximum Propagation Principle along the drift. We can suppose $\Omega \neq F(u)$, otherwise there is nothing to prove. Thus, from the Nagumo-Bony Theorem 2.20, we know that $F(u)$ is positively X_0-invariant if and only if

$$\langle X_0(y), \xi\rangle \le 0, \quad \text{for } y \in F(u)^* \text{ and for each } \xi \perp F(u) \text{ at } y.$$

We argue by contradiction assuming that there exist $y \in F(u)^*$ and $\xi \perp F(u)$ at y such that $\langle X_0(y), \xi\rangle > 0$. This allows us to apply Proposition 3.3 with $x_0 = y$, $\nu = \xi/\|\xi\|$ and

$\rho = \|\xi\|$, gaining that u attains its maximum at some point $\tilde{y} \in \Omega \cap B(y + \xi, \|\xi\|)$; this ensures that

$$F(u) \cap B(y + \xi, \|\xi\|) \neq \varnothing.$$

However ξ is externally orthogonal to $F(u)$ at y, so $F(u) \cap B(y + \xi, \|\xi\|) = \varnothing$. We have the desired contradiction. □

3.1.1 The full demonstration

We are now ready to approach the two preliminary results, proceeding step by step to our final aim: the proof of Proposition 3.3.

Proposition 3.4. *Let $u \in C^2(\Omega)$ be an L-subharmonic function on an open set Ω in $\mathbb{R}^N$. Let us assume that u attains its maximum at a point $x_0 \in \Omega$ and that there exists a function $f \in C^2(\Omega)$ satisying the following properties:*

(1) $f(x_0) = 0$ *and* $\nabla_{x_0} f \neq 0$*;*

(2) $Lf(x_0) > 0$.

Then, there is a neighborhood U_0 of x_0 such that, for every open neighborhood U of x_0 with closure contained in U_0, u attains its maximum not only at x_0 but also st some point of the set

$$\{x \in \partial U : f(x) > 0\}.$$

Proof. We consider the C^2-function

$$F : \Omega \to \mathbb{R}, \quad F(x) := f(x) - c\|x - x_0\|^2,$$

where $c > 0$ will be chosen in a moment. F satisfies three properties (obtainable with simple calculations):

1. $F(x_0) = 0$;
2. $\nabla_{x_0} F = \nabla_{x_0} f$;
3. $LF(x_0) = Lf(x_0) - 2c \operatorname{trace}(A(x_0))$.

From the first property of f and the second property of F, we can consider a ratio $r > 0$ such that, set $U_0 := B(x_0, r)$, $\overline{U_0} \subset \Omega$ and $\nabla_x F \neq 0$ for each $x \in \overline{U_0}$. By the second property of f and the third property of F, we can choose such a small $c > 0$ that $LF(x) > 0$ for every $x \in \overline{U_0}$.

Simply by definition of F we have that

$$\{x \in \Omega \setminus \{x_0\} : F(x) \geq 0\} = \{x \in \Omega : f(x) > 0\}.$$

Due to this fact, fixing an open neighborhood U of x_0 with $\overline{U} \subseteq U_0$, it is sufficient that we prove that u attains its maximum at some point of

$$\Sigma := \{x \in \partial U : F(x) \geq 0\}.$$

We argue by contradiction supposing that $\sup_\Sigma u < u(x_0)$. Thus it is possible to choose $\epsilon > 0$ such that

$$0 < \epsilon < \frac{u(x_0) - \sup_\Sigma u}{\sup_\Sigma F}, \tag{3.2}$$

also because $\sup_\Sigma F > 0$, since for every $x \in \overline{U_0}$ we have $\nabla_x F \neq 0$ and $F(x) \geq 0$.

We define $v := u + \epsilon F$ over Ω and we have:

i. $Lv = Lu + \epsilon LF \geq \epsilon LF > 0$ on U;

ii. $v = u + \epsilon F \overset{(3.2)}{<} u + u(x_0) - \sup_\Sigma u(x_0) \leq u(x_0)$ on Σ.

Having $v(x_0) = u(x_0)$ and $v \in C^2(U)$, it results that v is a strictly L-subharmonic function over U possessing a maximum point inside U, which is a contradiction by Proposition 1.5. □

Corollary 3.5. *Let $u \in C^2(\Omega)$ be an L-subharmonic function on an open set Ω in $\mathbb{R}^N$. Let us assume that u attains its maximum at a point $x_0 \in \Omega$ and that there exists a unit vector $\nu \in \mathbb{R}^N$ such that $\langle A(x_0)\nu, \nu\rangle > 0$. Then, for every $\rho > 0$, the function u attains its maximum not only at X_0 but also at some point of the set $\Omega \cap B(x_0 + \rho\nu, \rho)$.*

Proof. Let $\rho > 0$. As in Proposition 2.26, we consider the Hopf function $h \in C^\infty(\mathbb{R}^N)$ for L in the point x_0 along the vector $\rho\nu$. From the known properties of this kind of function we have $h(x_0) = 0$ and $Lh(x_0) > 0$. And, through a direct calculation, one can also show that $\frac{\partial}{\partial x_j} h(x_0) \neq 0$ for each $j = 1, \dots, N$, that is $\nabla_{x_0} h \neq 0$.

We are in a position to apply Proposition 3.4 to say that: there is a neighborhood U_0 of x_0 such that, for every open neighborhood U of x_0 with closure contained in U_0, u attains its maximum not only at x_0 but also at some point of the set $\{x \in \partial U : h(x) > 0\}$. But $\{x \in \mathbb{R}^N : h(x) > 0\} = B(x_0 + \rho\nu, \rho)$, then

$$\{x \in \partial U : h(x) > 0\} \subseteq \Omega \cap B(x_0 + \rho\nu, \rho),$$

and this is conclusive. □

To continue we have to specify how second order differential operators can be reparameterized through diffeomorphisms.

Remark 3.6. Let G be an operator of the form

$$G := \sum_{i,j=1}^{N} \alpha_{i,j}(x)\, \frac{\partial^2}{\partial x_i\, \partial x_j} + \sum_{i=1}^{N} \beta_i(x)\, \frac{\partial}{\partial\, x_i} + \gamma,$$

where $\alpha_{i,j} = \alpha_{j,i}$, β_i and γ are continuous function on a non-empty open set $D \subseteq \mathbb{R}^N$. Let $\psi : D \to \widetilde{D}$ be a C^2-diffeomorphism onto the open set $\widetilde{D} \subseteq \mathbb{R}^N$ and let $\mathrm{J}_x\psi$ be the Jacobian matrix of ψ at the point $x \in D$.

We also define, for any $x \in D$, the $N \times N$ real matrix $H(x) := (\alpha_{i,j}(x))_{i,j}$ and the vector $\beta(y) := (\beta_i(x))_i \in \mathbb{R}^N$.

We define a new operator $\widetilde{G}$ on $\widetilde{D}$ by setting

$$(\widetilde{G}u)(y) := (G(u \circ \psi))(\psi^{-1}(y)), \quad \text{with } y \in \widetilde{D} \text{ and } u \in C^2(\widetilde{D}).$$

By direct calculations one can prove that

$$\widetilde{G} = \sum_{i,j=1}^{N} \widetilde{\alpha}_{i,j}(y)\, \frac{\partial^2}{\partial y_i\, \partial y_j} + \sum_{i=1}^{N} \widetilde{\beta}_i(y)\, \frac{\partial}{\partial\, y_i} + \widetilde{\gamma},$$

where, for any $y \in \tilde{D}$,

$$\begin{aligned}
\widetilde{H}(y) &:= (\widetilde{\alpha}_{i,j}(y))_{i,j} = (\mathrm{J}_{\psi^{-1}(y)}\psi) \cdot H(\psi^{-1}(y)) \cdot (\mathrm{J}_{\psi^{-1}(y)}\psi)^{\intercal}, \\
\widetilde{\beta}(y) &:= (\widetilde{\beta}_i(y))_i^{\intercal} = (\mathrm{J}_{\psi^{-1}(y)}\psi) \cdot \beta(\psi^{-1}(y)) + (G_0\psi)(\psi^{-1}(y)), \\
\widetilde{\gamma}(y) &= (\gamma \circ \psi^{-1})(y),
\end{aligned}$$

in which $G_0\psi = \sum_{i,j=1}^{N} \alpha_{i,j} \frac{\partial^2}{\partial x_i\, \partial x_j}\psi$ over D (meaning that we are applying the principal part of G to each component of the vector-function ψ).

From the fact that, for each $x \in D$, the matrix $\mathrm{J}_x\psi$ is non-singular, we deduce that the principal matrices of G and $\widetilde{G}$ are congruent, so

$$G \text{ is semielliptic if and only if } \widetilde{G} \text{ is semielliptic.}$$

We also draw attention on the particular case of a linear change of coordinates

$$\psi(x) = M(x - x_0),$$

for a certain $x_0 \in \mathbb{R}^N$, $x \in D$, where M is an $N \times N$ non-singular matrix: in this case, for each $y \in \widetilde{D}$, we have

$$\widetilde{H}(y) = M \cdot H(\psi^{-1}(y)) \cdot M^{\intercal}, \quad \widetilde{\beta}(y) = M \cdot \beta(\psi^{-1}(y)), \quad \widetilde{\gamma}(y) = \gamma(\psi^{-1}(y)).$$

In the end, as a worthy conclusion of this thesis, we patiently give the proof of Proposition 3.3.

Proof of Proposition 3.3. We have $u \in C^2(\Omega)$ an L-subharmonic function and X_0 the drift of L. We suppose that u attains its maximum at a point $x_0 \in \Omega$ and that there is a unit vector $\nu \in \mathbb{R}^N$ such that

$$\langle X_0(x_0), \nu \rangle > 0. \tag{3.3}$$

We have to show that, for each $\rho > 0$, the function u attains its maximum not only at x_0 but also at some point of the set $\Omega \cap B(x_0 + \rho\nu, \rho)$.

For the sake of clarity we split the proof in three steps.

[STEP I]

If $\langle A(x_0)\nu, \nu\rangle > 0$, by Corollary 3.5, we have finished. Otherwise, being $A(x_0)$ positive semidefinte, $\langle A(x_0)\nu, \nu\rangle = 0$, that, thanks to Proposition 1.3, means $A(x_0)\nu = 0$, thus $r := \operatorname{rank}(A(x_0)) < N$. Being $A(x_0)$ othogonally diagonalizable, there exists a non-singular $N \times N$ matrix P such that

$$\text{(a): } P^{\intercal} e_N = \nu \quad \text{ and } \quad \text{(b): } PA(x_0)P^{\intercal} = \begin{pmatrix} \mathrm{Id}_{N\times N} & 0 \\ 0 & 0 \end{pmatrix}.$$

We consider the linear change of variables

$$\psi : \mathbb{R}^N \to \mathbb{R}^N, \quad y \equiv \psi(x) := P(x - x_0).$$

According to the previous Remark 3.6, L is turned by ψ into a linear differential semielliptic operator $\widetilde{L}$ over the open set $\widetilde{D} := \psi(D) \subseteq \mathbb{R}^N$, and $\widetilde{L}$ is of the following form:

$$\widetilde{L} := \sum_{i,j=1}^{N} \alpha_{i,j}(y) \frac{\partial^2}{\partial y_i\, \partial y_j} + \sum_{i=1}^{N} \beta_i(y) \frac{\partial}{\partial\, y_i},$$

where, for any $y \in \widetilde{D}$,

$$\widetilde{A}(y) := P \cdot A(\psi^{-1}(y)) \cdot P^{\intercal}, \quad \beta := P \cdot \beta(\psi^{-1}(y)).$$

Moreover, as we did with L in (3.1), we introduce the C^2 vector fields on $\widetilde{D}$ defined by

$$\begin{aligned} Y_i :&= \sum_{j=1}^{N} \alpha_{i,j}(y) \frac{\partial}{\partial y_j}, \quad \text{for } i = 1, \dots, N, \\ Y_0 :&= \sum_{j=1}^{N} (\beta_j(y) - \sum_{i=1}^{N} \frac{\partial \alpha_{i,j}}{\partial y_i}(y)) \frac{\partial}{\partial y_j}. \end{aligned} \tag{3.4}$$

Again from Remark 3.6, we have

$$(\widetilde{L}q)(y) := (L(q \circ \psi))(\psi^{-1}(y)), \qquad \text{with } y \in \widetilde{D} \text{ and } q \in C^2(\widetilde{D}),$$

from which, taking $q(y) = y_k$ for each $k \in \{1, \dots, N\}$, we gain

$$Y_0(y) = PX_0(\psi^{-1}(y)), \quad \text{for each } y \in \widetilde{D},$$

that gives

$$\langle Y_0(0), e_N \rangle = \langle PX_0(x_0), e_N \rangle = \langle X_0(x_0), P^{\intercal} e_N \rangle = \langle X_0(x_0), \nu \rangle > 0. \tag{3.5}$$

Moreover, from the property (b), we get

$$\alpha_{i,i}(0) = \langle \widetilde{A}(0)e_i, e_i \rangle = 0, \quad \text{for each } i \in \{r+1, \dots, N\},$$

that gives, through Lemma 1.17, for each $k \in \{1, \dots, N\}$

$$\frac{\partial \alpha_{i,j}}{\partial y_k}(0) = \left\langle \frac{\partial \widetilde{A}}{\partial y_k}(0)e_i, e_i \right\rangle = 0, \quad \text{for each } i, j \in \{r+1, \dots, N\}. \tag{3.6}$$

As a consequence, we obtain

$$\langle Y_0(0), e_N \rangle \overset{(3.4)}{=} \beta_N(0) - \sum_{i=1}^{N} \frac{\partial \alpha_{i,N}}{\partial y_i}(0) \overset{(3.6)}{=} \beta_N(0) - \sum_{i=1}^{r} \frac{\partial \alpha_{i,N}}{\partial y_i}(0), \tag{3.7}$$

thus, thanks to (3.5),

$$\beta_N(0) - \sum_{i=1}^{r} \frac{\partial \alpha_{i,N}}{\partial y_i}(0) > 0. \tag{3.8}$$

Now, we set $\widetilde{\Omega} := \psi(\Omega)$ and we consider the function $\widetilde{u} := u \circ \psi^{-1}$. Obviously $\widetilde{u} \in C^2(\widetilde{\Omega})$ and it attains its maximum at $0 = \psi(x_0)$. Furthermore, for each $y \in \widetilde{\Omega}$,

$$(\widetilde{L}\widetilde{u})(y) := (L(\widetilde{u} \circ \psi))(\psi^{-1}(y)) = (Lu)(\psi^{-1}(y)) \geq 0,$$

that means that $\widetilde{u}$ is $\widetilde{L}$-subharmonic over $\widetilde{\Omega}$.

[STEP II]

The aim of this second step is to show the thesis for $\widetilde{u}$ and $\widetilde{L}$. We want to prove that, for each $\rho > 0$, the function $\widetilde{u}$ attains its maximum at some point of the set $\widetilde{\Omega} \cap B(\rho e_N, \rho)$. In this regard, we consider the polynomial function

$$f(y) := y_N - \frac{1}{2} \sum_{k=1}^{r} y_k \Big(\frac{\partial \alpha_{k,N}}{\partial y_k}(0) y_k + 2 \sum_{j=k+1}^{N} \frac{\partial \alpha_{k,N}}{\partial y_j}(0) y_j \Big) - c \sum_{k=1}^{r} y_k^2 - C \sum_{k=r+1}^{N} y_k^2,$$

where c and C are suitable positive constants which we shall fix later on. We notice that $f(0) = 0$ and $\nabla_0 f = 0$. Moreover, a direct computation (based on property (b)) gives

$$\widetilde{L}f(0) = -2rc + \beta_N(0) - \sum_{k=1}^{r} \frac{\partial \alpha_{k,N}}{\partial y_k}(0).$$

By (3.8), we can suppose c so small that $\widetilde{L}f(0) > 0$. Owing to Proposition 3.4, it is possible to find a ratio $r_0 > 0$ such that $\overline{B(0, r_0)} \subseteq \Omega$ and, for each $0 < \epsilon < r_0$, the function attains its maximum at a point $z(0, \epsilon) \cap \{f > 0\}$, that is $\|z\| < \epsilon$ and

$$z_N > \frac{1}{2} \sum_{k=1}^{r} z_k \Big(\frac{\partial \alpha_{k,N}}{\partial y_k}(0) z_k + 2 \sum_{j=k+1}^{N} \frac{\partial \alpha_{k,N}}{\partial y_j}(0) z_j \Big) + c \sum_{k=1}^{r} z_k^2 + C \sum_{k=r+1}^{N} z_k^2. \tag{3.9}$$

Now, for each $y \in B(0, r_0)$, we can consider the integral curve $\gamma(t, Y_1, y)$ of Y_1 starting at y, which is C^2 in a suitable open interval containing zero. From ODE theory we know that (over a suitable domain) the function $(t, y) \mapsto \gamma(t, Y_1, y)$ is continuous, hence, shrinking r_0 if necessary, we can assume that, for every $y \in B(0, r)$,

$$y^{(1)} := \gamma(-y_1, Y, y) \text{ is well-posed and it belongs to } \widetilde{\Omega},$$

where y_1 is the first coordinate of the vector y. Afterwards, we can consider the integral curve $\gamma(t, Y_2, y^{(1)})$ of Y_2 starting at $y^{(1)}$. As above, we can assume r_0 so small that

$$y^{(2)} := \gamma(-y_2^{(1)}, Y_2, y^{(1)}) \text{ is well-posed and it belongs to } \widetilde{\Omega},$$

where $y_2^{(1)}$ is the second coordinate of $y^{(1)}$. By repeating this argument r times, we can suppose r_0 so small that

$$y^{(1)} := \gamma(-y_1, Y, y), \dots, y^{(r)} := \gamma(-y_r^{(r-1)}, Y_r, y^{(r-1)})$$

are all well-defined in $\tilde{\Omega}$. We can consider the map

$$\theta : B(0, r_0) \to \tilde{\Omega}, \quad \theta(y) := \gamma(-y_r^{(r-1)}, Y_r, y^{(r-1)}),$$

that is of class C^2 being $Y_1, \dots, Y_N$ of class C^2 (because integral curves of these vector fields are C^2 in both time and space). Since $Y_1, \dots, Y_N$ are principal vector fields (from Remark 3.1), thanks to the Maximum Propagation Principle Theorem 2.7, we deduce that, if $z \in B(0, \epsilon)$ is any maximum point for $\widetilde{u}$, then the same is true of $\overline{z} := \theta(z)$.

Now, taking $\rho > 0$, we want to show that, if ϵ is sufficiently small and C is sufficiently large, then $\theta(z) \in B(\rho e_N, \rho)$. Thanks to Remark 1.12, one is able to recognize that $\theta_1, \dots, \theta_r$ admit the following expansion at zero:

$$\theta_i(y) = -\frac{1}{2} \sum_{k=1}^{r} y_k \Big(\frac{\partial \alpha_{k,i}}{\partial y_k}(0) y_k + 2 \sum_{j=k+1}^{N} \frac{\partial \alpha_{k,i}}{\partial y_j}(0) y_j \Big) + \mathrm{o}(\|y\|^2);$$

and for $\theta_{r+1}, \dots, \theta_N$ it holds that

$$\theta_i(y) = y_i - \frac{1}{2}\sum_{k=1}^{r} y_k\Big(\frac{\partial \alpha_{k,i}}{\partial y_k}(0) y_k + 2\sum_{j=k+1}^{N}\frac{\partial \alpha_{k,i}}{\partial y_j}(0) y_j\Big) + \mathrm{o}(\|y\|^2). \tag{3.10}$$

From this, by shrinking r_0 if necessary, we infer the existence of a positive constant $M > 0$ such that, for each $y \in B(0, r_0)$,

$$\sum_{k=1}^{N-1} (\theta_k(y))^2 \le M\Big(\sum_{i=1}^{r} y_i^4 + \sum_{i=r+1}^{N} y_i^2\Big); \tag{3.11}$$

moreover, by (3.10) with $i = N$, we can assume that

$$\theta_N(y) > y_N - \frac{1}{2}\sum_{k=1}^{r} y_k\Big(\frac{\partial \alpha_{k,N}}{\partial y_k}(0) y_k + 2\sum_{j=k+1}^{N}\frac{\partial \alpha_{k,N}}{\partial y_j}(0) y_j\Big) - \frac{c}{2}\|y\|^2. \tag{3.12}$$

In particular, if $z \in B(0,\epsilon)\cap\{f > 0\}$ is a maximum point of $\widetilde{u}$, by combining inequality (3.12) with (3.9), then

$$\overline{z}_N = \theta_N(z) > \frac{c}{2}\sum_{i=1}^{r} z_i^2 + \Big(C - \frac{c}{2}\Big)\sum_{i=r+1}^{N} z_i^2. \tag{3.13}$$

Therefore, if we choose $\epsilon > 0$ in such a way that

1. $M\epsilon^2 < \rho c/2$;

2. $\|\theta(y)\| < \rho$ for every $y \in B(0, \epsilon)$,

and if let $C > 0$ be such that $C - \frac{c}{2} > M/\rho$, by (3.11) and (3.13), we obtain

$$\begin{aligned}
\overline{z}_N(2\rho - \overline{z}_N) &= \theta_N(z)(2\rho - \theta_N(z)) \quad \text{(from } \|z\| < \epsilon \text{ and (ii))} \\
&> \rho\theta_N(z) \overset{(3.13)}{>} \frac{\rho c}{2}\sum_{i=1}^{r} z_i^2 + \rho\Big(C - \frac{c}{2}\Big)\sum_{i=r+1}^{N} z_i^2 \quad \text{(owing to (ii))} \\
&> M\sum_{i=1}^{r} z_i^4 + \rho\Big(C - \frac{c}{2}\Big)\sum_{i=r+1}^{N} z_i^2 \quad \text{(by the choice of } C\text{)} \\
&> M\left(\sum_{i=1}^{r} z_i^4 + \sum_{i=r+1}^{N} z_i^2\right) \overset{(3.11)}{\ge} \sum_{k=1}^{N-1} (\theta_k(z))^2.
\end{aligned}$$

From $\theta_N(z)(2\rho - \theta_N(z)) > \sum_{k=1}^{N-1}(\theta_k(z))^2$, one can easily prove that, if we take $z \in B(0,\epsilon)\cap\{f > 0\}$ a maximum point of $\widetilde{u}$, the $\theta(z)$ (which is also a maximum point of $\widetilde{u}$) belongs to $\widetilde{\Omega} \cap B(\rho e_N, \rho)$.

[STEP III]

In this last step, we show that the result proved for $\widetilde{L}$ an $\widetilde{u}$ can be used to demonstrate an analogous fact for L and u. To this end, we consider $\rho > 0$. There exists $\delta > 0$ such that $B(\delta e_N, \delta) \subseteq \psi(B(x_0 + \rho\nu, \rho))$. Indeed: $B(x_0 + \rho\nu, \rho)$ has a smooth boundary and ψ is a smooth diffeomorphism of $\mathbb{R}^N$; if ν is the interior normal vector at x_0 to the ball $B(x_0 + \rho\nu, \rho)$, it can be recognized that e_N is the interior normal vector to $\psi(B(x_0+\rho\nu, \rho))$ at $\psi(x_0) = 0$. From the second step we know that $\widetilde{u}$ attains its maximum at some $\overline{y} \in \widetilde{\Omega} \cap B(\delta e_N, \delta)$; hence $\overline{x} := \psi^{-1}(\overline{y})$ is a maximum point of u belonging to $B(x_0 + \rho\nu, \rho)$.

This puts an end to the last proof. $\square$

Bibliography

[1] S. Biagi and A. Bonfiglioli: *An introduction to the Geometrical Analysis of Vector Fields with Applications to Maximum Principle and Lie Groups*, World Scientific, 2019.

[2] L. Hörmander: *Hypoelliptic second order differential equations*, Acta Math, 1967.

[3] E. Lanconelli: *Lezioni di ANALISI MATEMATICA 2 Seconda Parte*, Pitagora Editrice Bologna, 1997.

Ringraziamenti

Desidero esprimere la mia gratitudine a tutti coloro che hanno aiutato, direttamente o indirettamente, al conseguimento di questo risultato.

Non posso evitare di ringraziare il mio relatore di tesi, nonché il mio primo professore di Analisi Matematica, Andrea Bonfiglioli. Innanzitutto, per avermi trasmesso la passione per questa disciplina e per il tempo che mi ha dedicato durante la stesura di questo scritto. E soprattutto, per aver dimostrato, in diverse occasioni, un'umanità rara ed esemplare, schierandosi sempre a favore degli studenti.

Sempre parlando di sostegno accademico, sono grato a tutti i professori di cui ho seguito le lezioni, nel bene e nel male, per avermi fatto comprendere le mie affinità e anche ciò a cui affine proprio non mi sento di essere.

Passando oltre, è il turno dei miei amici e colleghi matematici, con cui ho passato dei bei momenti, che sia stato cercando di risolvere astrusi problemi alla lavagna o mangiando un boccone ai nostri soliti posti.

Poi, non può mancare una parola spesa per gli amici ravennati di più lunga data, i quali, sparsi in disverse zone d'Italia, hanno trovato ugualmente il modo di restare in contatto ed uniti.

Proseguendo, sono riconoscente alla mia famiglia, in particolare a mia mamma, mio babbo e mia nonna, che mi hanno finanziato ed appoggiato in questo percorso, e a mia sorella Ilaria, la quale mi è stata sempre vicino nei momenti di difficoltà.

Infine e più in generale, a ciascuna delle persone che ha contribuito a questo progetto e che è stata presente durante il viaggio di cui questa tesi è il culmine, rivolgo i miei più sentiti ringraziamenti.