DIPARTIMENTO DI MATEMATICA
"Francesco Brioschi"
POLITECNICO DI MILANO

# $L^p$ and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift

Bramanti, M.; Zhu, M.

Collezione dei *Quaderni di Dipartimento*, numero **QDD 92**
Inserito negli *Archivi Digitali di Dipartimento* in data 28-3-2011

Piazza Leonardo da Vinci, 32 - 20133 Milano (Italy)

# $L^p$ and Schauder estimates for nonvariational operators structured on Hörmander vector fields with drift*

Marco Bramanti and Maochun Zhu

March 26, 2011

**Abstract**

Let

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j + a_0(x) X_0,$$

where $X_0, X_1, ..., X_q$ are real smooth vector fields satisfying Hörmander's condition in some bounded domain $\Omega \subset \mathbb{R}^n$ $(n > q + 1)$, the coeffi cients $a_{ij} = a_{ji}, a_0$ are real valued, bounded measurable functions defined in $\Omega$, satisfying the uniform positivity conditions:

$$\mu |\xi|^2 \leq \sum_{i,j=1}^{q} a_{ij}(x) \xi_i \xi_j \leq \mu^{-1} |\xi|^2 ; \mu \leq a_0(x) \leq \mu^{-1}$$

for a.e. $x \in \Omega$, every $\xi \in \mathbb{R}^q$, some constant $\mu > 0$.

We prove that if the coeffi cients $a_{ij}, a_0$ belong to the Hölder space $C_X^{\alpha}(\Omega)$ with respect to the distance induced by the vector fields, then local Schauder estimates of the following kind hold:

$$\|X_i X_j u\|_{C_X^{\alpha}(\Omega')} + \|X_0 u\|_{C_X^{\alpha}(\Omega')} \leq c \left\{ \|Lu\|_{C_X^{\alpha}(\Omega)} + \|u\|_{L^{\infty}(\Omega)} \right\}$$

for any $\Omega' \Subset \Omega$;

if the coeffi cients $a_{ij}, a_0$ belong to the space $VMO_{X,loc}(\Omega)$ with respect to the distance induced by the vector fields, then local $L^p$ estimates of the following kind hold, for every $p \in (1, \infty)$:

$$\|X_i X_j u\|_{L^p(\Omega')} + \|X_0 u\|_{L^p(\Omega')} \leq c \left\{ \|Lu\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)} \right\}.$$

*$^{*}$***2000 AMS Classification**: Primary 35H20. Secondary: 35B45, 42B20, 53C17. **Keywords**: Hörmander's vector fields, Schauder estimates, $L^p$ estimates, drift

# Contents



# 1 Introduction

Let us consider a family of real smooth vector fields

$$X_i = \sum_{j=1}^{n} b_{ij}(x)\,\partial_{x_j}, \quad i = 0, 1, 2, ..., q$$

$(q + 1 < n)$ defined in some bounded domain $\Omega$ of $\mathbb{R}^n$ and satisfying Hörmander's condition: the Lie algebra generated by the $X_i$'s at any point of $\Omega$ span

$\mathbb{R}^n$. Under these assumptions, Hörmander's operators

$$\mathcal{L} = \sum_{i=1}^{q} X_i^2 + X_0$$

have been studied since the late 1960's. Hörmander [20] proved that $\mathcal{L}$ is hypoelliptic, while Rothschild-Stein [25] proved that for these operators *a priori* estimates of $L^p$ type for second order derivatives with respect to the vector fields hold, namely:

$$\sum_{i,j=1}^{q} \|X_i X_j u\|_{L^p(\Omega')} + \|X_0 u\|_{L^p(\Omega')} \leq c \left\{ \|\mathcal{L}u\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)} + \sum_{i=1}^{q} \|X_i u\|_{L^p(\Omega)} \right\} \tag{1.1}$$

for any $p \in (1, \infty)$, $\Omega' \Subset \Omega$.

Note that the "drift" vector field $X_0$ has weight two, compared with the vector fields $X_i$ for $i = 1, 2, ..., q$. For operators without the drift term ("sum of squares" of Hörmander type) many more results have been proved in the literature than for complete Hörmander's operators. On the other hand, complete operators owe their interest, for instance, to the class of Kolmogorov-Fokker-Planck operators, naturally arising in many fields of physics, natural sciences and finance, as the transport-diffusion equations satisfied by the transition probability density of stochastic systems of O.D.E.s which describe some real system led to a basically deterministic law perturbed by some kind of white noise. The study of Kolmogorov-Fokker-Planck operators in the framework of Hörmander's operators received a strong impulse by the paper [22] by Lanconelli-Polidoro, which started a lively line of research. We refer to [21] for a good survey on this field, with further motivations for the study of these equations and related references.

Let us also note that the study of Hörmander's operators is considerably easier when $\mathcal{L}$ is left invariant with respect to a suitable Lie group of translations and homogeneous of degree two with respect to a suitable family of dilations (which are group automorphisms of the corresponding group of translations). In this case we say that $\mathcal{L}$ has an underlying structure of homogeneous group and, by a famous result due to Folland [16], $\mathcal{L}$ possesses a homogeneous left invariant global fundamental solution, which turns out to be a precious tool in proving *a priori* estimates.

In the last ten years, more general classes of nonvariational operators structured on Hörmander's vector fields have been studied, namely

$$\mathcal{L}=\sum_{i,j=1}^{q}a_{ij}\left(x\right)X_iX_j \tag{1.2}$$

$$\mathcal{L}=\sum_{i,j=1}^{q}a_{ij}\left(x\right)X_iX_j-\partial_t \tag{1.3}$$

$$\mathcal{L}=\sum_{i,j=1}^{q}a_{ij}\left(x\right)X_iX_j+a_0\left(x\right)X_0 \tag{1.4}$$

where the matrix $\{a_{ij}\left(x\right)\}_{i,j=1}^{q}$ is symmetric positive definite, the coefficients are bounded ($a_0$ is bounded away from zero) and satisfy suitable mild regularity assumptions, for instance they belong to Hölder or $VMO$ spaces defined with respect to the distance induced by the vector fields. Since the $a_{ij}$'s are not $C^\infty$, these operators are no longer hypoelliptic. Nevertheless, *a priori* estimates on second order derivatives with respect to the vector fields are a natural result which does not in principle require smoothness of the coefficients. Namely, *a priori* estimates in $L^p$ (with coefficients $a_{ij}$ in $VMO\cap L^\infty$) have been proved in [3] for operators (1.2) and in [2] for operators (1.4) but in homogeneous groups; *a priori* estimates in $C^\alpha$ spaces (with coefficients $a_{ij}$ in $C^\alpha$) have been proved in [4] for operators (1.3) and in [19] for operators (1.4) but in homogeneous groups.

In the particular case of Kolmogorov-Fokker-Planck operators, which can be written as

$$\mathcal{L}=\sum_{i,j=1}^{q}a_{ij}\left(x\right)\partial_{x_ix_j}^2+X_0$$

for a suitable drift $X_0$, $L^p$ estimates (when $a_{ij}$ are $VMO$) have been proved in [7] in homogeneous groups, while Schauder estimates (when $a_{ij}$ are Hölder continuous) have been proved in [15], under more general assumptions (namely, assuming the existence of translations but not necessarily dilations, adapted to the operator). We recall that the idea of proving $L^p$ estimates for nonvariational operators with leading coefficients in $VMO\cap L^\infty$ (instead of assuming their uniform continuity) appeared for the first time in the papers [11], [12] by Chiarenza-Frasca-Longo, in the uniformly elliptic case.

The aim of the present paper is to prove both $L^p$ and $C^\alpha$ local estimates for general operators (1.4) structured on Hörmander's vector fields "with drift", without assuming the existence of any group structure, under the appropriate assumptions on the coefficients $a_{ij}, a_0$. Namely, our basic estimates read as follows:

$$\|u\|_{S_X^{2,p}(\Omega')}\leq c\left\{\|\mathcal{L}u\|_{L^p(\Omega)}+\|u\|_{L^p(\Omega)}\right\} \tag{1.5}$$

for $p\in(1,\infty)$ and any $\Omega'\Subset\Omega$ if the coefficients are $VMO_{X,loc}\left(\Omega\right)$, and

$$\|u\|_{C_X^{2,\alpha}(\Omega')}\leq c\left\{\|\mathcal{L}u\|_{C_X^{\alpha}(\Omega)}+\|u\|_{L^\infty(\Omega)}\right\} \tag{1.6}$$

for $\alpha \in (0,1)$ and $\Omega' \Subset \Omega$ if the coeffi cients are $C_X^\alpha(\Omega)$. The related Sobolev and Hölder spaces $S_X^{2,p}, C_X^{2,\alpha}$, are those induced by the vector fields $X_i$'s, and will be precisely defined in §3.4. Clearly, these estimates are more general than those contained in all the aforementioned papers.

At first sight, this kind of result could seem a straightforward generalization of existing theories. However, several diffi culties exist, sometimes hidden in subtle details. First of all, we have to remark that in the paper [25], although the results are stated for both sum of squares and complete Hörmander's operators, proofs are given only in the first case. While some adaptations are quite straightforward, this is not always the case. Therefore, some results proved in the present paper can be seen also as a detailed proof of results stated in [25], in the drift case. We will justify this statement later, when dealing with these details. For the moment we just point out that these diffi culties are mainly related to the proof of suitable representation formulas for second order derivatives $X_iX_ju$ of a test function, in terms of $u$ and $\mathcal{L}u$, via singular integrals and commutators of singular integrals. In turn, the reason why these representation formulas are harder to be proved in presence of a drift relies on the fact that a technical result which allows to exchange, in a suitable sense, the action of $X_i$-derivatives with that of suitable integral operators, assumes a more involved form when the drift is present.

Once the suitable representation formulas are established, a real variable machinery similar to that used in [3] and [4] can be applied, and this is the reason why we have chosen to give here a unified treatment to $L^p$ and $C^\alpha$ estimates. More specifically, one considers a bounded domain $\Omega$ endowed with the control distance induced by the vector fields $X_i$'s, which has been defined, in the drift case, by Nagel-Stein-Wainger in [23], and the Lebesgue measure, which is locally doubling with respect to these metric balls, as proved in [23]. However, a problem arises when trying to apply to this context known results about singular integrals in metric doubling spaces (or "spaces of homogeneous type", after [14]). Namely, what we should know to apply this theory is a doubling property as

$$\mu(B(x,2r)\cap\Omega') \leq c\mu(B(x,r)\cap\Omega') \text{ for any } x\in\Omega'\Subset\Omega, r>0 \tag{1.7}$$

while what we actually know in view of [23] is

$$\mu(B(x,2r)) \leq c\mu(B(x,r)) \text{ for any } x\in\Omega'\Subset\Omega, 0<r<r_0. \tag{1.8}$$

Now, it has been known since [18] that, when $\Omega'$ is for instance a metric ball, condition (1.7) follows from (1.8) as soon as the distance satisfies a kind of *segment property* which reads as follows: for any couple of points $x_1, x_2$ at distance $r$ and for any number $\delta < r$ and $\varepsilon > 0$ there exists a point $x_0$ having distance $\leq \delta$ from $x_1$ and $\leq r-\delta+\varepsilon$ from $x_2$. However, while when the drift term is lacking the distance induced by the $X_i$'s is easily seen to satisfy this property, this is no longer the case when the field $X_0$ with weight two enters the definition of distance and, as far as we know, a condition of kind (1.7) has never been proved in this context for $\Omega'$ a metric ball, or for any other special kind of bounded domain $\Omega$. Thus we are forced to apply a theory of singular integrals which does not require the full strength of the global doubling condition (1.7). A first possibility is to consider the context of *nondoubling spaces*, as studied by Tolsa, Nazarov-Treil-Volberg, and other authors (see for instance [28], [24], and references therein). Results of $L^p$ and $C^\alpha$ continuity for singular integrals of this kind, applicable to our context, have been proved in [1]. However, to prove our $L^p$ estimates (1.5), we also need some *commutator estimates*, of the kind of the well-known result proved by [13], that, as far as we know, are not presently available in the framework of general nondoubling quasimetric (or metric) spaces. For this reason, we have recently developed in [8] a theory of *locally homogeneous spaces* which is a quite natural framework where all the results we need about singular integrals and their commutators with BMO functions can be proved. To give a unified treatment to both $L^p$ and $C^\alpha$ estimates, here we have decided to prove both exploiting the results in [8]. We note that our Schauder estimates could also be obtained applying the results in [1], while $L^p$ estimates cannot.

The necessity of avoiding the use of a doubling property of type (1.7), as well as some modifications required by the presence of the drift $X_0$, also reflects in the way we have studied several properties of the function spaces $C_X^\alpha$ and $BMO$, as we will see in § 3.4.

Once the basic estimates on second order derivatives are established, a natural, but nontrivial, extension consists in proving similar estimates for derivatives of (weighted) order $k+2$, in terms of $k$ derivatives of $\mathcal{L}u$ (assuming, of course, that the coeffi cients of the operator possess the corresponding further regularity). In presence of a drift, it is reasonable to restrict this study to the case of $k$ even, as already appears from the analog result proved in homogeneous groups in [2]. Even in this case, a proof of this extensions seems to be a diffi cult task, and we have preferred not to address this problem in the present paper, in order not to further increase its length.

**Acknowledgements.** This research was mainly carried out while Maochun Zhu was visiting the Department of Mathematics of Politecnico di Milano, which we wish to thank for the hospitality. The project was supported by the National Natural Science Foundation of China (Grant No. 10871157), Specialized Research Fund for the Doctoral Program of Higher Education (No. 200806990032).

# 2 Assumptions and main results

We now state precisely our assumptions and main results. All the function spaces involved in the statements below will be defined precisely in § 3. Our basic assumption is:

**Assumption (H).** Let

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j + a_0(x) X_0,$$

where the $X_0, X_1, ..., X_q$ are real smooth vector fields satisfying Hörmander's condition in some bounded domain $\Omega \subset \mathbb{R}^n$, the coefficients $a_{ij} = a_{ji}, a_0$ are real valued, bounded measurable functions defined in $\Omega$, satisfying the uniform positivity conditions:

$$\mu|\xi|^2 \leq \sum_{i,j=1}^{q} a_{ij}(x)\xi_i\xi_j \leq \mu^{-1}|\xi|^2;$$

$$\mu \leq a_0(x) \leq \mu^{-1}$$

for a.e. $x \in \Omega$, every $\xi \in \mathbb{R}^q$, some constant $\mu > 0$.

Our main results are the following:

**Theorem 2.1** *In addition to assumption (H), assume that the coefficients $a_{ij}, a_0$ belong to $C_X^\alpha(\Omega)$ for some $\alpha \in (0,1)$. Then for every domain $\Omega' \Subset \Omega$, there exists a constant $c > 0$ depending on $\Omega', \Omega, X_i, \alpha, \mu, \|a_{ij}\|_{C_X^\alpha(\Omega)}$ and $\|a_0\|_{C_X^\alpha(\Omega)}$ such that, for every $u \in C_X^{2,\alpha}(\Omega)$, one has*

$$\|u\|_{C_X^{2,\alpha}(\Omega')} \leq c\left\{\|\mathcal{L}u\|_{C_X^\alpha(\Omega)} + \|u\|_{L^\infty(\Omega)}\right\}.$$

**Theorem 2.2** *In addition to assumption (H), assume that the coefficients $a_{ij}, a_0$ belong to the space $VMO_{X,loc}(\Omega)$. Then for every $p \in (1,\infty)$, any $\Omega' \Subset \Omega$, there exists a constant c depending on $X_i, n, q, p, \mu, \Omega', \Omega$ and the VMO moduli of $a_{ij}$ and $a_0$, such that for every $u \in S_X^{2,p}(\Omega)$,*

$$\|u\|_{S_X^{2,p}(\Omega')} \leq c\left\{\|\mathcal{L}u\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)}\right\}.$$

**Remark 2.3** *Under the assumptions of the previous theorems, it is not restrictive to assume $a_0(x)$ to be equal to 1, for we can always rewrite the equation*

$$\sum_{i,j=1}^{q} a_{ij}X_iX_j + a_0X_0 = f$$

*in the form*

$$\sum_{i,j=1}^{q} \frac{a_{ij}}{a_0}X_iX_j + X_0 = \frac{f}{a_0}$$

*and apply the a-priori estimates to this equation, controlling $C^\alpha$ or $VMO$ moduli of the new coefficients $\frac{a_{ij}}{a_0}$ in terms of the analogous moduli of $a_{ij}, a_0$ and the constant $\mu$. Therefore throughout the following we will always take $a_0 \equiv 1$.*

# 3 Known results and preparatory material from real analysis and geometry of vector fields

## 3.1 Some known facts about Hörmander's vector fields, lifting and approximation

Let $X_0, X_1, \ldots, X_q$ be a system of real smooth vector fields,

$$X_i = \sum_{j=1}^{n} b_{ij}(x)\,\partial_{x_j}, \quad i = 0, 1, 2, ..., q$$

$(q + 1 < n)$ defined in some bounded, open and connected subset $\Omega$ of $\mathbb{R}^n$. Let us assign to each $X_i$ a weight $p_i$, saying that

$$p_0 = 2 \text{ and } p_i = 1 \text{ for } i = 1, 2, ...q.$$

For any multiindex

$$I = (i_1, i_2, ..., i_k),$$

we define the weight of $I$ as

$$|I| = \sum_{j=1}^{k} p_{i_j}.$$

For any couple of vector fields $X, Y$, let $[X, Y] = XY - YX$ be their commutator. Now, for any multiindex $I = (i_1, i_2, ..., i_k)$ for $0 \leq i_k \leq q$ we set:

$$X_I = X_{i_1} X_{i_2} ... X_{i_k}$$

and

$$X_{[I]} = \left[X_{i_1}, \left[X_{i_2}, \ldots \left[X_{i_{k-1}}, X_{i_k}\right] \ldots\right]\right].$$

If $I = (i_1)$, then

$$X_{[I]} = X_{i_1} = X_I.$$

We will say that $X_{[I]}$ is a *commutator of weight* $|I|$. As usual, $X_{[I]}$ can be seen either as a differential operator or as a vector field. We will write

$$X_{[I]} f$$

to denote the differential operator $X_{[I]}$ acting on a function $f$, and

$$\left(X_{[I]}\right)_x$$

to denote the vector field $X_{[I]}$ evaluated at the point $x \in \Omega$.

We shall say that $X = \{X_0, X_1, \dots, X_q\}$ satisfy *Hörmander's condition of weight* $s$ if these vector fields, together with their commutators of weight $\leq s$, span the tangent space at every point $x \in \Omega$.

Let $\ell$ be the free Lie algebra of weight $s$ on $q+1$ generators, that is the quotient of the free Lie algebra with $q+1$ generators by the ideal generated by the commutators of weight at least $s+1$. We say that the vector fields $X_0, \dots, X_q$, which satisfy Hörmander's condition of weight $s$ at some point $x_0 \in \mathbb{R}^n$, are *free up to order* $s$ *at* $x_0$ if $n = \dim \ell$, as a vector space (note that inequality $\leq$ always holds). The famous Lifting Theorem proved by Rothschild-Stein in [25, p. 272] reads as follows:

**Theorem 3.1** *Let $X = (X_0, X_1, \dots, X_q)$ be $C^\infty$ real vector fields on a domain $\Omega \subset \mathbb{R}^n$ satisfying Hörmander's condition of weight $s$ in $\Omega$. Then, for any $\overline{x} \in \Omega$, in terms of new variables, $h_{n+1}, \dots, h_N$, there exist smooth functions $\lambda_{il}(x,h)$ $(0 \leq i \leq q,\, n+1 \leq l \leq N)$ defined in a neighborhood $\widetilde{U}$ of $\overline{\xi} = (\overline{x}, 0) \in \mathbb{R}^N$ such that the vector fields $\widetilde{X}_i$ given by*

$$\widetilde{X}_i = X_i + \sum_{l=n+1}^{N} \lambda_{il}(x,h) \frac{\partial}{\partial h_l}, \quad i = 0, \dots, q$$

*satisfy Hörmander's condition of weight $s$ and are free up to weight $s$ at every point in $\widetilde{U}$.*

Let $\widetilde{X} = \left(\widetilde{X}_0, \widetilde{X}_1, \dots, \widetilde{X}_q\right)$ be the lifted vector fields which are free up to weight $s$ at some point $\xi \in \mathbb{R}^N$ and $\ell$ be the free Lie algebra generated by $\widetilde{X}$. For each $j$, $1 \leq j \leq s$, we can select a family $\left\{\widetilde{X}_{j,k}\right\}_k$ of commutators of weight $j$, with $\widetilde{X}_{1,k} = \widetilde{X}_k, \widetilde{X}_{2,1} = \widetilde{X}_0, k = 1, 2, \dots, q$, such that $\left\{\widetilde{X}_{j,k}\right\}_{jk}$ is a basis of $\ell$, that is to say, there exists a set $A$ of double-indices $\alpha$ such that $\left\{\widetilde{X}_\alpha\right\}_{\alpha \in A}$ is a basis of $\ell$. Note that $\mathrm{Card} A = N$, which allows us to identify $\ell$ with $\mathbb{R}^N$.

Now, in $\mathbb{R}^N$ we can consider the group structure of $N(q+1, s)$, which is the simply connected Lie group associated to $\ell$. We will write $\circ$ for the Lie group operation (which we think as a *translation*) and will assume that the group identity is the origin. It is also possible to assume that $u^{-1} = -u$ (the group inverse is the Euclidean opposite). We can naturally define *dilations* in $N(q+1, s)$ by

$$D(\lambda)\left((u_\alpha)_{\alpha \in A}\right) = \left(\lambda^{|\alpha|} u_\alpha\right)_{\alpha \in A}. \tag{3.1}$$

These are group automorphisms, hence $N(q+1, s)$ is a *homogeneous group*, in the sense of Stein (see [27, p. 618-622]). We will call it $\mathbb{G}$, leaving the numbers $q, s$ implicitly understood.

We can define in $\mathbb{G}$ a *homogeneous norm* $\|\cdot\|$ as follows. For any $u \in \mathbb{G}$, $u \neq 0$, set

$$\|u\| = r \quad \Leftrightarrow \quad \left| D\left(\frac{1}{r}\right) u \right| = 1,$$

where $|\cdot|$ denotes the Euclidean norm.

The function

$$d_{\mathbb{G}}(u,v) = \left\| v^{-1} \circ u \right\|$$

is a *quasidistance*, that is:

$$\begin{aligned} & d_{\mathbb{G}}(u,v) \geq 0 \text{ and } d_{\mathbb{G}}(u,v) = 0 \text{ if and only if } u = v; \\ & d_{\mathbb{G}}(u,v) = d_{\mathbb{G}}(v,u); \\ & d_{\mathbb{G}}(u,v) \leq c\left(d_{\mathbb{G}}(u,z) + d_{\mathbb{G}}(z,v)\right), \end{aligned} \tag{3.2}$$

for every $u, v, z \in \mathbb{G}$ and some positive constant $c(\mathbb{G}) \geq 1$. We define the balls with respect to $d_{\mathbb{G}}$ as

$$B(u,r) \equiv \left\{ v \in \mathbb{R}^N : d_{\mathbb{G}}(u,v) < r \right\}.$$

It can be proved (see [27, p.619]) that the Lebesgue measure in $\mathbb{R}^N$ is the Haar measure of $\mathbb{G}$. Therefore, by (3.1),

$$|B(u,r)| = |B(u,1)|\, r^Q,$$

for every $u \in \mathbb{G}$ and $r > 0$, where $Q = \sum_{\alpha \in A} |\alpha|$. We will call $Q$ the *homogeneous dimension* of $\mathbb{G}$.

Next, we define the *convolution* of two functions in $\mathbb{G}$ as

$$(f * g)(u) = \int_{\mathbb{R}^N} f\left(u \circ v^{-1}\right) g(v)\, dv = \int_{\mathbb{R}^N} g\left(v^{-1} \circ u\right) f(v)\, dv,$$

for every couple of functions for which the above integrals make sense.

Let $\tau_u$ be the left translation operator acting on functions: $(\tau_u f)(v) = f(u \circ v)$. We say that a differential operator $P$ on $\mathbb{G}$ is *left invariant* if $P(\tau_u f) = \tau_u(Pf)$ for every smooth function $f$. From the above definition of convolution we read that if $P$ is any left invariant differential operator,

$$P(f * g) = f * Pg \tag{3.3}$$

(provided the integrals converge).

We say that *a differential operator* $P$ on $\mathbb{G}$ is *homogeneous of degree* $\delta > 0$ if

$$P(f(D(\lambda)u)) = \lambda^{\delta}(Pf)(D(\lambda)u)$$

for every test function $f$ and $\lambda > 0, u \in \mathbb{G}$. Also, we say that *a function* $f$ is *homogeneous of degree* $\delta \in \mathbb{R}$ if

$$f(D(\lambda)u) = \lambda^{\delta} f(u) \text{ for every } \lambda > 0, u \in \mathbb{G}.$$

Clearly, if $P$ is a differential operator homogeneous of degree $\delta_1$ and $f$ is a homogeneous function of degree $\delta_2$, then $Pf$ is a homogeneous function of degree $\delta_2 - \delta_1$, while $fP$ is a differential operator, homogeneous of degree $\delta_1 - \delta_2$.

Let $Y_\alpha$ be the left invariant vector field which agrees with $\frac{\partial}{\partial u_\alpha}$ at 0 and set $Y_{1,k} = Y_k, k = 1, \cdots, q, Y_{2,1} = Y_0$. The differential operator $Y_{i,k}$ is homogeneous of degree $i$, and $\{Y_\alpha\}_{\alpha \in A}$ is a basis of the free Lie algebra $\ell$.

A differential operator on $\mathbb{G}$ is said to have *local degree less than or equal to* $\lambda$ if, after taking the Taylor expansion at 0 of its coeffi cients, each term obtained is a differential operator homogeneous of degree $\leq \lambda$.

Also, a function on $\mathbb{G}$ is said to have *local degree greater than or equal to* $\lambda$ if, after taking the Taylor expansion at 0 of its coeffi cients, each term obtained is a homogeneous function of degree $\geq \lambda$.

For $\xi, \eta \in \widetilde{U}$, define the map

$$\Theta_\eta(\xi) = (u_\alpha)_{\alpha \in A}$$

with $\xi = \exp\left(\sum_{\alpha \in A} u_\alpha \widetilde{X}_\alpha\right) \eta$. We will also write $\Theta(\eta, \xi) = \Theta_\eta(\xi)$.

We can now state Rothschild-Stein's approximation theorem (see [25, p. 273]).

**Theorem 3.2** *In the coordinates given by $\Theta(\eta, \cdot)$ we can write $\widetilde{X}_i = Y_i + R_i^\eta$ on an open neighborhoods of 0, where $R_i^\eta$ is a vector field of local degree $\leq 0$ for $i = 1, \ldots, q$ ($\leq 1$ for $i = 0$) depending smoothly on $\eta$. Explicitly, this means that for every $f \in C_0^\infty(\mathbb{G})$:*

$$\widetilde{X}_i \left[f\left(\Theta\left(\eta, \cdot\right)\right)\right](\xi) = \left(Y_i f + R_i^\eta f\right)\left(\Theta\left(\eta, \xi\right)\right). \tag{3.4}$$

*More generally, for every double-index $(i, k) \in A$, we can write*

$$\widetilde{X}_{i,k} \left[f\left(\Theta\left(\eta, \cdot\right)\right)\right](\xi) = \left(Y_{i,k} f + R_{i,k}^\eta f\right)\left(\Theta\left(\eta, \xi\right)\right), \tag{3.5}$$

*where $R_{i,k}^\eta$ is a vector field of local degree $\leq i - 1$ depending smoothly on $\eta$.*

This theorem says that the lifted vector fields $\widetilde{X}_i$ can be locally approximated by the homogeneous, left invariant vector fields $Y_i$ on the group $\mathbb{G}$. Some other important properties of the map $\Theta$ are stated in the next theorem (see [25, p. 284-287]):

**Theorem 3.3** *Let $\overline{\xi} \in \mathbb{R}^N$ and $\widetilde{U}$ be a neighborhood of $\overline{\xi}$ such that for any $\eta \in \widetilde{U}$ the map $\Theta(\eta, \cdot)$ is well defined in $\widetilde{U}$. For $\xi, \eta \in \widetilde{U}$, define*

$$\rho(\eta, \xi) = ||\Theta(\eta, \xi)|| \tag{3.6}$$

*where $||\cdot||$ is the homogeneous norm defined above. Then:*

*(a)* $\Theta(\eta, \xi) = \Theta(\xi, \eta)^{-1} = -\Theta(\xi, \eta)$ *for every $\xi, \eta \in \widetilde{U}$;*

*(b) $\rho$ is a quasidistance in $\widetilde{U}$ (that is satisfies the three properties (3.2));*

*(c) under the change of coordinates $u = \Theta_\xi(\eta)$, the measure element becomes:*

$$d\eta = c(\xi) \cdot (1 + \omega(\xi, u))\, du, \tag{3.7}$$

*where $c(\xi)$ is a smooth function, bounded and bounded away from zero in $\widetilde{U}$, $\omega(\xi, u)$ is a smooth function in both variables, with*

$$|\omega(\xi, u)| \leq c\,\|u\|,$$

*and an analogous statement is true for the change of coordinates $u = \Theta_\eta(\xi)$.*

**Remark 3.4** *As we have recalled in the introduction, in the paper [25] detailed proofs are given only when the drift term $X_0$ is lacking. A proof of the lifting and approximation results explicitly covering the drift case can be found in [6], where the theory is also extended to the case of nonsmooth Hörmander's vector fields. We refer to the introduction of [6] for further bibliographic remarks about existing alternative proofs of the lifting and approximation theorems.*

## 3.2 Metric induced by vector fields

Let us start recalling the definition of control distance given by Nagel-Stein-Wainger in [23] for Hörmander's vector fields with drift:

**Definition 3.5** *For any $\delta > 0$, let $C(\delta)$ be the class of absolutely continuous mappings $\varphi$: $[0, 1] \to \Omega$ which satisfy*

$$\varphi'(t) = \sum_{|I| \leq s} \lambda_I(t) \left(X_{[I]}\right)_{\varphi(t)} \quad a.e.\ t \in (0, 1) \tag{3.8}$$

*with $|\lambda_I(t)| \leq \delta^{|I|}$. We define*

$$d(x, y) = \inf\left\{\delta : \exists \varphi \in C(\delta) \ \text{ with } \varphi(0) = x, \varphi(1) = y\right\}.$$

The finiteness of $d$ immediately follows by Hörmander's condition: since the vector fields $\left\{X_{[I]}\right\}_{|I| \leq s}$ span $\mathbb{R}^n$, we can always join any two points $x, y$ with a curve $\varphi$ of the kind (3.8); moreover, $d$ turns out to be a distance. Analogously to what Nagel-Stein-Wainger do in [23] when $X_0$ is lacking, in [5] the following notion is introduced:

**Definition 3.6** *For any $\delta > 0$, let $C_1(\delta)$ be the class of absolutely continuous mappings $\varphi : [0, 1] \to \Omega$ which satisfy*

$$\varphi'(t) = \sum_{i=0}^{q} \lambda_i(t) \left(X_i\right)_{\varphi(t)} \quad a.e.\ t \in (0, 1)$$

*with $|\lambda_0(t)| \leq \delta^2$ and $|\lambda_j(t)| \leq \delta$ for $j = 1, \cdots q$.*

*We define*

$$d_X(x, y) = \inf\left\{\delta : \exists \varphi \in C_1(\delta) \ \text{ with } \varphi(0) = x, \varphi(1) = y\right\}.$$

Note that the finiteness of $d_X(x,y)$ for any two points $x,y \in \Omega$ is not a trivial fact, but depends on a connectivity result ("Chow's theorem"); moreover, it can be proved that $d$ and $d_X$ are equivalent, and that $d_X$ is still a distance (see [5], where these results are proved in the more general setting of nonsmooth vector fields). From now on we will always refer to $d_X$ as to the control distance, induced by the system of Hörmander's vector fields $X$. It is well-known that this distance is topologically equivalent to the Euclidean one. For any $x \in \Omega$, we set

$$B_r(x) = \{y \in \Omega : d_X(x,y) < r\}.$$

The basic result about the measure of metric balls is the famous local doubling condition proved by Nagel-Stein-Wainger [23]:

**Theorem 3.7** *For every $\Omega' \Subset \Omega$ there exist positive constants $c, r_0$ such that for any $x \in \Omega', r \leq r_0$,*

$$|B(x,2r)| \leq c\,|B(x,r)|.$$

As already pointed out in the introduction, the distance $d_X$ does *not* satisfy the segment property: given two points at distance $r$, it is generally impossible to find a third point at distance $r/2$ from both. A weaker property which this distance actually satisfies is contained in the next lemma, and will be useful when dealing with the properties of Hölder spaces $C^\alpha$:

**Lemma 3.8** *For any $x,y \in \Omega$ and any positive integer $n$, we can join $x$ to $y$ with a curve $\gamma$ and find $n+1$ points $p_0 = x, p_1, p_2, ..., p_n = y$ on $\gamma$, such that*

$$d_X(p_j, p_{j+1}) \leq \frac{d_X(x,y)}{\sqrt{n}} \text{ for } j = 0, 2, ..., n-1.$$

**Proof.** For any $x,y \in \Omega$ with $d_X(x,y) = R$, any $\varepsilon > 0$, by Definition 3.6 we can join $x$ and $y$ with a curve $\gamma(t)$ satisfying

$$\gamma(0) = y, \gamma(1) = x$$

and

$$\gamma'(t) = \sum_{i=0}^{q} \lambda_i(t)(X_i)_{\gamma(t)},$$

with $|\lambda_i(t)| \leq R + \varepsilon$, for $i = 1, \ldots, q$ and $|\lambda_0(t)| \leq (R+\varepsilon)^2$.

Let $\gamma_j(t) = \gamma\left(\frac{t+j}{n}\right)$, for $j = 0, 1, 2, ..., n-1$. Then $\gamma_j(t)$ satisfies

$$\gamma_j(0) = \gamma\left(\frac{j}{n}\right) \equiv p_j, \gamma_j(1) = \gamma\left(\frac{j+1}{n}\right) = p_{j+1};$$

in particular, $p_0 = x$ and $p_n = y$; moreover,

$$\gamma_j'(t) = \frac{1}{n}\sum_{i=0}^{q} \lambda_i\left(\frac{t+j}{n}\right)(X_i)_{\gamma_j(t)},$$

with

$$\left|\frac{1}{n}\lambda_0\left(\frac{t+j}{n}\right)\right| \leq \left(\frac{R+\varepsilon}{\sqrt{n}}\right)^2, \left|\frac{1}{n}\lambda_i\left(\frac{t+j}{n}\right)\right| < \frac{R+\varepsilon}{\sqrt{n}}$$

for $i = 1, \ldots, q$, $j = 0, 2, ..., n-1$. Thus

$$d_X\left(p_j, p_{j+1}\right) \leq \frac{R+\varepsilon}{\sqrt{n}},$$

for $j = 0, 2, ..., n-1$ and any $\varepsilon > 0$, so we are done. ■

The free lifted vector fields $\widetilde{X}_i$ induce, in the neighborhood where they are defined, a control distance $d_{\widetilde{X}}$; we will denote by $\widetilde{B}(\xi, r)$ the corresponding metric balls. In this lifted setting we can also consider the quasidistance $\rho$ defined in (3.6). The two functions turn out to be equivalent:

**Lemma 3.9** *Let $\overline{\xi}, \widetilde{U}$ be as in Thm. 3.3. There exists $\widetilde{B}\left(\overline{\xi}, R\right) \subset \widetilde{U}$ such that the distance $d_{\widetilde{X}}$ is equivalent to the quasidistance $\rho$ in (3.6) in $\widetilde{B}\left(\overline{\xi}, R\right)$, and both are greater than the Euclidean distance; namely there exist positive constants $c_1, c_2, c_3$ such that*

$$c_1\left|\xi - \eta\right| \leq c_2\rho(\eta, \xi) \leq d_{\widetilde{X}}(\eta, \xi) \leq c_3\rho(\eta, \xi) \text{ for every } \xi, \eta \in \widetilde{B}\left(\overline{\xi}, R\right).$$

This fact is proved in [23], see also [6, Proposition 22].

### 3.3 Some known results about locally homogeneous spaces

We are now going to recall the notion of *locally homogeneous space*, introduced in [8]. This is the abstract setting which will allow us to apply suitable results about singular integrals. Roughly speaking, a locally homogeneous space is a set $\Omega$ endowed with a function $d$ which is a quasidistance on any compact subset, and a measure $\mu$ which is locally doubling, in a sense which will be made precise here below. In our concrete situation, our set is endowed with a function $d$ which is a *distance* in $\Omega$, and a locally doubling measure. We can therefore give the following definition, which is simpler than that given in [8]:

**Definition 3.10** *Let $\Omega$ be a set, endowed with a distance $d$. Let us denote by $B(x, r)$ the metric ball of center $x$ and radius $r$. We will endow $\Omega$ with the topology induced by the metric.*

*Let $\mu$ be a positive regular Borel measure in $\Omega$.*

*Assume there exists an increasing sequence $\{\Omega_n\}_{n=1}^{\infty}$ of bounded measurable subsets of $\Omega$, such that:*

$$\bigcup_{n=1}^{\infty}\Omega_n = \Omega \tag{3.9}$$

*and such for, any $n = 1, 2, 3, ...$:*

*(i) the closure of $\Omega_n$ in $\Omega$ is compact;*

*(ii) there exists $\varepsilon_n > 0$ such that*

$$\left\{x \in \Omega : d(x, y) < 2\varepsilon_n \text{ for some } y \in \Omega_n\right\} \subset \Omega_{n+1}; \tag{3.10}$$

*(iii) there exists $C_n > 1$ such that for any $x \in \Omega_n, 0 < r \leq \varepsilon_n$ we have*

$$0 < \mu\left(B\left(x, 2r\right)\right) \leq C_n \mu\left(B\left(x, r\right)\right) < \infty. \tag{3.11}$$

*(Note that for $x \in \Omega_n$ and $r \leq \varepsilon_n$ we also have $B\left(x, 2r\right) \subset \Omega_{n+1}$).*

*We will say that $\left(\Omega, \{\Omega_n\}_{n=1}^{\infty}, d, \mu\right)$ is a* (metric) locally homogeneous space *if the above assumptions hold.*

Any space satisfying the above definition *a fortiori* satisfies the definition of locally homogeneous space given in [8]. In the following, we will recall the statements of several results proved in [8].

Next, we introduce the notion of local singular kernel.

**Assumption (K).** For fixed $\Omega_n, \Omega_{n+1}$, and a fixed ball $B\left(\overline{x}, R_0\right)$, with $\overline{x} \in \Omega_n$ and $R_0 < 2\varepsilon_n$ (hence $B\left(\overline{x}, R_0\right) \subset \Omega_{n+1}$), let $K\left(x, y\right)$ be a measurable function defined for $x, y \in B\left(\overline{x}, R_0\right)$, $x \neq y$. Let $R > 0$ be any number satisfying

$$cR \leq R_0 \tag{3.12}$$

for some $c > 1$; let $a, b \in C_0^{\alpha}\left(\Omega_{n+1}\right)$, $B\left(\overline{x}, c_1 R\right) \prec a \prec B\left(\overline{x}, c_2 R\right)$, $B\left(\overline{x}, c_3 R\right) \prec b \prec B\left(\overline{x}, c_4 R\right)$ for some fixed constants $c_i \in (0,1)$, $i = 1, ..., 4$ (the symbol $B_1 \prec f \prec B_2$ means that $f = 1$ in $B_1$, vanishes outside $B_2$, and takes values in $[0,1]$). The new kernel

$$\widetilde{K}\left(x, y\right) = a\left(x\right) K\left(x, y\right) b\left(y\right) \tag{3.13}$$

can be considered defined in the whole $\Omega_{n+1} \times \Omega_{n+1} \setminus \{x = y\}$.

We now list a series of possible assumptions on the kernel $K$ which will be recalled in the following theorems.

(i) We say that $K$ satisfies the *standard estimates* for some $\nu \in [0,1)$ if the following hold:

$$\left|K\left(x, y\right)\right| \leq \frac{A d\left(x, y\right)^{\nu}}{\mu\left(B\left(x, d\left(x, y\right)\right)\right)} \tag{3.14}$$

for $x, y \in B\left(\overline{x}, R_0\right)$, $x \neq y$, and

$$\left|K(x_0, y) - K(x, y)\right| + \left|K(y, x_0) - K(y, x)\right| \leq \frac{B d\left(x_0, y\right)^{\nu}}{\mu(B(x_0, d(x_0, y)))} \left(\frac{d(x_0, x)}{d(x_0, y)}\right)^{\beta} \tag{3.15}$$

for any $x_0, x, y \in B\left(\overline{x}, R_0\right)$ with $d(x_0, y) > 2d(x_0, x)$, some $\beta > 0$.

(ii) We say that $K$ satisfies the *cancellation property* if the following holds:

there exists $C > 0$ such that for a.e. $x \in B\left(\overline{x}, R_0\right)$ and every $\varepsilon_1, \varepsilon_2$ such that $0 < \varepsilon_1 < \varepsilon_2$ and $B_{\rho}\left(x, \varepsilon_2\right) \subset \Omega_{n+1}$

$$\left|\int_{\Omega_{n+1}, \varepsilon_1 < \rho(x,y) < \varepsilon_2} K(x, y)\, d\mu(y)\right| + \left|\int_{\Omega_{n+1}, \varepsilon_1 < \rho(x,z) < \varepsilon_2} K(z, x)\, d\mu(z)\right| \leq C, \tag{3.16}$$

where $\rho$ is any *quasidistance,* equivalent to $d$ in $\Omega_{n+1}$ and $B_\rho$ denotes $\rho$-balls. This means that $\rho$ satisfies the axioms of distance, except for the triangle inequality, which is replaced by the weaker

$$\rho(x,y) \leq c\left[\rho(x,z)+\rho(z,y)\right]$$

for any $x,y,z\in\Omega_{n+1}$ and some constant $c\geq 1$; moreover,

$$c_1 d(x,y) \leq \rho(x,y) \leq c_2 d(x,y)$$

for any $x,y$ and some positive constants $c_1,c_2$.

(iii) We say that $K$ satisfies the *convergence condition* if the following holds: for a.e. $x\in B(\overline{x},R_0)$ such that $B_\rho(x,R)\subset\Omega_{n+1}$ there exists

$$h_R(x) \equiv \lim_{\varepsilon\to 0}\int_{\Omega_{n+1},\varepsilon<\rho(x,y)<R} K(x,y)d\mu(y), \tag{3.17}$$

where $\rho$ is any quasidistance equivalent to $d$ in $\Omega_{n+1}$.

All the following results in this section have been proved in [8]. In some statements we have introduced some slight simplifications (with respect to [8]) due to the fact that our space is assumed to be metric.

**Theorem 3.11 ($L^p$ and $C^\eta$ estimates for singular integrals)** *Let $K,\widetilde{K}$ be as in Assumption (K), with $K$ satisfying the standard estimates (i) with $\nu=0$, the cancellation property (ii) and the convergence condition (iii) stated above. If*

$$Tf(x) = \lim_{\varepsilon\to 0}\int_{B(\overline{x},R),\rho(x,y)>\varepsilon} \widetilde{K}(x,y)f(y)\,d\mu(y),$$

*then for any $p\in(1,\infty)$*

$$\|Tf\|_{L^p(B(\overline{x},R))} \leq c\,\|f\|_{L^p(B(\overline{x},R))}\,.$$

*The constant $c$ depends on $p,n$ and the constants of $K$ involved in the assumptions (but not on $R$).*

*Moreover, $T$ satisfies a weak 1-1 estimate:*

$$\mu\left(\left\{x\in B(\overline{x},R): |Tf(x)|>t\right\}\right) \leq \frac{c}{t}\,\|f\|_{L^1(B(\overline{x},R))} \text{ for any } t>0.$$

*Assume that, in addition, the kernel $K$ satisfies the condition*

$$\widetilde{h}(x) \equiv \lim_{\varepsilon\to 0}\int_{\rho(x,y)>\varepsilon} \widetilde{K}(x,y)d\mu(y) \in C^\gamma(\Omega_{n+1}) \tag{3.18}$$

*for some $\gamma>0$ (where $\rho$ is the same appearing in the assumed convergence condition (iii)). Then*

$$\|Tf\|_{C^\eta(B(\overline{x},R))} \leq c\,\|f\|_{C^\eta(B(\overline{x},HR))} \tag{3.19}$$

*for any positive $\eta < \min(\alpha, \beta, \gamma)$ and some constant $H > 1$ independent of $R$. (Recall that $\alpha$ is the Hölder exponent related to the cutoff functions defining $\widetilde{K}, \beta$ appears in the standard estimates (i) and $\gamma$ is the number in (3.18)).*

*The constant $c$ depends on $\eta, n, R$, the constants involved in the assumptions on $K$, and the $C^\gamma$ norm of $\widetilde{h}$.*

**Remark 3.12 (Estimates for $C_0^\eta$ functions)** *Applying the Hölder continuity result to functions $f \in C_0^\eta(B(\overline{x}, r))$ with $r < R$ we can get a a bound*

$$\|Tf\|_{C^\eta(B(\overline{x},r))} \le c\,\|f\|_{C^\eta(B(\overline{x},r))}$$

*with $c$ depending on $R$ but not on $r$.*

**Theorem 3.13 ($L^p - L^q$ estimate for fractional integrals)** *Let $K, \widetilde{K}$ be as in Assumption (K), with $K$ satisfying the growth condition*

$$0 \le K(x, y) \le \frac{c}{\mu(B(x, d(x, y)))^{1-\nu}} \tag{3.20}$$

*for some $\nu \in (0, 1)$, $c > 0$, any $x, y \in B(\overline{x}, R_0)$, $x \ne y$. If*

$$I_\nu f(x) = \int_{B(\overline{x},R)} \widetilde{K}(x, y) f(y)\, d\mu(y)$$

*then, for any $p \in \left(1, \frac{1}{\nu}\right)$, $\frac{1}{q} = \frac{1}{p} - \nu$ there exists $c$ such that*

$$\|I_\nu f\|_{L^q(B(\overline{x},R))} \le c\,\|f\|_{L^p(B(\overline{x},R))}$$

*for any $f \in L^p(B(\overline{x}, R))$. The constant $c$ depends on $p, n$, and the constants of $K$ involved in the assumptions (but not on $R$).*

**Theorem 3.14 ($C^\eta$ estimate for fractional integrals)** *Let $K, \widetilde{K}$ be as in Assumption (K), with $K$ satisfying (3.14) and (3.15) for some $\nu \in (0, 1)$, $\beta > 0$. If*

$$I_\nu f(x) = \int_{B(\overline{x},R)} \widetilde{K}(x, y) f(y)\, d\mu(y),$$

*then, for any $\eta < \min(\alpha, \beta, \nu)$*

$$\|I_\nu f\|_{C^\eta(B(\overline{x},R))} \le c\,\|f\|_{C^\eta(B(\overline{x},HR))}.$$

*The constant $c$ depends on $\eta, n, R$ and the constants of $K$ involved in the assumptions; the number $H$ only depends on $n$.*

Reasoning as in Remark 3.12, we can also say that for functions $f \in C_0^\eta(B(\overline{x}, r))$ with $r < R$ the following bound holds

$$\|I_\nu f\|_{C^\eta(B(\overline{x},r))} \le c\,\|f\|_{C^\eta(B(\overline{x},r))}$$

with $c$ depending on $R$ but not on $r$.

To state the commutator theorems that we will need, we have first to recall the following

**Definition 3.15 (Local $BMO$ and $VMO$ spaces)** *Let $(\Omega, \{\Omega_n\}_{n=1}^{\infty}, d, \mu)$ be a locally homogeneous space. For any function $u \in L^1(\Omega_{n+1})$, and $r > 0$, with $r \le \varepsilon_n$, set*

$$\eta^*_{u,\Omega_n,\Omega_{n+1}}(r) = \sup_{t \le r} \sup_{x_0 \in \Omega_n} \frac{1}{\mu(B(x_0,t))} \int_{B(x_0,t)} |u(x) - u_B| \, d\mu(x),$$

*where $u_B = \mu(B(x_0,t))^{-1} \int_{B(x_0,t)} u$. We say that $u \in BMO_{loc}(\Omega_n, \Omega_{n+1})$ if*

$$\|u\|_{BMO_{loc}(\Omega_n,\Omega_{n+1})} = \sup_{r \le \varepsilon_n} \eta^*_{u,\Omega_n,\Omega_{n+1}}(r) < \infty.$$

*We say that $u \in VMO_{loc}(\Omega_n, \Omega_{n+1})$ if $u \in BMO_{loc}(\Omega_n, \Omega_{n+1})$ and*

$$\eta^*_{u,\Omega_n,\Omega_{n+1}}(r) \to 0 \text{ as } r \to 0.$$

*The function $\eta^*_{u,\Omega_n,\Omega_{n+1}}$ will be called $VMO$ local modulus of $u$ in $(\Omega_n, \Omega_{n+1})$.*

Note that in the previous definition we integrate $u$ over balls centered at points of $\Omega_n$ and enclosed in $\Omega_{n+1}$. This is a fairly natural definition if we want to avoid integrating over the *intersection* $B(x_0,t) \cap \Omega_n$.

**Theorem 3.16 (Commutators of local singular integrals)** *Let $K, \widetilde{K}$ be as in Assumption (K), with $K$ satisfying the standard estimates (i) with $\nu = 0$, the cancellation property (ii) and the convergence condition (iii). If*

$$Tf(x) = \lim_{\varepsilon \to 0} \int_{B(\overline{x},R), \rho(x,y) > \varepsilon} \widetilde{K}(x,y) f(y) \, d\mu(y)$$

*and, for $a \in BMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$, we set*

$$C_a f(x) = T(af)(x) - a(x) Tf(x),$$

*then for any $p \in (1,\infty)$ there exists $c > 0$ such that*

$$\|C_a f\|_{L^p(B(\overline{x},R))} \le c \|a\|_{BMO_{loc}(\Omega_{n+2},\Omega_{n+3})} \|f\|_{L^p(B(\overline{x},R))}.$$

*Moreover, if $a \in VMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$ for any $\varepsilon > 0$ there exists $r > 0$ such that for any $f \in L^p(B(\overline{x},r))$ we have*

$$\|C_a f\|_{L^p(B(\overline{x},r))} \le \varepsilon \|f\|_{L^p(B(\overline{x},r))}.$$

*The constant $c$ depends on $p, n$ and the constants of $K$ involved in the assumptions (but not on $R$); the constant $r$ also depends on the $VMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$ modulus of $a$.*

**Theorem 3.17 (Positive commutators of local fractional integrals)** *Let $K, \widetilde{K}$ be as in Assumption (K), with $K$ satisfying the growth condition (3.20) for some $\nu > 0$. If*

$$I_\nu f(x) = \int_{B(\overline{x},R)} \widetilde{K}(x,y) f(y) \, d\mu(y)$$

*and, for* $a \in BMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$*, we set*

$$C_{\nu,a}f(x) = \int_{B(\overline{x},R)} \widetilde{K}(x,y)\,|a(x) - a(y)|\,f(y)\,d\mu(y) \tag{3.21}$$

*then, for any* $p \in \left(1, \frac{1}{\nu}\right), \frac{1}{q} = \frac{1}{p} - \nu$ *there exists* $c$ *such that*

$$\|C_{\nu,a}f\|_{L^q(B(\overline{x},R))} \leq c\,\|a\|_{BMO_{loc}(\Omega_{n+2},\Omega_{n+3})}\,\|f\|_{L^p(B(\overline{x},R))}$$

*for any* $f \in L^p(B(\overline{x},R))$.

*Moreover, if* $a \in VMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$ *for any* $\varepsilon > 0$ *there exists* $r > 0$ *such that for any* $f \in L^p(B(\overline{x},r))$ *we have*

$$\|C_{\nu,a}f\|_{L^q(B(\overline{x},r))} \leq \varepsilon\,\|f\|_{L^p(B(\overline{x},r))}.$$

*The constant* $c$ *depends on* $p, \nu, n$ *and the constants involved in the assumptions on* $K$ *(but not on* $R$*); the constant* $r$ *also depends on the* $VMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$ *modulus of* $a$.

**Theorem 3.18 (Positive commutators of nonsingular integrals)** *Let* $K, \widetilde{K}$ *be as in Assumption (K), with* $K$ *satisfying condition (3.15) with* $\nu = 0$*. Assume that the operator*

$$Tf(x) = \int_{B(\overline{x},R)} \widetilde{K}(x,y)f(y)\,d\mu(y)$$

*is continuous on* $L^p(B(\overline{x},R))$ *for any* $p \in (1,\infty)$*. For* $a \in BMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$*, set*

$$C_a f(x) = \int_{B(\overline{x},R)} \widetilde{K}(x,y)\,|a(x) - a(y)|\,f(y)\,d\mu(y), \tag{3.22}$$

*then*

$$\|C_a f\|_{L^p(B(\overline{x},R))} \leq c\,\|a\|_{BMO_{loc}(\Omega_{n+2},\Omega_{n+3})}\,\|f\|_{L^p(B(\overline{x},R))}$$

*for any* $f \in L^p(B(\overline{x},R)), p \in (1,\infty)$.

*Moreover, if* $a \in VMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$ *for any* $\varepsilon > 0$ *there exists* $r > 0$ *such that for any* $f \in L^p(B(\overline{x},r))$ *we have*

$$\|C_a f\|_{L^p(B(\overline{x},r))} \leq \varepsilon\,\|f\|_{L^p(B(\overline{x},r))}.$$

*The constant* $c$ *depends on* $n$*, the constants involved in the assumptions on* $K$*, and the* $L^p$*-*$L^p$ *norm of the operator* $T$ *(but not explicitly on* $R$*); the constant* $r$ *also depends on the* $VMO_{loc}(\Omega_{n+2}, \Omega_{n+3})$ *modulus of* $a$.

**Remark 3.19** *In the statements of Theorems 3.11, 3.13, 3.14, 3.16, 3.17, 3.18 we wrote that the constant depends on the kernel only through the constants involved in the assumptions. In the following we will need some additional information about this dependence. A standard sublinearity argument allows us to say that if, for example, our assumptions on the kernel are (3.14), (3.15), (3.16), then the constant in our upper bound will have the form*

$$c \cdot (A + B + C)$$

*where* $A, B, C$ *are the constants appearing in (3.14), (3.15), (3.16), and* $c$ *does not depend on the kernel.*

We will also need the notion of *local maximal operator* in locally homogeneous spaces.

**Definition 3.20** *Fix* $\Omega_n, \Omega_{n+1}$ *and, for any* $f \in L^1(\Omega_{n+1})$ *define the* local maximal function

$$M_{\Omega_n,\Omega_{n+1}} f(x) = \sup_{r \le r_n} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y)| \, d\mu(y) \quad \text{for } x \in \Omega_n$$

*where* $r_n = 2\varepsilon_n/5$.

**Theorem 3.21** *Let* $f$ *be a measurable function defined on* $\Omega_{n+1}$. *The following hold:*

*(a) If* $f \in L^p(\Omega_{n+1})$ *for some* $p \in [1,\infty]$, *then* $M_{\Omega_n,\Omega_{n+1}} f$ *is finite almost everywhere in* $\Omega_n$*;*

*(b) if* $f \in L^1(\Omega_{n+1})$, *then for every* $t > 0$,

$$\mu\left(\left\{x \in \Omega_n : \left(M_{\Omega_n,\Omega_{n+1}} f\right)(x) > t\right\}\right) \le \frac{c_n}{t} \int_{\Omega_{n+1}} |f(y)| \, d\mu(y);$$

*(c) if* $f \in L^p(\Omega_{n+1})$, $1 < p \le \infty$, *then* $M_{\Omega_n,\Omega_{n+1}} f \in L^p(\Omega_n)$ *and*

$$\left\| M_{\Omega_n,\Omega_{n+1}} f \right\|_{L^p(\Omega_n)} \le c_{n,p} \|f\|_{L^p(\Omega_{n+1})}.$$

Finally, we need to discuss an integral characterization of Hölder continuous, analogous to the one classically introduced by Campanato [10], in our abstract and local setting.

**Definition 3.22 (Local Campanato spaces)** *Let* $(\Omega, \{\Omega_n\}_{n=1}^{\infty}, d, \mu)$ *be a locally homogeneous space. For any function* $u \in L^1(\Omega_{n+1})$, $\alpha \in (0,1)$, *let*

$$M_{\alpha,\Omega_n,\Omega_{n+1}} u = \sup_{x \in \Omega_n, r \le \varepsilon_n} \inf_{c \in \mathbb{R}} \frac{1}{r^\alpha |B(x,r)|} \int_{B(x,r)} |u(y) - c| \, d\mu(y).$$

*Set*

$$\mathcal{L}^\alpha(\Omega_n, \Omega_{n+1}) = \left\{ u \in L^1(\Omega_{n+1}) : M_{\alpha,\Omega_n,\Omega_{n+1}} u < \infty \right\}.$$

If $u \in C^\alpha(\Omega_{n+1})$ then clearly

$$M_{\alpha,\Omega_n,\Omega_{n+1}} u \le |u|_{C^\alpha(\Omega_{n+1})}.$$

A converse result is contained in the following:

**Theorem 3.23** *For any* $u \in \mathcal{L}^\alpha(\Omega_n, \Omega_{n+1})$, *there exists a function* $u^*$, *equal to* $u$ *a.e. in* $\Omega_n$,*such that* $u^*$ *belongs to* $C^\alpha(\Omega_n)$. *Namely, for any* $x, y \in \Omega_n$ *with* $2d(x,y) \le \varepsilon_n$ *we have*

$$|u^*(x) - u^*(y)| \le c M_{\alpha,\Omega_n,\Omega_{n+1}} u \, d(x,y)^\alpha. \tag{3.23}$$

*If* $2d(x,y) > \varepsilon_n$ *then*

$$|u^*(x) - u^*(y)| \le c \left\{ M_{\alpha,\Omega_n,\Omega_{n+1}} u + \|u\|_{L^1(\Omega_{n+1})} \right\} d(x,y)^\alpha. \tag{3.24}$$

*The constant* $c$ *in (3.23), (3.24) depends on* $C_n$ *but not on* $\varepsilon_n$.

**Application of the abstract theory to our setting**

Letš now explain the way how this abstract setting will be used to describe our concrete situation. The a-priori estimates we will prove in Theorems 2.1, 2.2, involve a fixed subdomain $\Omega' \Subset \Omega$. Fix once and for all this $\Omega'$. For any $\overline{x} \in \Omega'$ we can perform in a suitable neighborhood of $\overline{x}$ the lifting and approximation procedure as explained in § 3.1. Let $\overline{\xi} = (\overline{x}, 0) \in \mathbb{R}^N$ and $\widetilde{B}\left(\overline{\xi}, R\right)$ be as in Lemma 3.9. We can then choose

$$\widetilde{\Omega} = \widetilde{B}\left(\overline{\xi}, R\right); \widetilde{\Omega}_k = \widetilde{B}\left(\overline{\xi}, \frac{kR}{k+1}\right) \text{ for } k = 1, 2, 3, ...$$

By the properties of $d_{\widetilde{X}}$ that we have listed in § 3.2, and particularly Theorem 3.7,

$$\left(\widetilde{\Omega}, \left\{\widetilde{\Omega}_k\right\}_{k=1}^{\infty}, d_{\widetilde{X}}, d\xi\right)$$

is a metric locally homogeneous space. The function $\rho(\xi, \eta) = \|\Theta(\eta, \xi)\|$ will play the role of the quasidistance appearing in conditions (3.16) and (3.17), in view of Lemma 3.9. This will be the basic setting where we will apply singular integral estimates.

In the space of the original variables $(\Omega, d_X, dx)$, instead, we will not apply singular integral estimates, but we will use again the local doubling condition, when we will establish some important properties of function spaces $C^\alpha$ and $VMO$ (see § 3.4). Note that, if $\Omega_k$ is an increasing sequence of domains with $\Omega_k \Subset \Omega_{k+1} \Subset \Omega$, we can say that

$$(\Omega, \{\Omega_k\}_k, d_X, dx)$$

is a metric locally homogeneous space.

## 3.4 Function spaces

The aim of this section is twofold. First, we want to define the basic function spaces we will need and point out their main properties; second, we want to find a relation between function spaces defined over a ball $B(\overline{x}, r) \subset \Omega \subset \mathbb{R}^n$ and on the corresponding lifted ball $\widetilde{B}\left(\overline{\xi}, r\right) \subset \mathbb{R}^N$. More precisely, we need to know that $f(x)$ belongs to some function space on $B$ if and only $\widetilde{f}(x, h) = f(x)$ belongs to the analogous function space on $\widetilde{B}$. This last fact relies on the following known result (see [23, Lemmas 3.1 and 3.2, p. 139]):

**Theorem 3.24** *Let us denote by $B, \widetilde{B}$ the balls defined with respect to $d_X$ in $\Omega$ and $d_{\widetilde{X}}$ in $\widetilde{\Omega}$, respectively. There exist constants $\delta_0 \in (0,1), r_0, c_1, c_2 > 0$ such that*

$$\begin{aligned} c_1 vol\left(\widetilde{B}_r(x,h)\right) &\le vol\left(B_r(x)\right) \cdot vol\left\{h' \in \mathbb{R}^{N-n} : (z, h') \in \widetilde{B}_r(x,h)\right\} \\ &\le c_2 vol\left(\widetilde{B}_r(x,h)\right) \end{aligned} \quad (3.25)$$

*for every $x \in \Omega, z \in B_{\delta_0 r}(x)$ and $r \leq r_0$. (Here "vol" stands for the Lebesgue measure in the appropriate dimension, $x$ denotes a point in $\mathbb{R}^n$ and $h$ a point in $\mathbb{R}^{N-n}$). More precisely, the condition $z \in B_{\delta_0 r}(x)$ is needed only for the validity of the first inequality in (3.25). Moreover:*

$$d_{\widetilde{X}}\left((x,h),(x',h')\right) \geq d_X(x,x'). \tag{3.26}$$

*Finally, the projection of the lifted ball $\widetilde{B}_r(x,h)$ on $\mathbb{R}^n$ is just the ball $B(x,r)$, and this projection is onto.*

A consequence of the above theorem is the following

**Corollary 3.25** *For any positive function $g$ defined in $B_r(x) \subset \Omega, r \leq r_0$, one has*

$$\frac{c_1}{|B_{\delta_0 r}(x)|}\int_{B_{\delta_0 r}(x)} g(y)\,dy \leq \frac{1}{\left|\widetilde{B}_r(x,h)\right|}\int_{\widetilde{B}_r(x,h)} g(y)\,dydh' \leq \frac{c_2}{|B_r(x)|}\int_{B_r(x)} g(y)\,dy. \tag{3.27}$$

*where $\delta_0$ is the constant in Theorem 3.24.*

**Proof.** By (3.25) and the locally doubling condition, we have, for some fixed $\delta_0 < 1$ as in Theorem 3.24,

$$\begin{aligned}
&\frac{1}{\left|\widetilde{B}_r(x,h)\right|}\int_{\widetilde{B}_r(x,h)} g(y)\,dydh' \\
&= \frac{1}{\left|\widetilde{B}_r(x,h)\right|}\int_{B_r(x)} g(y)\,dy \int_{h'\in\mathbb{R}^{N-n}:(y,h')\in\widetilde{B}_r(x,h)} dh' \\
&\geq \frac{c_1}{\left|\widetilde{B}_r(x,h)\right|}\int_{B_{\delta_0 r}(x)} \frac{\left|\widetilde{B}_r(x,h)\right|}{|B_r(x)|} g(y)\,dy \\
&\geq \frac{c}{|B_{\delta_0 r}(x)|}\int_{B_{\delta_0 r}(x)} g(y)\,dy
\end{aligned}$$

where in the last inequality we exploited the doubling condition $|B_r(x)| \leq c\,|B_{\delta_0 r}(x)|$, which holds because $B_r(x) \subset \Omega$ and $r \leq r_0$. The proof of the second inequality in (3.27) is analogous but easier, since it involves the second inequality in (3.25), which does not require the condition $y \in B_{\delta_0 r}(x)$. ■

### 3.4.1 Hölder spaces

**Definition 3.26 (Hölder spaces)** *For any $0 < \alpha < 1, u : \Omega \to \mathbb{R}$, let:*

$$\begin{aligned}
|u|_{C_X^\alpha(\Omega)} &= \sup\left\{\frac{|u(x)-u(y)|}{d_X(x,y)^\alpha} : x,y \in \Omega, x \neq y\right\}, \\
\|u\|_{C_X^\alpha(\Omega)} &= |u|_{C^\alpha(\Omega)} + \|u\|_{L^\infty(\Omega)}, \\
C_X^\alpha(\Omega) &= \left\{u : \Omega \to \mathbb{R} : \|u\|_{C^\alpha(\Omega)} < \infty\right\}.
\end{aligned}$$

*For any positive integer k and* $0 < \alpha < 0$*, let*

$$C_X^{k,\alpha}(\Omega) = \left\{ u : \Omega \to \mathbb{R} : \|u\|_{C^{k,\alpha}(\Omega)} < \infty \right\},$$

*with*

$$\|u\|_{C_X^{k,\alpha}(\Omega)} = \sum_{|I|=1}^{k} \sum_{j_i=0}^{q} \|X_{j_1} \dots X_{j_l} u\|_{C^\alpha(\Omega)} + \|u\|_{C^\alpha(\Omega)}$$

*where* $I = (j_1, j_2, ..., j_l)$.

*We will set* $C_{X,0}^{\alpha}(\Omega)$ *and* $C_{X,0}^{k,\alpha}(\Omega)$ *for the subspaces of* $C_X^\alpha(\Omega)$ *and* $C_X^{k,\alpha}(\Omega)$ *of functions which are compactly supported in* $\Omega$*, and set* $C_{\widetilde{X}}^{\alpha}\left(\widetilde{\Omega}\right)$, $C_{\widetilde{X}}^{k,\alpha}\left(\widetilde{\Omega}\right)$, $C_{\widetilde{X},0}^{\alpha}\left(\widetilde{\Omega}\right)$ *and* $C_{\widetilde{X},0}^{k,\alpha}\left(\widetilde{\Omega}\right)$ *for the analogous function spaces over* $\widetilde{\Omega}$ *defined by the* $\widetilde{X}_i$*'s.*

*We will also write* $C_X^{k,0}(\Omega)$ *to denote the space of functions with continuous* $X$*-derivatives up to weight* $k$*.*

*Finally, whenever there is no risk of confusion, we will drop the index* $X$*, writing* $C^\alpha(\Omega)$ *instead of* $C_X^\alpha(\Omega)$*, and so on.*

The next Proposition, adapted from [4, Proposition 4.2], collects some properties of $C^\alpha$ functions which will be useful later. We will apply these properties mainly in the context of lifted variables, that is for the vector fields $\widetilde{X}_i$ on a ball $\widetilde{B}\left(\overline{\xi}, R\right)$.

**Proposition 3.27** *Let* $B(\overline{x}, 2R)$ *be a fixed ball where the vector fields* $X_i$ *and the control distance* $d$ *are well defined.*

*(i) For any* $\delta \in (0,1)$*, for any* $f \in C^1(B(\overline{x}, (1+\delta)R))$*, one has*

$$|f(x) - f(y)| \le \frac{c}{\delta} d_X(x,y) \left( \sum_{i=1}^{q} \sup_{B(\overline{x},(1+\delta)R)} |X_i f| + d_X(x,y) \sup_{B(\overline{x},(1+\delta)R)} |X_0 f| \right) \tag{3.28}$$

*for any* $x, y \in B(\overline{x}, R)$.

*If* $f \in C_0^1(B(\overline{x}, R))$*, one can simply write, for any* $x, y \in B(\overline{x}, R)$*,*

$$|f(x) - f(y)| \le c d_X(x,y) \left( \sum_{i=1}^{q} \sup_{B(\overline{x},R)} |X_i f| + d_X(x,y) \sup_{B(\overline{x},R)} |X_0 f| \right). \tag{3.29}$$

*In particular, for* $f \in C_0^1(B(\overline{x}, R)$*,*

$$|f|_{C^\alpha(B(\overline{x},R))} \le c R^{1-\alpha} \cdot \left( \sum_{i=1}^{q} \sup_{B(\overline{x},R)} |X_i f| + R \sup_{B(\overline{x},R)} |X_0 f| \right). \tag{3.30}$$

*Here* $C^1$ *(and* $C_0^1$ *) stands for the classical space of (compactly supported) continuously differentiable functions. The assumption* $f \in C^1$ *(or* $C_0^1$ *) can be replaced by* $f \in C_X^2$ *(or* $C_{X,0}^2$*, respectively).*

*(ii) For any couple of functions $f, g \in C_X^\alpha(B(\overline{x}, R))$, one has*

$$|fg|_{C_X^\alpha(B(\overline{x},R))} \leq |f|_{C_X^\alpha(B(\overline{x},R))} \|g\|_{L^\infty(B(\overline{x},R))} + |g|_{C_X^\alpha(B(\overline{x},R))} \|f\|_{L^\infty(B(\overline{x},R))}$$

*and*

$$\|fg\|_{C_X^\alpha(B(\overline{x},R))} \leq 2 \|f\|_{C_X^\alpha(B(\overline{x},R))} \|g\|_{C_X^\alpha(B(\overline{x},R))} . \tag{3.31}$$

*Moreover, if both $f$ and $g$ vanish at least at a point of $B(\overline{x}, R)$, then*

$$|fg|_{C_X^\alpha(B(\overline{x},R))} \leq cR^\alpha |f|_{C_X^\alpha(B(\overline{x},R))} |g|_{C_X^\alpha(B(\overline{x},R))} . \tag{3.32}$$

*(iii) Let $B(x_i, r)$ $(i = 1, 2, \cdots, k)$ be a finite family of balls of the same radius $r$, such that $\cup_{i=1}^k B(x_i, 2r) \subset \Omega$. Then for any $f \in C_X^\alpha(\Omega)$,*

$$\|f\|_{C_X^\alpha(\cup_{i=1}^k B(x_i,r))} \leq c \sum_{i=1}^k \|f\|_{C_X^\alpha(B(x_i,2r))} \tag{3.33}$$

*with $c$ depending on the family of balls, but not on $f$.*

*(iv) There exists $r_0 > 0$ such that for any $f \in C_{X,0}^{2,\alpha}(B(\overline{x}, R))$ and $0 < r \leq r_0$, we have the following interpolation inequality:*

$$\|X_0 f\|_{L^\infty(B(\overline{x},R))} \leq r^{\alpha/2} |X_0 f|_{C_X^\alpha(B(\overline{x},R))} + \frac{2}{r} \|f\|_{L^\infty(B(\overline{x},R))} . \tag{3.34}$$

**Proof.** The proof for (ii)-(iii) is similar to that in [4, Proposition 4.2], hence we will only prove (i) and (iv).

Throughout the proof we will write $d$ for $d_X$.

(i) Fix $\delta \in (0, 1)$ and let $R' = (1 + \delta) R$. Let us distinguish two cases:

(a) $d(x, y) < R' - \max(d(\overline{x}, x), d(\overline{x}, y))$. Let $\varepsilon > 0$ such that also

$$d(x, y) + \varepsilon < R' - \max(d(\overline{x}, x), d(\overline{x}, y)), \tag{3.35}$$

hence by Definition 3.6 there exists a curve $\varphi(t)$, such that $\varphi(0) = x, \varphi(1) = y$, and

$$\varphi'(t) = \sum_{i=0}^q \lambda_i(t) (X_i)_{\varphi(t)}$$

with $|\lambda_i(t)| \leq (d(x, y) + \varepsilon), |\lambda_0(t)| \leq (d(x, y) + \varepsilon)^2$ for $i = 1, \ldots q$. By (3.35),

$$B(x, d(x, y) + \varepsilon) \subset B(\overline{x}, R')$$

hence every point $\gamma(t)$ for $t \in (0, 1)$ belongs to $B(\overline{x}, R')$. Then we can write:

$$\begin{aligned} |f(x) - f(y)| &= \left| \int_0^1 \frac{d}{dt} f(\varphi(t)) \, dt \right| = \left| \int_0^1 \sum_{i=0}^q \lambda_i(t) (X_i f)_{\varphi(t)} \, dt \right| \\ &\leq (d(x, y) + \varepsilon) \sum_{i=1}^q \sup_{B(\overline{x},R')} |X_i f| + (d(x, y) + \varepsilon)^2 \sup_{B(\overline{x},R')} |X_0 f|, \end{aligned}$$

and since $\varepsilon$ is arbitrary this implies

$$|f(x)-f(y)| \leq d(x,y)\left\{\sum_{i=1}^{q}\sup_{B(\overline{x},R')}|X_i f| + d(x,y)\sup_{B(\overline{x},R')}|X_0 f|\right\}.$$

Note that the above argument relies on the differentiability of $f$ along the curve $\varphi$, which holds under either the assumption $f \in C^1(B(\overline{x},(1+\delta)R))$ or $f \in C_X^2(B(\overline{x},(1+\delta)R))$ (since $X_0$ has weight two).

(b) Let now $d(x,y) \geq R' - \max(d(\overline{x},x), d(\overline{x},y))$, and let us write

$$|f(x)-f(y)| \leq |f(x)-f(\overline{x})| + |f(\overline{x})-f(y)| = A+B.$$

Each of the terms $A, B$ can be bounded by an argument similar to that in case (a) (since both $x$ and $y$ can be joined to $\overline{x}$ by curves contained in $B(\overline{x},R)$), getting

$$|f(x)-f(y)| \leq [d(x,\overline{x})+d(y,\overline{x})]\left\{\sum_{i=1}^{q}\sup_{B(\overline{x},R)}|X_i f| + [d(x,\overline{x})+d(y,\overline{x})]\sup_{B(\overline{x},R)}|X_0 f|\right\}.$$

Now it is enough to show that

$$d(x,\overline{x})+d(y,\overline{x}) \leq \frac{c}{\delta}d(x,y).$$

To show this, let $r \equiv \max(d(\overline{x},x), d(\overline{x},y))$. Then:

$$d(x,\overline{x})+d(y,\overline{x}) \leq 2r \leq \frac{2}{\delta}(R'-r) \leq \frac{2}{\delta}d(x,y)$$

where the second inequality holds since $r < R$ and $R' = (1+\delta)R$, and the last inequality is assumption (b). This completes the proof of (3.28), which immediately implies (3.29) and (3.30).

Let us now prove (vi). Let $f \in C_{X,0}^{2,\alpha}(B(\overline{x},R))$. For any $x \in B(\overline{x},R)$, let $\gamma(t)$ be the curve such that

$$\gamma'(t) = (X_0)_{\gamma(t)},\ \gamma(0) = x.$$

This $\gamma(t)$ will be defined at least for $t \in [0,r_0]$ where $r_0 > 0$ is a number only depending on $B(\overline{x},R)$ and $X_0$. Then, for any $r \in (0,r_0)$ we can write, for some $\theta \in (0,1)$:

$$\begin{aligned}
(X_0 f)(x) &= (X_0 f)(\gamma(0)) = \frac{d}{dt}[f(\gamma(t))]_{t=0} \\
&= \frac{d}{dt}[f(\gamma(t))]_{t=0} - [f(\gamma(r)) - f(\gamma(0))] + [f(\gamma(r)) - f(\gamma(0))] \\
&= \frac{d}{dt}[f(\gamma(t))]_{t=0} - r\frac{d}{dt}[f(\gamma(t))]_{t=\theta r} + [f(\gamma(r)) - f(\gamma(0))] \\
&= \frac{d}{dt}[f(\gamma(t))]_{t=0}(1-r) + r\left(\frac{d}{dt}[f(\gamma(t))]_{t=0} - \frac{d}{dt}[f(\gamma(t))]_{t=\theta r}\right) + \\
&\quad + [f(\gamma(r)) - f(\gamma(0))] \\
&= (1-r)(X_0 f)(x) + r[(X_0 f)(\gamma(0)) - (X_0 f)(\gamma(\theta r))] \\
&\quad + [f(\gamma(r)) - f(\gamma(0))],
\end{aligned}$$

hence

$$\begin{aligned} r\left|\left(X_0 f\right)(x)\right| &\leq r\left|\left(X_0 f\right)(\gamma(0))-\left(X_0 f\right)(\gamma(\theta r))\right|+2\|f\|_{L^{\infty}} \\ &= r \frac{(\theta r)^{\alpha / 2}\left|\left(X_0 f\right)(\gamma(0))-\left(X_0 f\right)(\gamma(\theta r))\right|}{(\theta r)^{\alpha / 2}}+2\|f\|_{L^{\infty}} . \end{aligned}$$

Since, by definition of $\gamma$ and $d$, $d(\gamma(0), \gamma(\theta r)) \leq (\theta r)^{1/2}$,

$$\begin{aligned} \left|\left(X_0 f\right)(x)\right| &\leq (\theta r)^{\alpha / 2}\left|X_0 f\right|_{C_X^\alpha(B(\overline{x}, R))}+\frac{2}{r}\|f\|_{L^{\infty}(B(\overline{x}, R))} \\ &\leq r^{\alpha / 2}\left|X_0 f\right|_{C_X^\alpha(B(\overline{x}, R))}+\frac{2}{r}\|f\|_{L^{\infty}(B(\overline{x}, R))}, \end{aligned}$$

and we are done. ■

Next, we are going to study the relation between the spaces $C_X^\alpha(B_R)$ and $C_{\widetilde{X}}^\alpha\left(\widetilde{B}_R\right)$.

**Proposition 3.28** *Let $\widetilde{B}\left(\overline{\xi}, R\right)$ be a lifted ball (as described at the end of § 3.3), with $\overline{\xi}=(\overline{x}, 0)$. If $f$ is a function defined in $B(\overline{x}, R)$ and $\widetilde{f}(x, h)=f(x)$ is regarded as a function defined on $\widetilde{B}_R\left(\overline{\xi}, R\right)$, then the following inequalities hold (whenever the right-hand side is finite):*

$$\begin{aligned} \left|\widetilde{f}\right|_{C_{\widetilde{X}}^\alpha\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)} &\leq |f|_{C_X^\alpha(B(\overline{x}, R))}, \\ |f|_{C_X^\alpha(B(\overline{x}, s))} &\leq \frac{c}{(t-s)^2}\left|\widetilde{f}\right|_{C_{\widetilde{X}}^\alpha\left(\widetilde{B}\left(\overline{\xi}, t\right)\right)} \text{ for } 0<s<t<R \end{aligned} \tag{3.36}$$

*where $c$ also depends on $R$. Moreover,*

$$\left|\widetilde{X}_{i_1} \widetilde{X}_{i_2} \cdots \widetilde{X}_{i_k} \widetilde{f}\right|_{C_{\widetilde{X}}^\alpha\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)} \leq\left|X_{i_1} X_{i_2} \cdots X_{i_k} f\right|_{C_X^\alpha(B(\overline{x}, R))}, \tag{3.37}$$

$$\left|X_{i_1} X_{i_2} \cdots X_{i_k} f\right|_{C_X^\alpha(B(\overline{x}, s))} \leq \frac{c}{(t-s)^2}\left|\widetilde{X}_{i_1} \widetilde{X}_{i_2} \cdots \widetilde{X}_{i_k} \widetilde{f}\right|_{C_{\widetilde{X}}^\alpha\left(\widetilde{B}\left(\overline{\xi}, t\right)\right)} \tag{3.38}$$

*for $0<s<t<R$ and $i_j=0,1,2, \ldots, q$.*

As already done in [4, Proposition 8.3], to prove the above relation between Hölder spaces over $B$ and $\widetilde{B}$ we have to exploit an equivalent integral characterization of Hölder continuous functions, analogous to the one established in the classical case by Campanato in [10]. However, to avoid integration over sets of the kind $\Omega \cap B(x, r)$ (with the related problem of assuring a suitable doubling condition) we need to apply the local version of this result which has been established in [8] and recalled in § 3.3. We are going to apply Definition 3.22 in our context.

**Definition 3.29** *For $\overline{x} \in \Omega', B(\overline{x}, R) \subset \Omega, f \in L^1(B(\overline{x}, R)), \alpha \in (0,1)$, $0< s<t \leq 1$, let*

$$M_{\alpha, B_{sR}, B_{tR}}(f)=\sup_{x \in B(\overline{x}, sR), r \leq(t-s) R} \inf_{c \in \mathbb{R}} \frac{1}{r^\alpha\left|B_r(x)\right|} \int_{B_r(x)}|f(y)-c| dy .$$

If $f \in C_X^\alpha (B(\overline{x}, R))$ then

$$M_{\alpha, B_{sR}, B_{tR}}(f) \leq |f|_{C^\alpha(B_R(x_0))}.$$

Moreover:

**Lemma 3.30** *For $\overline{x} \in \Omega', B(\overline{x}, 2R_0) \subset \Omega,\ R < R_0,\ \alpha \in (0,1),\ 0 < s < t \leq 1$, if $f \in L^1(B(\overline{x}, tR))$ is a function such that $M_{\alpha, B_{sR}, B_{tR}}(f) < \infty$, then there exists a function $f^*$, a.e. equal to $f$, such that $f^* \in C_X^\alpha(B(\overline{x}, sR))$ and*

$$|f^*|_{C_X^\alpha(B(\overline{x}, sR))} \leq \frac{c}{(t-s)^2} M_{\alpha, B_{sR}, B_{tR}}(f)$$

*for some $c$ independent of $f, s, t$.*

**Proof.** We can apply Theorem 3.23 choosing $\Omega_k = B(\overline{x}, sR), \Omega_{k+1} = B(\overline{x}, tR), \varepsilon_n = R(t-s)$. The locally doubling constant can be chosen independently of $R$, since $B(\overline{x}, 2R_0) \subset \Omega,\ R < R_0$. We conclude there exists a function $f^*$, a.e. equal to $f$, such that

$$|f^*(x) - f^*(y)| \leq c M_{\alpha, B_{sR}, B_{tR}}(f)\, d_X(x,y)^\alpha$$

for any $x, y \in B(\overline{x}, sR)$ with $d_X(x,y) \leq R(t-s)/2$

If now $x, y$ are any two points in $B_{sR}(x_0)$, and $r = d_X(x,y)$, by Lemma 3.8 we can find $n+1$ points $x_0 = x, x_1, x_2, ..., x_n = y$ in $B_{sR}(x_0)$ such that

$$d_X(x_i, x_{i-1}) \leq \frac{r}{\sqrt{n}}.$$

Let $n$ be the least integer such that $\frac{r}{\sqrt{n}} \leq R(t-s)/2$, then

$$\begin{aligned} |f^*(x) - f^*(y)| &\leq \sum_{i=1}^n |f^*(x_i) - f^*(x_{i-1})| \leq \sum_{i=1}^n c M_{\alpha, B_{sR}, B_{tR}}(f)\, d_X(x_i, x_{i-1})^\alpha \\ &\leq nc M_{\alpha, B_{sR}, B_{tR}}(f)\, d_X(x,y)^\alpha. \end{aligned}$$

Let us find an upper bound on $n$. We know that

$$\sqrt{n} \leq c \frac{d_X(x,y)}{R(t-s)} \leq \frac{c}{t-s}$$

since $d_X(x,y) \leq 2R$ for $x, y \in B_{tR}(x_0)$. Hence $n \leq c/(t-s)^2$ and the lemma is proved. ■

**Proof of Proposition 3.28.** The first inequality immediately follows by (3.26), so let us prove the second one.

Let $0 < s < t < 1$, $x \in B(\overline{x}, \delta_0 sR)$, where $\delta_0$ is the number in Theorem 3.24, $r \leq R(t-s)$, $\overline{\xi} = (\overline{x}, 0)$. Since the projection $\pi : \widetilde{B}((x,s), \delta) \to B(x, \delta)$ is onto (see Theorem 3.24), there exists $h \in \mathbb{R}^{N-n}$ such that $\xi = (x,h) \in \widetilde{B}\left(\overline{\xi}, \delta_0 sR\right)$. Then we have the following inequalities:

$$\frac{1}{r^{\alpha}} \frac{c}{\left|B_{\delta_0 r}(x)\right|} \int_{B_{\delta_0 r}(x)} |f(y) - k|\, dy$$

(by Corollary 3.25)

$$\leq \frac{c}{r^{\alpha}} \frac{1}{\left|\widetilde{B}(\xi, r)\right|} \int_{\widetilde{B}(\xi,r)} \left|\widetilde{f}(\eta) - k\right| d\eta$$

choosing $k = f(x) = \widetilde{f}(\xi)$

$$\leq \frac{c}{r^{\alpha}} \left|\widetilde{f}\right|_{C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}(\xi,r)\right)} r^{\alpha} = c \left|\widetilde{f}\right|_{C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}(\xi,r)\right)}. \tag{3.39}$$

Since $r \leq R(t-s)$ and $d\left(\xi, \overline{\xi}\right) < \delta_0 sR$, we have the inclusion

$$\widetilde{B}(\xi, r) \subset \widetilde{B}\left(\overline{\xi}, \delta_0 sR + R(t-s)\right) \equiv \widetilde{B}\left(\overline{\xi}, R'\right)$$

so that (3.39) implies

$$M_{\alpha, B(\overline{x}, \delta_0 sR), B(\overline{x}, \delta_0 tR)}(f) \leq c \left|\widetilde{f}\right|_{C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, R'\right)\right)},$$

and by Lemma 3.30, we conclude

$$|f^*|_{C^{\alpha}_{X}(B(\overline{x}, \delta_0 sR))} \leq \frac{c}{(t-s)^2} \left|\widetilde{f}\right|_{C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, R'\right)\right)}.$$

Note that $R' - \delta_0 sR = R(t-s)$, hence changing our notation as

$$\delta_0 sR = s'$$
$$R' = t'$$

we get

$$|f^*|_{C^{\alpha}_{X}(B(\overline{x}, s'))} \leq \frac{c}{(t'-s')^2} \left|\widetilde{f}\right|_{C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, t'\right)\right)}$$

for $0 < s' < t' < R$, with $c$ also depending on $R$. This is (3.36).

Now, inequalities (3.37) and (3.38) are also consequences of what we have proved because $\widetilde{X}_i \widetilde{f} = \widetilde{X_i f}$, hence the same reasoning can be iterated to higher order derivatives. ■

### 3.4.2 $L^p$ and Sobolev spaces

We are going to define the Sobolev spaces $S^{k,p}_X(\Omega)$ in the present context as in [25].

**Definition 3.31 (Sobolev spaces)** *If $X = (X_0, X_1, \ldots, X_q)$ is any system of smooth vector fields satisfying Hörmander's condition in a domain $\Omega \subset \mathbb{R}^n$, the*

*Sobolev space* $S_X^{k,p}(\Omega)$ *($1 \leq p \leq \infty, k$ positive integer) consists of $L^p$-functions with $k$ (weighted) derivatives with respect to the vector fields $X_i$'s, in $L^p$. Explicitly,*

$$\|u\|_{S_X^{k,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \sum_{i=1}^{k} \left\|D^i u\right\|_{L^p(\Omega)}, \text{ where}$$

$$\left\|D^k u\right\|_{L^p(\Omega)} = \sum_{|I|=k} \|X_I u\|_{L^p(\Omega)}.$$

*Also, we can define the spaces of functions vanishing at the boundary saying that $u \in S_{0,X}^{k,p}(\Omega)$ if there exists a sequence $\{u_k\}$ of $C_0^\infty(\Omega)$ functions converging to $u$ in $S_X^{k,p}(\Omega)$. Similarly, we can define the Sobolev spaces $S_{\widetilde{X}}^{k,p}\left(\widetilde{B}\right)$, $S_{\widetilde{X},0}^{k,p}\left(\widetilde{B}\right)$ over a lifted ball $\widetilde{B}$, induced by the $\widetilde{X}$'s.*

It can be proved (see [3, Proposition 3.5]) that:

**Proposition 3.32** *If $u \in S_X^{2,p}(\Omega)$ and $\varphi \in C_0^\infty(\Omega)$, then $u\varphi \in S_{0,X}^{2,p}(\Omega)$, and an analogous property holds for the space $S_{0,\widetilde{X}}^{2,p}\left(\widetilde{B}\right)$.*

Moreover:

**Theorem 3.33** *Let $f \in L^p(B(x,r))$, $\widetilde{f}(x,h) = f(x)$, $\widetilde{B}(\xi,r)$ be the lifted ball of $B(x,r)$, with $\xi = (x,0) \in \mathbb{R}^N$. Then*

$$c_1 \|f\|_{L^p(B(x,\delta_0 r))} \leq \left\|\widetilde{f}\right\|_{L^p\left(\widetilde{B}(\xi,r)\right)} \leq c_2 \|f\|_{L^p(B(x,r))}$$

$$c_1 \|f\|_{S_X^{2,p}(B(x,\delta_0 r))} \leq \left\|\widetilde{f}\right\|_{S_{\widetilde{X}}^{2,p}\left(\widetilde{B}(\xi,r)\right)} \leq c_2 \|f\|_{S_X^{2,p}(B(x,r))}$$

*where $\delta_0 < 1$ is the number appearing in Theorem 3.24.*

**Proof.** The first inequality follows by Theorem 3.24; the second follows by the first one, since

$$\widetilde{X}_i \widetilde{f} = X_i \widetilde{f} = \widetilde{(X_i f)}.$$

■

### 3.4.3 Vanishing mean oscillation

The definition of $VMO_{loc}(\Omega_k, \Omega_{k+1})$ in an abstract locally homogeneous space has been recalled in § 3.3 (see Definition 3.15); let us endow our domain $\Omega$ with the structure

$$(\Omega, \{\Omega_k\}_k, d_X, dx)$$

of locally homogeneous space described at the end of § 3.3. Then:

**Definition 3.34 (Local $VMO$)** *We say that $a \in VMO_{X,loc}(\Omega)$ if*

$$a \in VMO_{loc}(\Omega_k, \Omega_{k+1}) \text{ for every } k.$$

More explicitly, this means that for any fixed $\Omega' \Subset \Omega$, the function

$$\eta^*_{u,\Omega',\Omega}(r) = \sup_{t \le r} \sup_{x_0 \in \Omega'} \frac{1}{|B_t(x_0)|} \int_{B_t(x_0)} |u(x) - u_{B_t(x_0)}| \, dx,$$

is finite for $r \le r_0$ and vanishes for $r \to 0$, where $r_0$ is the number such that the local doubling condition of Theorem 3.7 holds:

$$|B(x, 2r)| \le c |B(x, r)| \text{ for any } x \in \Omega', r \le r_0.$$

As for Hölder continuous and Sobolev functions, we need a comparison result for $VMO$ functions in the original variables and the lifted ones. By Corollary 3.25 we immediately have the following:

**Proposition 3.35** *Let $a \in VMO_{X,loc}(\Omega)$ then for any $\Omega' \Subset \Omega, x_0 \in \Omega', B(x_0, R)$ and $\widetilde{\Omega}_k = \widetilde{B}\left(\xi_0, \frac{kR}{k+1}\right)$ as before, we have that $\widetilde{a}(x, h) = a(x)$ belongs to the class $VMO_{loc}\left(\widetilde{\Omega}_k, \widetilde{\Omega}_k\right)$ for every $k$, with*

$$\eta^*_{\widetilde{a},\widetilde{\Omega}_k,\widetilde{\Omega}_{k+1}}(r) \le c\eta^*_{a,\Omega',\Omega}(r).$$

In other words, the $VMO_{loc}$ modulus of the original function $a$ controls the $VMO_{loc}$ modulus of its lifted version.

# 4 Operators of type $\lambda$ and representation formulas

## 4.1 Differential operators and fundamental solutions

We now define various differential operators that we will handle in the following. Our main interest is to study the operator

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j + X_0,$$

under the Assumption (H) in § 2. Recall that in view of Remark 2.3 we have set $a_0(x) \equiv 1$.

For any $\overline{x} \in \Omega$ we can apply the "lifting theorem" to the vector fields $X_i$ (see § 3.1 for the statement and notation), obtaining new vector fields $\widetilde{X}_i$ which are free up to weight $s$ and satisfy Hörmander's condition of weight $s$ in a neighborhood of $\overline{\xi} = (\overline{x}, 0) \in \mathbb{R}^N$. For $\xi = (x, t) \in \widetilde{B}\left(\overline{\xi}, R\right)$, with $\widetilde{B}\left(\overline{\xi}, R\right)$ as in Lemma 3.9, set

$$\widetilde{a}_{ij}(x, t) = a_{ij}(x),$$

and let

$$\widetilde{\mathcal{L}} = \sum_{i,j=1}^{q} \widetilde{a}_{ij}(\xi)\widetilde{X}_i\widetilde{X}_j + \widetilde{X}_0 \tag{4.1}$$

be the lifted operator, defined in $\widetilde{B}\left(\overline{\xi}, R\right)$. Next, we freeze $\widetilde{\mathcal{L}}$ at some point $\xi_0 \in \widetilde{B}\left(\overline{\xi}, R\right)$, and consider the frozen lifted operator:

$$\widetilde{\mathcal{L}}_0 = \sum_{i,j=1}^{q} \widetilde{a}_{ij}(\xi_0)\widetilde{X}_i\widetilde{X}_j + \widetilde{X}_0. \tag{4.2}$$

To study $\widetilde{\mathcal{L}}_0$, in view of the "approximation theorem" (Thm. 3.2), we will consider the approximating operator, defined on the homogeneous group $\mathbb{G}$:

$$\mathcal{L}_0^* = \sum_{i,j=1}^{q} \widetilde{a}_{ij}(\xi_0)Y_iY_j + Y_0$$

and its transpose:

$$\mathcal{L}_0^{*T} = \sum_{i,j=1}^{q} \widetilde{a}_{ij}(\xi_0)Y_iY_j - Y_0$$

where $\{Y_i\}$ are the left invariant vector fields on the group $\mathbb{G}$ defined in § 3.1.

We will apply to $\mathcal{L}_0^*$ and $\mathcal{L}_0^{*T}$ several results proved in [2], which in turn are based on results due to Folland [16, Thm. 2.1 and Corollary 2.8] and Folland-Stein [17, Proposition 8.5]. They are collected in the following theorem:

**Theorem 4.1** *Assume that the homogeneous dimension of $\mathbb{G}$ is $Q \geq 3$. For every $\xi_0 \in \widetilde{B}\left(\overline{\xi}, R\right)$ the operator $\mathcal{L}_0^*$ has a unique fundamental solution $\Gamma(\xi_0; \cdot)$ such that: :*

(a) $\Gamma(\xi_0; \cdot) \in C^\infty\left(\mathbb{R}^N \setminus \{0\}\right)$;

(b) $\Gamma(\xi_0; \cdot)$ *is homogeneous of degree* $(2 - Q)$;

(c) *for every test function $f$ and every $v \in \mathbb{R}^N$,*

$$f(v) = \left(\mathcal{L}_0^* f * \Gamma(\xi_0; \cdot)\right)(v) = \int_{\mathbb{R}^N} \Gamma\left(\xi_0; u^{-1} \circ v\right) \mathcal{L}_0^* f(u)\, du;$$

*moreover, for every $i, j = 1, \ldots, q$, there exist constants $\alpha_{ij}(\xi_0)$ such that*

$$Y_iY_j f(v) = P.V. \int_{\mathbb{R}^N} Y_iY_j\Gamma\left(\xi_0; u^{-1} \circ v\right) \mathcal{L}_0^* f(u) du + \alpha_{ij}(\xi_0) \cdot \mathcal{L}_0^* f(v); \tag{4.3}$$

(d) $Y_iY_j\Gamma(\xi_0; \cdot) \in C^\infty\left(\mathbb{R}^N \setminus \{0\}\right)$;

(e) $Y_iY_j\Gamma(\xi_0; \cdot)$ *is homogeneous of degree* $-Q$;

(f)

$$\int_{r<\|u\|<R} Y_iY_j\Gamma(\xi_0; u)\, du = \int_{\|u\|=1} Y_iY_j\Gamma(\xi_0; u)\, d\sigma(u) = 0 \quad \textit{for every } R > r > 0.$$

In (4.3) the notation $P.V. \int_{\mathbb{R}^N} (...) \, du$ stands for $\lim_{\varepsilon \to 0} \int_{\|u^{-1} \circ v\| > \varepsilon} (...) \, du$.

**Remark 4.2** *By [16, Remark on p.174], we know that the fundamental solution of the transposed operator* $\mathcal{L}_0^{*T}$ *is*

$$\Gamma^T(\xi_0; u) = \Gamma\left(\xi_0; u^{-1}\right) = \Gamma(\xi_0; -u).$$

*(However, beware that* $Y_i \Gamma^T(\xi_0; u) \neq Y_i \Gamma(\xi_0; -u)$*).*

Throughout the following, we will set, for $i, j = 1, \ldots, q$,

$$\begin{aligned} \Gamma_{ij}(\xi_0; u) &= Y_i Y_j \, [\Gamma(\xi_0; \cdot)](u); \\ \Gamma_{ij}^T(\xi_0; u) &= Y_i Y_j \, \left[\Gamma^T(\xi_0; \cdot)\right](u). \end{aligned}$$

A second fundamental result we need contains a bound on the derivatives of $\Gamma$, uniform with respect to $\xi_0$, and is proved in [2, Thm. 12]:

**Theorem 4.3** *For every multi-index* $\beta$*, there exists a constant* $c = c(\beta, \mathbb{G}, \mu)$ *such that*

$$\sup_{\substack{\|u\|=1 \\ \xi \in \widetilde{B}(\overline{\xi}, R)}} \left| \left(\frac{\partial}{\partial u}\right)^{\beta} \Gamma_{ij}(\xi\,;\,u) \right| \leq c,$$

*for any* $i, j = 1, \ldots, q$*; moreover, for the* $\alpha_{ij}$*'s appearing in (4.3), the uniform bound*

$$\sup_{\xi \in \widetilde{B}(\overline{\xi}, R)} |\alpha_{ij}(\xi)| \leq c_2$$

*holds for some constant* $c_2 = c_2(\mathbb{G}, \mu)$*.*

**Remark 4.4** *Theorems 4.1 and 4.3 still hold when we replace* $\Gamma$ *by* $\Gamma^T$ *and* $\Gamma_{ij}$ *by* $\Gamma_{ij}^T$*.*

## 4.2 Operators of type $\lambda$

As in [25] and [3], we are going to build a parametrix for $\widetilde{\mathcal{L}}$ shaped on the homogeneous fundamental solution of $\mathcal{L}_0^*$. More generally, we need to define a class of integral operators with different degrees of singularity. The next definition is adapted from [3], the difference being the necessity, in the present case, to consider integral kernels shaped on the fundamental solutions of both $\mathcal{L}_0^*$ and $\mathcal{L}_0^{*T}$.

**Definition 4.5** *For any* $\xi_0 \in \widetilde{B}\left(\overline{\xi}, R\right)$*, we say that* $k(\xi_0; \xi, \eta)$ *is a* frozen kernel of type $\lambda$ *(over the ball* $\widetilde{B}\left(\overline{\xi}, R\right)$*), for some nonnegative integer* $\lambda$*, if for every* *positive integer m we can write, for* $\xi,\eta\in\widetilde{B}\left(\overline{\xi},R\right)$,

$$\begin{aligned}k(\xi_0;\xi,\eta) &= k'(\xi_0;\xi,\eta)+k''(\xi_0;\xi,\eta)\\ &= \left\{\sum_{i=1}^{H_m}a_i(\xi)b_i(\eta)D_i\Gamma(\xi_0;\cdot)+a_0(\xi)b_0(\eta)D_0\Gamma(\xi_0;\cdot)\right\}(\Theta(\eta,\xi))\\ &+\left\{\sum_{i=1}^{H_m}a_i'(\xi)b_i'(\eta)D_i'\Gamma^T(\xi_0;\cdot)+a_0'(\xi)b_0'(\eta)D_0'\Gamma^T(\xi_0;\cdot)\right\}(\Theta(\eta,\xi))\end{aligned}$$

*where* $a_i,b_i,a_i',b_i'\in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$ $(i=0,1,\dots H_m)$, $D_i$ *and* $D_i'$ *are differential operators such that: for* $i=1,\dots,H_m$, $D_i$ *and* $D_i'$ *are homogeneous of degree* $\leq 2-\lambda$ *(so that* $D_i\Gamma(\xi_0;\cdot)$ *and* $D_i'\Gamma^T(\xi_0;\cdot)$ *are homogeneous functions of degree* $\geq \lambda-Q$*);* $D_0$ *and* $D_0'$ *are differential operators such that* $D_0\Gamma(\xi_0;\cdot)$ *and* $D_0'\Gamma^T(\xi_0;\cdot)$ *have m (weighted) derivatives with respect to the vector fields* $Y_i$ $(i=0,1,\dots,q)$. *Moreover, the coeffi cients of the differential operators* $D_i,D_i'$ *for* $i=0,1,...,H_m$ *possibly depend also on the variables* $\xi,\eta$, *in such a way that the joint dependence on* $(\xi,\eta,u)$ *is smooth.*

In order to simplify notation, we will not always express explicitly this dependence of the coeffi cients of $D_i$ on $\xi,\eta$. Only when it is necessary we will write, for instance, $a_i(\xi)b_i(\eta)D_i^{\xi,\eta}\Gamma(\xi_0;\Theta(\eta,\xi))$ to recall this dependence.

**Remark 4.6** *Note that if a smooth function* $c(\xi,\eta,u)$ *is* $D(\lambda)$*-homogeneous of some degree* $\beta$ *with respect to* $u$, *then any* $\xi$ *or* $\eta$ *derivative of* $c$ *has the same homogeneity with respect to* $u$, *since*

$$c(\xi,\eta,D(\lambda)u)=\lambda^\beta c(\xi,\eta,u)\ \text{ implies }\ \frac{\partial c}{\partial\xi_i}(\xi,\eta,D(\lambda)u)=\lambda^\beta\frac{\partial c}{\partial\xi_i}(\xi,\eta,u).$$

*Hence any derivative*

$$\left(\frac{\partial}{\partial\xi_i}D_i^{\xi,\eta}\right)\Gamma(\xi_0;\cdot),\left(\frac{\partial}{\partial\eta_i}D_i^{\xi,\eta}\right)\Gamma(\xi_0;\cdot)$$

*has the same homogeneity as*

$$D_i^{\xi,\eta}\Gamma(\xi_0;\cdot).$$

**Definition 4.7** *For any* $\xi_0\in\widetilde{B}\left(\overline{\xi},R\right)$, *we say that* $T(\xi_0)$ *is a* frozen operator of type $\lambda\geq 1$ *(over the ball* $\widetilde{B}\left(\overline{\xi},R\right)$*) if* $k(\xi_0;\xi,\eta)$ *is a frozen kernel of type* $\lambda$ *and*

$$T(\xi_0)f(\xi)=\int_{\widetilde{B}}k(\xi_0;\xi,\eta)\,f(\eta)\,d\eta$$

*for* $f\in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$. *We say that* $T(\xi_0)$ *is a* frozen operator of type 0 *if* $k(\xi_0;\xi,\eta)$ *is a frozen kernel of type* 0 *and*

$$T(\xi_0)f(\xi)=P.V.\int_{\widetilde{B}}k(\xi_0;\xi,\eta)\,f(\eta)\,d\eta+\alpha(\xi_0,\xi)\,f(\xi),$$

*where $\alpha$ is a bounded measurable function, smooth in $\xi$, and the principal value integral exists. Explicitly, this principal value is defined by:*

$$P.V. \int_{\widetilde{B}} k(\xi_0;\xi,\eta)\, f(\eta)\, d\eta = \lim_{\varepsilon\to 0} \int_{\|\Theta(\eta,\xi)\|>\varepsilon} k(\xi_0;\xi,\eta)\, f(\eta)\, d\eta.$$

**Definition 4.8** *If $k(\xi_0;\xi,\eta)$ is a frozen kernel of type $\lambda \geq 0$, we say that $k(\xi;\xi,\eta)$ is a* variable kernel of type $\lambda$ *(over the ball $\widetilde{B}\left(\overline{\xi},R\right)$), and*

$$Tf(\xi) = \int_{\widetilde{B}} k(\xi;\xi,\eta) f(\eta)\, d\eta$$

*is a* variable operator of type $\lambda$. *If $\lambda = 0$, the integral must be taken in principal value sense and a term $\alpha\left(\xi,\xi\right) f\left(\xi\right)$ must be added.*

With reference to Definition 4.5, we will call the $k', k''$ parts of $k$ "frozen kernel of type $\lambda$ modeled on $\Gamma, \Gamma^T$", respectively. Analogously we will sometimes speak of frozen operators of type $\lambda$ modeled on $\Gamma$ or $\Gamma^T$, to denote that the kernel has this special form.

A common operation on frozen operators is *transposition*:

**Definition 4.9** *If $T\left(\xi_0\right)$ is a frozen operator of type $\lambda \geq 0$ over $\widetilde{B}\left(\overline{\xi},R\right)$,we will denote by $T\left(\xi_0\right)^T$ the transposed operator, formally defined by*

$$\int_{\widetilde{B}} f\left(\xi\right) T\left(\xi_0\right)^T g\left(\xi\right) d\xi = \int_{\widetilde{B}} g\left(\xi\right) T\left(\xi_0\right) f\left(\xi\right) d\xi$$

*for any $f,g \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$.*

Clearly, if $k\left(\xi_0,\xi,\eta\right)$ is the kernel of $T\left(\xi_0\right)$, then $k\left(\xi_0,\eta,\xi\right)$ is the kernel of $T\left(\xi_0\right)^T$. It is useful to note that:

**Proposition 4.10** *If $T\left(\xi_0\right)$ is a frozen operator of type $\lambda \geq 0$ over $\widetilde{B}\left(\overline{\xi},R\right)$, modeled on $\Gamma$ or $\Gamma^T$, then $T\left(\xi_0\right)^T$ is a frozen operator of type $\lambda$, modeled on $\Gamma^T, \Gamma$, respectively. In particular, the transposed of a frozen operator of type $\lambda$ is still a frozen operator of type $\lambda$.*

**Proof.** Let $D$ be any differential operator on the group $\mathbb{G}$. For any $f \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$, let $f'\left(u\right) = f\left(-u\right)$. Let $D'$ be the differential operator defined by the identity

$$D'f = \left(D\left(f'\right)\right)'.$$

Clearly, if $D$ is homogeneous of some degree $\beta$, the same is true for $D'$; if $D\Gamma(\xi_0;\cdot)$ or $D\Gamma^T(\xi_0;\cdot)$ have $m$ (weighted) derivatives with respect to the vector fields $Y_i$ $(i = 0, 1, \dots, q)$, the same is true for $D'\Gamma(\xi_0;\cdot)$ or $D'\Gamma^T(\xi_0;\cdot)$. Also, recalling that $\Gamma^T(\xi_0;u) = \Gamma(\xi_0;-u)$, we have

$$\begin{aligned}(D'\Gamma)(u) &= \left(D\Gamma^T\right)(-u)\\ \left(D'\Gamma^T\right)(u) &= (D\Gamma)(-u).\end{aligned}$$

Moreover, these identities can be iterated, for instance:

$$(D_1D_2\Gamma)(-u) = (D_1(D_2\Gamma))(-u) = \left(D_1'(D_2\Gamma)'\right)(u) = \left(D_1'D_2'\Gamma^T\right)(u).$$

Then, if

$$k'(\xi_0,\xi,\eta) = \left\{\sum_{i=1}^{H_m} a_i(\xi)b_i(\eta)D_i\Gamma(\xi_0;\cdot) + a_0(\xi)b_0(\eta)D_0\Gamma(\xi_0;\cdot)\right\}(\Theta(\eta,\xi))$$

is a frozen kernel of type $\lambda$ modeled on $\Gamma$,

$$\begin{aligned}k'(\xi_0,\eta,\xi) &= \left\{\sum_{i=1}^{H_m} a_i(\eta)b_i(\xi)D_i\Gamma(\xi_0;\cdot) + a_0(\xi)b_0(\eta)D_0\Gamma(\xi_0;\cdot)\right\}(-\Theta(\eta,\xi))\\ &= \left\{\sum_{i=1}^{H_m} a_i(\eta)b_i(\xi)D_i'\Gamma^T(\xi_0;\cdot) + a_0(\xi)b_0(\eta)D_0'\Gamma^T(\xi_0;\cdot)\right\}(\Theta(\eta,\xi))\end{aligned}$$

is a frozen kernel of type $\lambda$ modeled on $\Gamma^T$. Analogously one can prove the converse. ■

We have now to deal with the relations between operators of type $\lambda$ and the differential operators represented by the vector fields $\widetilde{X}_i$. This is a study which has been carried out in [25, § 14], and adapted to nonvariational operators in [3]. We are interested in two main results. Roughly speaking, the first says that the composition, in any order, of an operator of type $\lambda$ with the $\widetilde{X}_i$ or $\widetilde{X}_0$ derivative is an operator of type $\lambda-1$ or $\lambda-2$, respectively. The second says that the $\widetilde{X}_i$ derivative of an operator of type $\lambda$ can be rewritten as the sum of other operators of type $\lambda$, each acting on a different $\widetilde{X}_j$ derivative, plus a suitable remainder. In [25] these results are proved only for a system of Hörmander's vector fields of weight one (that is, without the drift), and some proofs are quite condensed. Hence we need to extend and modify some arguments in [25, § 14] to cover the present situation. Moreover, as in [3], we need to keep under careful control the dependence of any quantity on the frozen point $\xi_0$ appearing in $\Gamma(\xi_0,\cdot)$. For these and other technical reasons, we prefer to write complete proofs of these properties, even though they are not so different from known results. The first result is the following:

**Theorem 4.11** *(See [25, Thm. 8]). Suppose $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$. Then $\widetilde{X}_kT(\xi_0)$ and $T(\xi_0)\widetilde{X}_k$ $(k = 1, 2, ..., q)$ are operators of type $\lambda - 1$. If $\lambda \geq 2$, then $\widetilde{X}_0T(\xi_0)$ and $T(\xi_0)\widetilde{X}_0$ are operators of type $\lambda - 2$.*

To prove this, we begin by stating the following two lemmas:

**Lemma 4.12** *If $k(\xi_0;\xi,\eta)$ is a frozen kernel of type $\lambda \geq 1$ over $\widetilde{B}\left(\overline{\xi},R\right)$, then $\left(\widetilde{X}_j k\right)(\xi_0;\cdot,\eta)(\xi)$ $(j=1,2,...,q)$ is a frozen kernel of type $\lambda-1$. If $\lambda \geq 2$, then $\left(\widetilde{X}_0 k\right)(\xi_0;\cdot,\eta)(\xi)$ is a frozen kernel of type $\lambda-2$.*

**Proof.** This basically follows by the definition of kernel of type $\lambda$ and Theorem 3.2 in § 3.1. When the $\widetilde{X}_j$ derivative acts on the $\xi$ variable of a kernel $D_i^{\xi}\Gamma(\xi_0,\cdot)$, one also has to take into account Remark 4.6.

Here we just want to point out the following fact. The prototype of frozen kernel of type 2 is the function

$$a(\xi)\,\Gamma(\xi_0;\Theta(\eta,\xi))\,b(\eta).$$

Note that the computation

$$\begin{aligned}
&\widetilde{X}_i\left[a(\cdot)\,\Gamma(\xi_0;\Theta(\eta,\cdot))\,b(\eta)\right](\xi)\\
&=a(\xi)\left[(Y_i+R_i^{\eta})\,\Gamma(\xi_0;\cdot)\right](\Theta(\eta,\xi))\,b(\eta)+\left(\widetilde{X}_i a\right)(\xi)\,\Gamma(\xi_0;\Theta(\eta,\xi))\,b(\eta)
\end{aligned}$$

in particular generates the term

$$a(\xi)\,(R_i^{\eta}\Gamma)(\xi_0;\cdot)\,(\Theta(\eta,\xi))\,b(\eta)$$

where the differential operator $R_i^{\eta}$ has coeffi cients depending on$\eta$. In the proof of Theorem 4.11 we will see another basic computation on frozen kernels which generates differential operators with coeffi cients also depending on$\xi$. This is the reason why Definition 4.5 allows for this kind of dependence. ∎

**Lemma 4.13** *If $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$ over $\widetilde{B}\left(\overline{\xi},R\right)$, then $\widetilde{X}_i T(\xi_0)$ $(i=1,2,...,q)$ is a frozen operator of type $\lambda-1$. If $\lambda \geq 2$, then $\widetilde{X}_0 T(\xi_0)$ is a frozen operator of type $\lambda-2$.*

**Proof.** With reference to Definition 4.5, it is enough to consider the part $k'$ of the kernel of $T$, the proof for $k''$ being completely analogous. So, let us consider the operator $\widetilde{X}_i T(\xi_0)$ $(i=1,2,...,q)$, where $T(\xi_0)$ has kernel $k'$.

If $\lambda > 1$, the result immediately follows by the previous lemma. If $\lambda = 1$, then

$$T(\xi_0)f(\xi)=\int_{\widetilde{B}\left(\overline{\xi},R\right)} a(\xi)b(\eta)D_1\Gamma(\xi_0;\Theta(\eta,\xi))\,f(\eta)\,d\eta+T'(\xi_0)f(\xi)$$

where $T'(\xi_0)$ is a frozen operator of type 2, and $D_1$ is a 1-homogeneous differential operator. We already know that $\widetilde{X}_i T'(\xi_0)$ is a frozen operator of type 1, so we are left to show that

$$\widetilde{X}_i\int_{\widetilde{B}\left(\overline{\xi},R\right)} a(\xi)b(\eta)D_1\Gamma(\xi_0;(\Theta(\eta,\xi)))f(\eta)\,d\eta$$

is a frozen operator of type 0. To do this, we have to apply a distributional argument, which will be used several times in the following: let us compute, for any $\omega \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$,

$$\int_{\widetilde{B}\left(\overline{\xi},R\right)} \widetilde{X}_i^T \omega\left(\xi\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} a(\xi)b(\eta) D_1^\xi \Gamma(\xi_0; (\Theta(\eta,\xi))) f\left(\eta\right) d\eta d\xi$$
$$= \lim_{\varepsilon\to 0} \int_{\widetilde{B}\left(\overline{\xi},R\right)} \widetilde{X}_i^T \omega\left(\xi\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} a(\xi)b(\eta)\varphi_\varepsilon\left(\Theta(\eta,\xi)\right) D_1^\xi \Gamma(\xi_0; (\Theta(\eta,\xi))) f\left(\eta\right) d\eta d\xi$$

where $\varphi_\varepsilon\left(u\right) = \varphi\left(D\left(\varepsilon^{-1}\right) u\right)$ and $\varphi \in C_0^\infty\left(\mathbb{R}^N\right), \varphi\left(u\right) = 0$ for $\|u\| < 1, \varphi\left(u\right) = 1$ for $\|u\| > 2$. Here we have written $D_1^\xi$ to recall that the coeffi cients of the differential operator $D_1$ also depend (smoothly) on $\xi$ as a parameter. By Theorem 3.2,

$$\int_{\widetilde{B}\left(\overline{\xi},R\right)} \widetilde{X}_i^T \omega\left(\xi\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} a(\xi)b(\eta)\varphi_\varepsilon\left(\Theta(\eta,\xi)\right) D_1^\xi \Gamma(\xi_0; (\Theta(\eta,\xi))) f\left(\eta\right) d\eta d\xi$$
$$= \int_{\widetilde{B}\left(\overline{\xi},R\right)} b(\eta) f\left(\eta\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} \left(\widetilde{X}_i^T \omega\right)\left(\xi\right) a(\xi)\varphi_\varepsilon\left(\Theta(\eta,\xi)\right) D_1^\xi \Gamma(\xi_0; (\Theta(\eta,\xi))) d\xi d\eta$$
$$= \int_{\widetilde{B}\left(\overline{\xi},R\right)} b(\eta) f\left(\eta\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} \omega\left(\xi\right) \left(\widetilde{X}_i a\right)\left(\xi\right)\varphi_\varepsilon\left(\Theta(\eta,\xi)\right) D_1^\xi \Gamma(\xi_0; (\Theta(\eta,\xi))) d\xi d\eta$$
$$+ \int_{\widetilde{B}\left(\overline{\xi},R\right)} b(\eta) f\left(\eta\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} \omega\left(\xi\right) a(\xi)\varphi_\varepsilon\left(\Theta(\eta,\xi)\right) \left(\widetilde{X}_i D_1^\xi\right) \Gamma(\xi_0; (\Theta(\eta,\xi))) d\xi d\eta$$
$$+ \int_{\widetilde{B}\left(\overline{\xi},R\right)} b(\eta) f\left(\eta\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} \omega\left(\xi\right) a(\xi) \left[\left(Y_i + R_i^\eta\right)\left(\varphi_\varepsilon D_1^\xi \Gamma(\xi_0; \cdot)\right)\right]\left(\Theta(\eta,\xi)\right) d\xi d\eta$$
$$\equiv A_\varepsilon + B_\varepsilon + C_\varepsilon. \tag{4.4}$$

Now,

$$A_\varepsilon \to \int_{\widetilde{B}\left(\overline{\xi},R\right)} b(\eta) f\left(\eta\right) \int_{\widetilde{B}\left(\overline{\xi},R\right)} \omega\left(\xi\right) \left(\widetilde{X}_i a\right)\left(\xi\right) D_1 \Gamma(\xi_0; (\Theta(\eta,\xi))) d\xi d\eta$$
$$= \int_{\widetilde{B}\left(\overline{\xi},R\right)} f\left(\eta\right) S_1\left(\xi_0\right) \omega\left(\eta\right) d\eta$$
$$= \int_{\widetilde{B}\left(\overline{\xi},R\right)} \omega\left(\eta\right) S_1\left(\xi_0\right)^T f\left(\eta\right) d\eta \tag{4.5}$$

where $S_1(\xi_0)$ is a frozen operator of type 1, and $S_1(\xi_0)^T$, its transpose, is still a frozen operator of type 1 (see Proposition 4.10).

$$\begin{aligned}
B_\varepsilon &\to \int_{\widetilde{B}(\overline{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi)\, a(\xi) \left(\widetilde{X}_i D_1^{\xi}\right) \Gamma(\xi_0; (\Theta(\eta,\xi))) d\xi d\eta \\
&= \int_{\widetilde{B}(\overline{\xi},R)} f(\eta)\, S_1'(\xi_0)\, \omega(\eta)\, d\eta \\
&= \int_{\widetilde{B}(\overline{\xi},R)} \omega(\eta)\, S_1'(\xi_0)^T f(\eta)\, d\eta \qquad (4.6)
\end{aligned}$$

where, by Remark 4.6, $S_1'(\xi_0)$ is a frozen operator of type 1, and the same is true for $S_1'(\xi_0)^T$ by Proposition 4.10.

$$\begin{aligned}
C_\varepsilon &= \int_{\widetilde{B}(\overline{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi)\, a(\xi) \left[\varphi_\varepsilon Y_i D_1 \Gamma(\xi_0; \cdot)\right] (\Theta(\eta,\xi))\, d\xi d\eta \\
&+ \int_{\widetilde{B}(\overline{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi)\, a(\xi) \left[\varphi_\varepsilon R_i^\eta D_1 \Gamma(\xi_0; \cdot)\right] (\Theta(\eta,\xi))\, d\xi d\eta \\
&+ \int_{\widetilde{B}(\overline{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi)\, a(\xi) \left[(Y_i + R_i^\eta)\, \varphi_\varepsilon D_1 \Gamma(\xi_0; \cdot)\right] (\Theta(\eta,\xi))\, d\xi d\eta \\
&\equiv C_\varepsilon^1 + C_\varepsilon^2 + C_\varepsilon^3. \qquad (4.7)
\end{aligned}$$

Now:

$$\begin{aligned}
C_\varepsilon^1 &\to \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi) \left\{ P.V. \int_{\widetilde{B}(\overline{\xi},R)} a(\xi) Y_i D_1 \Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta)\, d\eta \right\} d\xi \\
&= \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi)\, T(\xi_0)\, f(\xi)\, d\xi \qquad (4.8)
\end{aligned}$$

with $T(\xi_0)$ frozen operator of type 0. Note that the principal value exists because the kernel $Y_i D_1 \Gamma(\xi_0; u)$ has vanishing integral over spherical shells $\{u \in \mathbb{G} : r_1 < \|u\| < r_2\}$ (see Theorem 4.1).

$$\begin{aligned}
C_\varepsilon^2 &\to \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi) \left\{ \int_{\|u\|<R} a(\xi) R_i^\eta D_1 \Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta)\, d\eta \right\} d\xi \\
&= \int_{\widetilde{B}(\overline{\xi},R)} \omega(\xi)\, S(\xi_0)\, f(\xi)\, d\xi \qquad (4.9)
\end{aligned}$$

with $S(\xi_0)$ frozen operator of type 1.

To handle $C_\varepsilon^3$, let us perform the change of variables $u = \Theta(\eta,\xi)$ which, by Theorem 3.3 gives

$$\begin{aligned}
C_\varepsilon^3 = \int_{\widetilde{B}(\overline{\xi},R)} (bf)(\eta) \int_{\|u\|<R} (\omega a) \left(\Theta(\eta,\cdot)^{-1}(u)\right) \left[(Y_i + R_i^\eta)\, \varphi_\varepsilon D_1 \Gamma(\xi_0; \cdot)\right](u) \cdot \\
\cdot\, c(\eta) \left(1 + O(\|u\|)\right) du d\eta
\end{aligned}$$

On the other hand, $Y_i\varphi_\varepsilon(u) = \frac{1}{\varepsilon}Y_i\varphi\left(D\left(\frac{1}{\varepsilon}\right)u\right)$, while $R_i^\eta\varphi_\varepsilon(u)$ is uniformly bounded in $\varepsilon$. Hence the change of variables $D\left(\frac{1}{\varepsilon}\right)u = v$ gives

$$\begin{aligned}
C_\varepsilon^3 &= \int_{\widetilde{B}(\overline{\xi},R)} (bf)(\eta) \int_{\|v\|<\frac{R}{\varepsilon}} (\omega a)\left(\Theta(\eta,\cdot)^{-1}(D(\varepsilon)v)\right)\left[\frac{1}{\varepsilon}Y_i\varphi(v) + O(1)\right]\cdot \\
&\quad \cdot c(\eta)\,\varepsilon^{1-Q} D_1^\eta\Gamma(\xi_0;v)\,(1+O(\varepsilon\|v\|))\,\varepsilon^Q dvd\eta \\
&\to \int_{\widetilde{B}(\overline{\xi},R)} (bcf)(\eta) \int_{\|v\|<\frac{R}{\varepsilon}} (\omega a)\left(\Theta(\eta,\cdot)^{-1}(0)\right) Y_i\varphi(v)\, D_1^\eta\Gamma(\xi_0;v)\,dvd\eta \\
&= \int_{\widetilde{B}(\overline{\xi},R)} (\omega abcf)(\eta) \int_{\|v\|<\frac{R}{\varepsilon}} Y_i\varphi(v)\, D_1^\eta\Gamma(\xi_0;v)\,dvd\eta \\
&= \int_{\widetilde{B}(\overline{\xi},R)} (\omega abcf)(\eta)\,\alpha(\xi_0,\eta)\,d\eta, \qquad (4.10)
\end{aligned}$$

which is the integral of $\omega$ times the multiplicative part of a frozen operator of type 0. It is worthwhile (although not logically necessary to prove the theorem) to realize that the quantity $\alpha(\xi_0,\eta)$ appearing in (4.10) actually does not depend on the function $\varphi$. Namely, recalling that $Y_i\varphi(v)$ is supported in the spherical shell $1 \leq \|v\| \leq 2$, with $\varphi(u) = 1$ for $\|u\| = 2$ and $\varphi(u) = 0$ for $\|u\| = 1$, an integration by parts gives

$$\begin{aligned}
&\int_{1\leq\|v\|\leq 2} Y_i\varphi(v)\, D_1^\eta\Gamma(\xi_0;v)\,dv \\
&= -\int_{1\leq\|v\|\leq 2} \varphi(v)\, Y_i D_1^\eta\Gamma(\xi_0;v)\,dv + \int_{\|v\|=2} D_1^\eta\Gamma(\xi_0;v)\,n_i d\sigma(v)
\end{aligned}$$

with $n_i = \sum_{j=1}^N b_{ij}(u)\,\nu_j$, where $Y_i = \sum_{j=1}^N b_{ij}(u)\,\partial_{u_j}$ and $\nu$ is the outer normal on $\|v\| = 2$. The vanishing property of the kernel $Y_i D_1^\xi\Gamma(\xi_0;\cdot)$ implies that if $\varphi$ is a radial function the first integral vanishes. Therefore

$$\alpha(\xi_0,\eta) = \int_{\|v\|=2} D_1^\eta\Gamma(\xi_0;v)\,n_i d\sigma(v)$$

which also shows that $\alpha(\xi_0,\eta)$ smoothly depends on $\eta$ and is bounded in $\xi_0$ (by Theorem 4.3). By (4.4), (4.5), (4.6), (4.8), (4.9), (4.10) we have therefore proved that

$$\widetilde{X}_i T(\xi_0) f(\xi) = S_1(\xi_0)^T f(\xi) + S_1'(\xi_0)^T f(\xi) + T(\xi_0) f(\xi) + \alpha(\xi_0,\xi)(abcf)(\xi)$$

which is a frozen operator of type 0.

This completes the proof of the first statement of the Lemma. The proof of the fact that if $\lambda \geq 2$ then $\widetilde{X}_0 T(\xi_0)$ is a frozen operator of type $\lambda - 2$ is completely analogous. ■

The above two lemmas imply the assertion on $\widetilde{X}_k T(\xi_0)$ and $\widetilde{X}_0 T(\xi_0)$ in Theorem 4.11. To prove the assertions about $T(\xi_0)\widetilde{X}_k, T(\xi_0)\widetilde{X}_0$ we need a way to express $\xi$-derivatives of the integral kernel in terms of $\eta$-derivatives of the kernel, in order to integrate by parts. This will involve the use of *right invariant vector fields* on the group $\mathbb{G}$: throughout the following, we will denote by

$$Y_{i,k}^{R}$$

the right invariant vector field on $\mathbb{G}$ satisfying $Y_{i,k}^{R}f(0) = Y_{i,k}f(0)$. We have the following:

**Lemma 4.14** *For any $f \in C_0^{\infty}(\mathbb{G})$ and $\eta, \xi$ in a neighborhood of $\xi_0$, we can write, for any $i = 1, 2, ..., s$, $k = 1, 2, ..., k_i$ (recall $s$ is the step of the Lie algebra)*

$$\widetilde{X}_{i,k}\left[f\left(\Theta\left(\cdot,\xi\right)\right)\right](\eta) = -\left(Y_{i,k}^{R}f\right)\left(\Theta\left(\eta,\xi\right)\right) + \left(\left(R_{i,k}^{\xi}\right)' f\right)\left(\Theta\left(\eta,\xi\right)\right), \quad (4.11)$$

*where $\left(R_{i,k}^{\xi}\right)'$ is a vector field of local degree $\leq i-1$ smoothly depending on $\xi$.*

**Proof.** We start with the following

**Claim.** For any function $f$ defined on $\mathbb{G}$, let

$$f'\left(u\right) = f\left(-u\right)$$

(recall that $-u = u^{-1}$); then the following identities hold:

$$Y_{i,k}\left(f'\right) = -\left(Y_{i,k}^{R}f\right)'. \quad (4.12)$$

To prove this, let us define the vector fields $\widehat{Y}_{i,k}$ by

$$Y_{i,k}\left(f'\right) = -\left(\widehat{Y}_{i,k}f\right)', \quad (4.13)$$

then for any $a \in \mathbb{G}$, denoting by $L_a, R_a$ the corresponding operators of left and right translation, respectively (acting on functions), we have

$$\begin{aligned}(\widehat{Y}_{i,k}R_a f)' &= -Y_{i,k}((R_a f)') = -Y_{i,k}(L_{-a}f') = \\ &= -L_{-a}Y_{i,k}f' = L_{-a}(-Y_{i,k}f') = \\ &= L_{-a}(\widehat{Y}_{i,k}f)' = (R_a\widehat{Y}_{i,k}f)',\end{aligned}$$

hence $\widehat{Y}_{i,k}$ are right invariant vector fields. Also, note that for any vector field $Y = \sum a_j\left(u\right)\partial_{u_j}$ we have

$$Y\left(f'\right)(0) = -\left(Yf\right)(0)$$

because

$$\begin{aligned}&Y\left(f'\right)(u) = \sum a_j\left(u\right)\partial_{u_j}\left[f\left(-u\right)\right] = -\sum a_j\left(u\right)\left(\partial_{u_j}f\right)\left(-u\right) \text{ implies} \\ &Y\left(f'\right)(0) = -\sum a_j\left(0\right)\left(\partial_{u_j}f\right)(0) = -\left(Yf\right)(0)\end{aligned}$$

hence by (4.13) we know that $\widehat{Y}_k f(0) = Y_k f(0)$. Therefore $\widehat{Y}_k$ is the right invariant vector field which coincides with $Y_k$ at the origin, that is $\widehat{Y}_k = Y_k^R$, and the Claim is proved.

By (3.5) and (4.12),

$$\begin{aligned}\widetilde{X}_{i,k}\left[f\left(\Theta\left(\cdot,\xi\right)\right)\right](\eta) &= \widetilde{X}_{i,k}\left[f'\left(\Theta\left(\xi\,,\cdot\right)\right)\right](\eta) = \\ &= \left(Y_{i,k}f' + R^{\xi}_{i,k}f'\right)\left(\Theta\left(\xi\,,\eta\right)\right) = \\ &= -\left(Y^R_{i,k}f\right)'\left(\Theta\left(\xi\,,\eta\right)\right) + R^{\xi}_{i,k}f'\left(\Theta\left(\xi\,,\eta\right)\right) = \\ &= -\left(Y^R_{i,k}f\right)\left(\Theta\left(\eta,\xi\right)\right) + \left(\left(R^{\xi}_{i,k}\right)' f\right)\left(\Theta\left(\eta,\xi\right)\right),\end{aligned} \tag{4.14}$$

where

$$\left(\left(R^{\xi}_{i,k}\right)' f\right)(u) = \left(R^{\xi}_{i,k}f'\right)(-u)$$

is a differential operator of degree $\leq i-1$. This proves (4.11). ■

**Proof of Theorem 4.11.** As we noted after Lemma 4.13, we are left to prove the assertion about $T\left(\xi_0\right)\widetilde{X}_i$ and $T\left(\xi_0\right)\widetilde{X}_0$. We only give the proof for the case $\lambda \geq 1$, $i = 1,\cdots,q$, the proof for $\lambda \geq 2$, $i = 0$ being very similar. Like in the proof of Lemma 4.13, it is enough to consider the part $k'$ of the kernel of $T$, the proof for $k''$ being completely analogous (see Definition 4.5). Let us expand

$$k'(\xi_0;\xi,\eta) = \left\{\sum_{j=1}^{H_m} a_j(\xi)b_j(\eta)D_j\Gamma(\xi_0;\cdot) + a_0(\xi)b_0(\eta)D_0\Gamma(\xi_0;\cdot)\right\}(\Theta(\eta,\xi))$$

where $D_0\Gamma(\xi_0;\cdot)$ has bounded $Y_i$-derivatives ($i = 1,2,...,q$). We can consider each of the terms

$$T_j\left(\xi_0\right)\widetilde{X}_i f\left(\xi\right) \equiv \int a_j\left(\xi\right) b_j\left(\eta\right) D^{\eta}_j\Gamma(\xi_0;\Theta(\eta,\xi))\widetilde{X}_i f\left(\eta\right) d\eta$$

(this time it is important to recall the $\eta$-dependence of the coeffi cients of$D_j$) and distinguish 2 cases:

(i) $D_j\Gamma$ is homogeneous of degree $\geq 2-Q$ or it is regular (i.e. $D_j\Gamma$ has bounded $Y_i$-derivatives);

(ii) $T_j\left(\xi_0\right)$ is a frozen operator of type 1 and $D_j\Gamma$ is homogeneous of degree $1-Q$.

Case (i). We can integrate by parts, recalling that the transpose of $\widetilde{X}_i$ is

$$\left(\widetilde{X}_i\right)^T g\left(\eta\right) = -\widetilde{X}_i g\left(\eta\right) + c_i\left(\eta\right) g\left(\eta\right)$$

with $c_i$ smooth functions:

$$
\begin{aligned}
T_j(\xi_0)\widetilde{X}_i f(\xi) &= \int c_i(\eta)\, a_j(\xi)\, b_j(\eta)\, D_j^\eta \Gamma(\xi_0;\Theta(\eta,\xi)) f(\eta)\, d\eta \\
&- \int a_j(\xi) \left(\widetilde{X}_i b_j\right)(\eta)\, D_j^\eta \Gamma(\xi_0;\Theta(\eta,\xi)) f(\eta)\, d\eta \\
&- \int a_j(\xi)\, b_j(\eta)\, \widetilde{X}_i \left[D_j^\eta \Gamma(\xi_0;\Theta(\cdot,\xi))\right](\eta)\, f(\eta)\, d\eta \\
&- \int a_j(\xi)\, b_j(\eta) \left(\widetilde{X}_i^\eta D_j^\eta\right) \Gamma(\xi_0;\Theta(\eta,\xi)) f(\eta)\, d\eta \\
&= A(\xi) + B(\xi) + C(\xi) + D(\xi).
\end{aligned}
$$

Now, $A(\xi) + B(\xi)$ is still an operator of type $\lambda$, applied to $f$; in particular, it can be seen as operator of type $\lambda - 1$; the same is true for $D(\xi)$, by Remark 4.6. To study $C(\xi)$, we apply Lemma 4.14,

$$
\widetilde{X}_i \left[D_j^\eta \Gamma(\xi_0;\Theta(\cdot,\xi))\right](\eta) = -\left(Y_i^R D_j^\eta \Gamma\right)(\xi_0, \Theta(\eta,\xi)) + \left(\left(R_i^\xi\right)' D_j^\eta \Gamma\right)(\xi_0, \Theta(\eta,\xi)).
$$

Since $Y_i^R$ is homogeneous of degree 1, $a_j(\xi)\, b_j(\eta)\, Y_i^R D_j^\eta \Gamma(\xi_0, \Theta(\eta,\xi))$ is a kernel of type $\lambda - 1$. Since $\left(R_i^\xi\right)'$ is a differential operator of degree $\leq 0$, the kernel $a_j(\xi)\, b_j(\eta) \left(\left(R_i^\xi\right)' D_j^\eta \Gamma\right)(\xi_0, \Theta(\eta,\xi))$ is of type $\lambda$.

Note that, even when the coeffi cients of the differential operator $D_j$ (in the expression $D_j \Gamma(\xi_0;\Theta(\eta,\xi))$) do not depend on $\xi$ and $\eta$, this procedure introduces, with the operator $\left(R_i^\xi\right)'$, a new $\xi$-dependence of the coeffi cients. Compare with what we have remarked in the proof of Lemma 4.12.

Case (ii). In this case the kernel $\left(Y_i^R D_j \Gamma\right)$ is singular, so that the computation must be handled with more care. We can write

$$
\begin{aligned}
&T_j(\xi_0)\widetilde{X}_i f(\xi) = \\
&= \lim_{\varepsilon \to 0} \int a_j(\xi)\, b_j(\eta)\, \varphi_\varepsilon(\Theta(\xi,\eta))\, D_j \Gamma(\xi_0;\Theta(\eta,\xi)) \widetilde{X}_i f(\eta)\, d\eta \equiv \lim_{\varepsilon \to 0} T_\varepsilon(\xi)
\end{aligned}
$$

with $\varphi_\varepsilon$ as in the proof of Lemma 4.13. Note that, choosing a radial $\varphi$, we have $\varphi_\varepsilon(\Theta(\xi,\eta)) = \varphi_\varepsilon(\Theta(\eta,\xi))$. Then

$$
\begin{aligned}
T_\varepsilon(\xi) &= \int c_i(\eta)\, a_j(\xi)\, b_j(\eta)\, \varphi_\varepsilon(\Theta(\xi,\eta))\, D_j \Gamma(\xi_0;\Theta(\eta,\xi)) f(\eta)\, d\eta \\
&- \int a_j(\xi) \left(\widetilde{X}_i b_j\right)(\eta)\, \varphi_\varepsilon(\Theta(\xi,\eta))\, D_j \Gamma(\xi_0;\Theta(\eta,\xi)) f(\eta)\, d\eta \\
&- \int a_j(\xi)\, b_j(\eta)\, \widetilde{X}_i \left[\varphi_\varepsilon(\Theta(\cdot,\xi))\, D_j \Gamma(\xi_0;\Theta(\cdot,\xi))\right](\eta)\, f(\eta)\, d\eta \\
&- \int a_j(\xi)\, b_j(\eta)\, \varphi_\varepsilon(\Theta(\xi,\eta)) \left(\widetilde{X}_i^\eta D_j^\eta\right) \Gamma(\xi_0;\Theta(\eta,\xi)) f(\eta)\, d\eta \\
&= A_\varepsilon(\xi) + B_\varepsilon(\xi) + C_\varepsilon(\xi) + D_\varepsilon(\xi).
\end{aligned}
$$

Now $A_\varepsilon(\xi)+B_\varepsilon(\xi)+D_\varepsilon(\xi)$ converge to an operator of type $\lambda$, as $A(\xi), B(\xi), D(\xi)$ are in Case (i), while by Theorem 3.2 and Lemma 4.14

$$
\begin{aligned}
C_\varepsilon(\xi) &= -\int a_j(\xi)\, b_j(\eta)\, f(\eta)\, (Y_i\varphi_\varepsilon)(\Theta(\eta,\xi))\, D_j\Gamma(\xi_0;\Theta(\eta,\xi))d\eta \\
&- \int a_j(\xi)\, b_j(\eta)\, f(\eta)\left(R_i^\xi\varphi_\varepsilon\right)(\Theta(\eta,\xi))\, D_j\Gamma(\xi_0;\Theta(\eta,\xi))d\eta \\
&+ \int a_j(\xi)\, b_j(\eta)\, f(\eta)\,\varphi_\varepsilon(\Theta(\eta,\xi))\left(Y_i^R D_j\Gamma\right)(\xi_0,\Theta(\eta,\xi))\,d\eta \\
&- \int a_j(\xi)\, b_j(\eta)\, f(\eta)\,\varphi_\varepsilon(\Theta(\eta,\xi))\left(\left(R_i^\xi\right)' D_j\Gamma\right)(\xi_0,\Theta(\eta,\xi))\,d\eta \\
&\equiv E_\varepsilon(\xi)+F_\varepsilon(\xi)+G_\varepsilon(\xi)+H_\varepsilon(\xi).
\end{aligned}
$$

Now: $H_\varepsilon(\xi)$ tends to an operator of type 1; $G_\varepsilon(\xi)$ tends to

$$
P.V.\int a_j(\xi)\, b_j(\eta)\, f(\eta)\left(Y_i^R D_j\Gamma\right)(\xi_0,\Theta(\eta,\xi))\,d\eta,
$$

which is an operator of type 0. As to $E_\varepsilon(\xi)$, the same computation performed in the proof of Lemma 4.13 gives

$$
E_\varepsilon(\xi)\to\alpha(\xi_0,\xi)(abcf)(\xi)
$$

with

$$
\alpha(\xi_0,\xi)=\int Y_i\varphi(v)\, D_1^\xi\Gamma(\xi_0;v)\,dv
$$

which is the multiplicative part of an operator of type 0. A similar computation shows that $F_\varepsilon(\xi)\to 0$, so we are done. ∎

Let us come to the second important result of this section. In [25, Corollary p. 296], the following fact is proved in the case of a family of Hörmanderś vector fields of weight one (that is, without the drift $\widetilde{X}_0$): for any frozen operator $T(\xi_0)$ of type 1, $i=1,2,...,q$, there exist operators $T_{ij}(\xi_0), T_i(\xi_0)$ of type 1 such that

$$
\widetilde{X}_i T(\xi_0)=\sum_{j=1}^{q} T_{ij}(\xi_0)\,\widetilde{X}_j+T_i(\xi_0).
$$

This possibility of exchanging the order of integral and differential operators will be crucial in the proof of representation formulas. However, such an identity cannot be proved in this form when the drift $\widetilde{X}_0$ is present. Instead, we are going to prove the following, which will be enough for our purposes:

**Theorem 4.15** *If $T(\xi_0)$ is a frozen operator of type $\lambda\geq 1$, $k_0=1,2\ldots q$, then*

$$
\widetilde{X}_{k_0}T(\xi_0)=\sum_{k=1}^{q} T_k^{k_0}(\xi_0)\,\widetilde{X}_k+\sum_{h,j=1}^{q}\widetilde{a}_{hj}(\xi_0)\,T^{hk_0}(\xi_0)\,\widetilde{X}_j+T_0^{k_0}(\xi_0)+T^{k_0}(\xi_0)\,\widetilde{\mathcal{L}}_0, \tag{4.15}
$$

*where* $T_k^{k_0}(\xi_0)$ $(k = 0, 1, ..., q)$ *and* $T^{hk_0}(\xi_0)$ *are frozen operators of type* $\lambda$, $T^{k_0}(\xi_0)$ *are frozen operators of type* $\lambda + 1$, *and* $\widetilde{a}_{hj}(\xi_0)$ *are the frozen coefficients of* $\widetilde{\mathcal{L}}_0$.

*If* $T(\xi_0)$ *is a frozen operator of type* $\lambda \geq 2$, *then*

$$\widetilde{X}_0 T(\xi_0) = \sum_{k=1}^{q} T_k(\xi_0)\widetilde{X}_k + \sum_{h,j=1}^{q} \widetilde{a}_{hj}(\xi_0) T^h(\xi_0)\widetilde{X}_j + T_0(\xi_0) + T(\xi_0)\widetilde{\mathcal{L}}_0, \quad (4.16)$$

*where* $T_k(\xi_0)$ $(k = 0, 1, ..., q)$ *and* $T^h(\xi_0)$ *are frozen operators of type* $\lambda - 1$, $T(\xi_0)$ *is a frozen operator of type* $\lambda$.

We start with the following lemma, which is similar to that proved in [25, p. 296]. Again, we prefer to present a detailed proof since the one given in [25] is very condensed.

**Lemma 4.16** *For any vector field* $\widetilde{X}_{j_0,k_0}$ $(j_0 = 1, 2, ..., s,\ k_0 = 1, 2, ..., k_{j_0})$ *there exist smooth functions* $\left\{a_{jk}^{j_0 k_0 \eta}\right\}_{\substack{j=1,2,...,s \\ k=1,2,...h_j}}$ *having local degree* $\geq \max\{j - j_0, 0\}$ *and smoothly depending on* $\eta$, *such that for any* $f \in C_0^\infty(\mathbb{G})$, *one can write*

$$\begin{aligned} &\widetilde{X}_{j_0,k_0}[f(\Theta(\eta,\cdot))](\xi) = \qquad (4.17) \\ &= \sum_{\substack{j=1,2,...,s \\ k=1,2,...,k_j}} a_{jk}^{j_0 k_0 \eta}(\Theta(\eta,\xi))\widetilde{X}_{j,k}[f(\Theta(\cdot,\xi))](\eta) + \left(R_{j_0}^{\xi,\eta} f\right)(\Theta(\eta,\xi)) \end{aligned}$$

*where* $R_{j_0}^{\xi,\eta}$ *is a vector field of local degree* $\leq j_0 - 1$, *smoothly depending on* $\xi, \eta$.

**Proof.** By Theorem 3.2 we know that

$$\widetilde{X}_{j_0,k_0}[f(\Theta(\eta,\cdot))](\xi) = \left(Y_{j_0,k_0} f + R_{j_0,k_0}^{\eta} f\right)(\Theta(\eta,\xi)) \equiv \left(Z_{j_0}^{\eta} f\right)(\Theta(\eta,\xi)), \quad (4.18)$$

where $Z_{j_0}^{\eta}$ is a vector field of local degree $\leq j_0$, smoothly depending on $\eta$. To rewrite $\left(Z_{j_0}^{\eta} f\right)$ in the suitable form, we start from the following identities:

$$Y_{i,k} = \frac{\partial}{\partial u_{ik}} + \sum_r \sum_{i<l\leq s} g_{lr}^{ik}(u)\frac{\partial}{\partial u_{lr}} \quad (4.19)$$

for any $i = 1, 2, ..., s$ and $k = 1, 2, ..., k_i$;

$$Y_{i,k} = \sum g_{lr}^{ik}(u) Y_{l,r}^{R}, \quad (4.20)$$

where $g_{lr}^{ik}(u)$ are homogeneous of degree $l - i$ (see [25, p. 295]). Hence we can write

$$Z_{j_0}^{\eta} = \sum a_{jk}^{\eta}(u)\frac{\partial}{\partial u_{jk}},$$

where $a_{jk}$ has local degree $\geq j - j_0$ and smoothly depends on $\eta$. By inverting (for any $i, k$) the triangular system (4.19), we obtain

$$\frac{\partial}{\partial u_{jk}} = Y_{j,k} + \sum_{j<l\leq s} f_{lr}^{jk}(u)\, Y_{l,r},$$

where each $f_{lr}^{jk}(u)$ is homogeneous of degree $l - j$. Using also (4.20), we have

$$\begin{aligned} \left(Z_{j_0}^{\eta} f\right)(u) &= \sum a_{jk}^{\eta}(u) \left[ \left(Y_{j,k} f\right)(u) + \sum_{j<l\leq s} f_{lr}^{jk}(u) \left(Y_{l,r} f\right)(u) \right] \\ &= \sum b_{lr}^{\eta}(u) \left(Y_{l,r}^{R} f\right)(u), \end{aligned} \tag{4.21}$$

where

$$b_{lr}^{\eta} \text{ has local degree } \geq \max\{l - j_0, 0\} \tag{4.22}$$

and smoothly depends on $\eta$. By Lemma 4.14, then

$$\begin{aligned} \left(Z_{j_0}^{\eta} f\right)(\Theta(\eta,\xi)) &= \sum_{l,r} -b_{lr}^{\eta}(\Theta(\eta,\xi))\, \widetilde{X}_{l,r}\left[f(\Theta(\cdot,\xi))\right](\eta) \\ &\quad + \sum_{l,r} \left(b_{lr}^{\eta} \left(R_{l,r}^{\xi}\right)' f\right)(\Theta(\eta,\xi)), \end{aligned} \tag{4.23}$$

where $\left(R_{l,r}^{\xi}\right)'$ is a differential operator of local degree $\leq l - 1$, hence the differential operator on $\mathbb{G}$

$$R_{j_0}^{\xi,\eta} \equiv \sum_{l,r} b_{lr}^{\eta} \left(R_{l,r}^{\xi}\right)' \text{ has local degree } \leq j_0 - 1, \tag{4.24}$$

and depends smoothly on $\xi, \eta$. Collecting (4.18), (4.22), (4.23), (4.24), the Lemma is proved, with $a_{jk}^{j_0 k_0 \eta} = -b_{jk}^{\eta}$. ■

With this lemma in hand, we can prove the following, similar to [25, Thm 9]:

**Theorem 4.17** *(i) Suppose $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$. Given a vector field $\widetilde{X}_i$ for $i = 1, 2, ..., q$, there exist $T^i(\xi_0)$, frozen operator of type $\lambda$, and $T_{jk}^i(\xi_0)$, frozen operators of type $\lambda + j - 1$, such that:*

$$\widetilde{X}_i T(\xi_0) = \sum_{j,k} T_{jk}^i(\xi_0)\, \widetilde{X}_{j,k} + T^i(\xi_0); \tag{4.25}$$

*(ii) Suppose $T(\xi_0)$ is a frozen operator of type $\lambda \geq 2$. There exist $T^0(\xi_0), T_{jk}^i(\xi_0)$ frozen operators of type $\lambda - 1$, $\lambda + \max\{j - 2, 0\}$, respectively, such that:*

$$\widetilde{X}_0 T(\xi_0) = \sum_{j,k} T_{jk}^0(\xi_0)\, \widetilde{X}_{j,k} + T^0(\xi_0); \tag{4.26}$$

**Proof.** First of all, it is enough to consider the part $k'$ of the kernel of $T(\xi_0)$, the proof for $k''$ being completely analogous (see Definition 4.5).

(i) If $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$ with kernel $k'$ we can write it as

$$T(\xi_0) f(\xi) = \int a(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta) d\eta + T'(\xi_0) f(\xi)$$

where $D\Gamma(\xi_0, \cdot)$ is homogeneous of degree $\lambda - Q$ and $T'(\xi_0)$ is a frozen operator of degree $\lambda + 1$. Since $\widetilde{X}_i T'(\xi_0)$ is a frozen operator of type $\lambda$, it has already the form $T^i(\xi_0)$ required by the theorem, hence it is enough to prove that

$$\widetilde{X}_i \int a(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta) d\eta$$

can be rewritten in the form

$$\sum_{j,k} T^i_{jk}(\xi_0) \widetilde{X}_{j,k} f(\xi) + T^i(\xi_0) f(\xi)$$

with $T^i_{jk}(\xi_0), T^i(\xi_0)$ frozen operators of type $\lambda + j - 1$ and $\lambda$, respectively. Next, we have to distinguish two cases.

Case 1: $\lambda \geq 2$. In this case the $\widetilde{X}_i$ derivative can be taken under the integral sign, writing:

$$\begin{aligned}
&\widetilde{X}_i \int a(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta) d\eta \\
&= \int \left(\widetilde{X}_i a\right)(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta) d\eta \\
&+ \int a(\xi) \widetilde{X}_i [D\Gamma(\Theta(\eta, \cdot))](\xi) b(\eta) f(\eta) d\eta \\
&= A(\xi) + B(\xi).
\end{aligned}$$

Now $A(\xi)$ is frozen operator of type $\lambda$, while applying Lemma 4.16 with $j_0 = 1$ we get

$$\begin{aligned}
B(\xi) &= \int a(\xi) \sum_{l,r} a^i_{lr}(\Theta(\eta, \xi)) \widetilde{X}_{l,r} [D\Gamma(\xi_0; \Theta(\cdot, \xi))](\eta) b(\eta) f(\eta) d\eta \\
&+ \int a(\xi) \sum_{l,r} a^i_{lr}(\Theta(\eta, \xi)) \left(R^\xi_{rs} D\Gamma\right)(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta) d\eta \\
&\equiv C(\xi) + D(\xi)
\end{aligned}$$

where $R^\xi_{rs}$ are differential operators of local degree $\leq 0$, and the $a^i_{lr}$'s have local degree $\geq l - 1$. Hence $D$ is a frozen operator of type $\lambda$, while, since the transposed vector field of $\widetilde{X}_{l,r}$ is

$$\widetilde{X}^T_{l,r} = -\widetilde{X}_{l,r} + c_{l,r}$$

with $c_{l,r}$ smooth functions,

$$C(\xi) = -a(\xi)\sum_{l,r}\int \widetilde{X}_{l,r}\left[a^i_{lr}(\Theta(\cdot,\xi))\,b(\cdot)\right](\eta)\,D\Gamma(\xi_0;\Theta(\eta,\xi))\,f(\eta)\,d\eta$$
$$+\,a(\xi)\sum_{l,r}\int a^i_{lr}(\Theta(\eta,\xi))\,D\Gamma(\xi_0;\Theta(\eta,\xi))\,c_{l,r}(\eta)\,b(\eta)\,f(\eta)\,d\eta$$
$$-\,a(\xi)\sum_{l,r}\int a^i_{lr}(\Theta(\eta,\xi))\,D\Gamma(\xi_0;\Theta(\eta,\xi))\,b(\eta)\,\widetilde{X}_{l,r}f(\eta)\,d\eta.$$

The first two terms in the last expression are still frozen operators of type $\lambda$ applied to $f$, while the third is a sum of operators of type $\lambda + l - 1$ applied to $\widetilde{X}_{l,r}f$, as required by the theorem.

Case 2. $\lambda = 1$. In this case we have to compute the derivative of the integral in distributional sense, as already done in the proof of Lemma 4.13: with the same meaning of $\varphi_\varepsilon$, let us compute

$$\lim_{\varepsilon\to 0}\widetilde{X}_i\int a(\xi)\,\varphi_\varepsilon(\Theta(\eta,\xi))\,D\Gamma(\xi_0;\Theta(\eta,\xi))\,b(\eta)\,f(\eta)\,d\eta.$$

Actually, this gives exactly the same result as in case 1:

$$\widetilde{X}_i\int a(\xi)\,\varphi_\varepsilon(\Theta(\eta,\xi))\,D\Gamma(\xi_0;\Theta(\eta,\xi))\,b(\eta)\,f(\eta)\,d\eta$$
$$=\int\left(\widetilde{X}_i a\right)(\xi)\,\varphi_\varepsilon(\Theta(\eta,\xi))\,D\Gamma(\xi_0;\Theta(\eta,\xi))\,b(\eta)\,f(\eta)\,d\eta$$
$$+\int a(\xi)\,\widetilde{X}_i\left[(\varphi_\varepsilon D\Gamma)(\Theta(\eta,\cdot))\right](\xi)\,b(\eta)\,f(\eta)\,d\eta$$
$$= A_\varepsilon(\xi) + B_\varepsilon(\xi)$$

where

$$A_\varepsilon(\xi) \to \int\left(\widetilde{X}_i a\right)(\xi)\,D\Gamma(\xi_0;\Theta(\eta,\xi))\,b(\eta)\,f(\eta)\,d\eta$$

and

$$B_\varepsilon(\xi) = \int a(\xi)\sum_{l,r}a^i_{lr}(\Theta(\eta,\xi))\,\widetilde{X}_{l,r}\left[\varphi_\varepsilon(\Theta(\cdot,\xi))\,D\Gamma(\xi_0;\Theta(\cdot,\xi))\right](\eta)\,b(\eta)\,f(\eta)\,d\eta$$
$$+\int a(\xi)\sum_{l,r}a^i_{lr}(\Theta(\eta,\xi))\left(R^\xi_{rs}(\varphi_\varepsilon D\Gamma)\right)(\xi_0;\Theta(\eta,\xi))\,b(\eta)\,f(\eta)\,d\eta$$
$$\equiv C_\varepsilon(\xi) + D_\varepsilon(\xi)$$

where $C_\varepsilon(\xi)$ converges to the expression called $C(\xi)$ in the computation of case 1; as to $D_\varepsilon(\xi)$,

$$R^\xi_{rs}(\varphi_\varepsilon D\Gamma) = \left(R^\xi_{rs}\varphi_\varepsilon\right)D\Gamma + \varphi_\varepsilon R^\xi_{rs}D\Gamma.$$

Now, $\varphi_\varepsilon R^\xi_{rs} D\Gamma \to R^\xi_{rs} D\Gamma$ while $\left(R^\xi_{rs}\varphi_\varepsilon\right) D\Gamma \to 0$, being $R^\xi_{rs}$ a vector field of local degree $\leq 0$. Hence also $D_\varepsilon(\xi)$ converges to the expression called $D(\xi)$ in the computation of case 1, and we are done.

(ii) Let now $T(\xi_0)$ be a frozen operator of type $\lambda \geq 2$ with kernel $k'$. As in the case (i), it is enough to prove that

$$\widetilde{X}_0 \int a(\xi)\, D\Gamma(\xi_0; \Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta,$$

where $D\Gamma$ is homogeneous of degree $\lambda - Q$, can be rewritten in the form

$$\sum_{j,k} T^0_{jk}(\xi_0)\, \widetilde{X}_{j,k} f(\xi) + T^0(\xi_0)\, f(\xi)$$

with $T^0_{jk}(\xi_0), T^0(\xi_0)$ frozen operators of type $\lambda + j - 2$ and $\lambda - 1$, respectively. Let us consider only the case $\lambda \geq 3$, the case $\lambda = 2$ being handled with the modification seen in (i), Case 2.

$$\begin{aligned}
&\widetilde{X}_0 \int a(\xi)\, D\Gamma(\xi_0; \Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta \\
&= \int \left(\widetilde{X}_0 a\right)(\xi)\, D\Gamma(\xi_0; \Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta + \\
&+ \int a(\xi)\, \widetilde{X}_0 \left[D\Gamma(\xi_0; \Theta(\eta,\cdot))\right](\xi)\, b(\eta)\, f(\eta)\, d\eta \\
&= \int \left(\widetilde{X}_0 a\right)(\xi)\, D\Gamma(\xi_0; \Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta + \\
&+ \int a(\xi) \sum_{l,r} a^0_{lr}(\Theta(\eta,\xi))\, \widetilde{X}_{l,r}\left[D\Gamma(\xi_0; \Theta(\cdot,\xi))\right](\eta)\, b(\eta)\, f(\eta)\, d\eta \\
&+ \int a(\xi) \sum_{l,r} a^0_{lr}(\Theta(\eta,\xi)) \left(R^\xi_{rs} D\Gamma\right)(\xi_0; \Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta \\
&\equiv A(\xi) + C(\xi) + D(\xi).
\end{aligned}$$

where $R^\xi_{rs}$ are now differential operators of local degree $\leq 1$, and the $a^0_{lr}$'s have local degree $\geq \max\{j-2, 0\}$. Then $A(\xi)$ is a frozen operator of type $\lambda$, applied to $f$; $D(\xi)$ is a frozen operator of type $\lambda - 1$, applied to $f$. Moreover,

$$\begin{aligned}
C(\xi) = &-a(\xi) \sum_{l,r} \int \widetilde{X}_{l,r}\left[a^0_{lr}(\Theta(\cdot,\xi))\, b(\cdot)\right](\eta)\, D\Gamma(\xi_0; \Theta(\eta,\xi))\, f(\eta)\, d\eta \\
&+ a(\xi) \sum_{l,r} \int a^0_{lr}(\Theta(\eta,\xi))\, D\Gamma(\xi_0; \Theta(\eta,\xi))\, c_{l,r}(\eta)\, b(\eta)\, f(\eta)\, d\eta \\
&- a(\xi) \sum_{l,r} \int a^0_{lr}(\Theta(\eta,\xi))\, D\Gamma(\xi_0; \Theta(\eta,\xi))\, b(\eta)\, \widetilde{X}_{l,r} f(\eta)\, d\eta
\end{aligned}$$

where the first two terms are still frozen operators of type $\lambda$, applied to $f$, while the third is the sum of frozen operators of type $\lambda + \max\{j-2, 0\}$ applied to $\widetilde{X}_{l,r} f$. ■

We can now proceed to the:

**Proof of Theorem 4.15.** It suffi ces to prove the formula (4.15), for the second is similar. Let us consider one of the terms $T^i_{jk}(\xi_0)\widetilde{X}_{j,k}$ appearing in (4.25).

If $j=1$, the term is already in the form required by the Theorem.

If $j=2$, then $\widetilde{X}_{2,k}$ can be written as a combination of commutators of the vector fields $\widetilde{X}_1,\widetilde{X}_2,...,\widetilde{X}_q$, plus (possibly) the field $\widetilde{X}_0$. Then $T^i_{2k}(\xi_0)\widetilde{X}_{2,k}$ contains terms $T^i_{2k}(\xi_0)\widetilde{X}_h\widetilde{X}_j$ and possibly a term $T^i_{2k}(\xi_0)\widetilde{X}_0$. By the above theorem we know $T^i_{2k}$ is a frozen operator of type $\lambda+1$. Now:

$$T^i_{2k}(\xi_0)\widetilde{X}_h\widetilde{X}_j=\left(T^i_{2k}(\xi_0)\widetilde{X}_h\right)\widetilde{X}_j=T^i_k(\xi_0)\widetilde{X}_j,$$

where by Theorem 4.11, $T^i_k(\xi_0)$ is a frozen operator of type $\lambda$; on the other hand, by (4.2),

$$\begin{aligned}T^i_{2k}(\xi_0)\widetilde{X}_0&=T^i_{2k}(\xi_0)\left(\widetilde{\mathcal{L}}_0-\sum_{h,j=1}^{q}\widetilde{a}_{hj}(\xi_0)\widetilde{X}_h\widetilde{X}_j\right)\\&=T^i_{2k}(\xi_0)\widetilde{\mathcal{L}}_0-\sum_{h,j=1}^{q}\widetilde{a}_{hj}(\xi_0)\left(T^i_{2k}(\xi_0)\widetilde{X}_h\right)\widetilde{X}_j\\&=T^i_{2k}(\xi_0)\widetilde{\mathcal{L}}_0-\sum_{h,j=1}^{q}\widetilde{a}_{hj}(\xi_0)T^i_{h,k}(\xi_0)\widetilde{X}_j,\end{aligned}$$

with $T^i_{2k}(\xi_0)$, $T^i_{h,k}(\xi_0)$, frozen operators of type $\lambda+1,\lambda$, which is in the form allowed by the thesis of the Theorem.

Finally, if $j>2$, it is enough to look at the final part of the differential operator $\widetilde{X}_{j,k}$: it is always possible to rewrite $\widetilde{X}_{j,k}$ either as $\widetilde{X}_{j-1,k}\widetilde{X}_{1,k}$ or as $\widetilde{X}_{j-2,k}\widetilde{X}_{2,k}$. In the first case, we have

$$T^i_{jk}(\xi_0)\widetilde{X}_{j,k}=\left(T^i_{jk}(\xi_0)\widetilde{X}_{j-1,k}\right)\widetilde{X}_{1,k}=T'^i_{jk}(\xi_0)\widetilde{X}_{1,k},$$

with $T'^i_{jk}(\xi_0)$ frozen operator of type $\lambda$, which is already in the proper form; in the second case, we have

$$T^i_{jk}(\xi_0)\widetilde{X}_{j,k}=\left(T^i_{jk}(\xi_0)\widetilde{X}_{j-2,k}\right)\widetilde{X}_{2,k}=T'^i_j(\xi_0)\widetilde{X}_{2,k}$$

with $T'^i_{jk}(\xi_0)$ frozen operator of type $\lambda+1$, and then we can proceed as in the case $j=2$. So the Theorem is proved. ■

## 4.3 Parametrix and representation formulas

Throughout this subsection we will make extensive use of computations on frozen operators of type $\lambda$. To make more readable our formulas, we will use the symbols

$$T(\xi_0),S(\xi_0),P(\xi_0)$$

(with possibly other indexes) to denote frozen operators of type $0,1,2$, respectively.

In order to prove representation formulas for second order derivatives, we start with the following parametrix identities, analogous to [25, Thm. 10], [3, Thm. 3.1].

**Theorem 4.18** *Given* $a \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$*, there exist* $S_{ij}(\xi_0)$*,* $S_0\left(\xi_0\right)$*,* $S_{ij}^*(\xi_0)$*,* $S_0^*\left(\xi_0\right)$*, frozen operators of type* 1 *and* $P(\xi_0), P^*(\xi_0)$*, frozen operators of type* 2 *(over the ball* $\widetilde{B}\left(\overline{\xi},R\right)$*) such that:*

$$aI = \widetilde{\mathcal{L}}_0^T P^*\left(\xi_0\right) + \sum_{i,j=1}^q \widetilde{a}_{ij}\left(\xi_0\right) S_{ij}^*\left(\xi_0\right) + S_0^*\left(\xi_0\right);$$

$$aI = P(\xi_0)\widetilde{\mathcal{L}}_0 + \sum_{i,j=1}^q \widetilde{a}_{ij}\left(\xi_0\right) S_{ij}\left(\xi_0\right) + S_0\left(\xi_0\right) \tag{4.27}$$

*where* $I$ *denotes the identity. Moreover,* $S_{ij}^*\left(\xi_0\right), S_0^*\left(\xi_0\right), P^*\left(\xi_0\right)$ *are modeled on* $\Gamma^T$*, while* $S_{ij}\left(\xi_0\right), S_0\left(\xi_0\right), P\left(\xi_0\right)$ *are modeled on* $\Gamma$*. Explicitly,*

$$P^*\left(\xi_0\right) f\left(\xi\right) = -\frac{a(\xi)}{c(\xi)} \int_{\widetilde{B}} \Gamma^T(\xi_0;\Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta$$

$$P\left(\xi_0\right) f\left(\xi\right) = -b\left(\xi\right) \int_{\widetilde{B}} \frac{a(\eta)}{c(\eta)} \Gamma(\xi_0;\Theta(\eta,\xi))\, f(\eta)\, d\eta$$

*where* $c$ *is the function appearing in Theorem 3.3 (c).*

The proof of this result is similar to that of [25], [3, Thm. 3.1]. However, we will write a detailed version.

**Proof.** Let us define

$$P^*(\xi_0)f(\xi) = \frac{a(\xi)}{c(\xi)} \int_{\widetilde{B}} \Gamma^T(\xi_0;\Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta$$

where $a, b \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$ such that $ab = a$, and $c\left(\xi\right)$ is the function appearing in the formula of change of variables (3.7), and let us compute $\widetilde{\mathcal{L}}_0^T P^*(\xi_0)f$, for $f \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$. We have to apply a distributional argument like in the proof of Lemma 4.13: for $\omega \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$, let us evaluate

$$\int_{\widetilde{B}} \widetilde{\mathcal{L}}_0\omega\left(\xi\right) P^*(\xi_0)f\left(\xi\right) d\xi = \lim_{\varepsilon\to 0} \int_{\widetilde{B}} \widetilde{\mathcal{L}}_0\omega\left(\xi\right) P_\varepsilon^*(\xi_0)f\left(\xi\right) d\xi$$

where

$$P_\varepsilon^*(\xi_0)f(\xi) = \frac{a(\xi)}{c(\xi)} \int_{\widetilde{B}} \varphi_\varepsilon\left(\Theta(\eta,\xi)\right) \Gamma^T(\xi_0;\Theta(\eta,\xi))\, b(\eta)\, f(\eta)\, d\eta$$

with $\varphi_\varepsilon$ as in the proof of the quoted Lemma.

$$\int_{\widetilde{B}} \widetilde{\mathcal{L}}_0\omega\left(\xi\right) P_\varepsilon^*(\xi_0) f\left(\xi\right) d\xi$$
$$= \int_{\widetilde{B}} bf(\eta) \left( \int_{\widetilde{B}} \frac{a(\xi)}{c(\xi)} \varphi_\varepsilon\left(\Theta(\eta,\xi)\right) \Gamma^T(\xi_0;\Theta(\eta,\xi)) \widetilde{\mathcal{L}}_0\omega\left(\xi\right) d\xi \right) d\eta$$
$$= \int_{\widetilde{B}} bf(\eta) \left( \int_{\widetilde{B}} \widetilde{\mathcal{L}}_0^T \left( \frac{a(\xi)}{c(\xi)} \right) \varphi_\varepsilon\left(\Theta(\eta,\xi)\right) \Gamma^T(\xi_0;\Theta(\eta,\xi)) \omega\left(\xi\right) d\xi \right) d\eta$$
$$+ \int_{\widetilde{B}} bf(\eta) \left( \int_{\widetilde{B}} \frac{a(\xi)}{c(\xi)} \widetilde{\mathcal{L}}_0^T \left[ \varphi_\varepsilon\left(\Theta(\eta,\cdot)\right) \Gamma^T(\xi_0;\Theta(\eta,\cdot)) \right] (\xi)\, \omega\left(\xi\right) d\xi \right) d\eta$$
$$+ \int_{\widetilde{B}} bf(\eta) \left( \int_{\widetilde{B}} 2 \sum_{h,k=1}^{q} \widetilde{a}_{hk}\left(\xi_0\right) \widetilde{X}_h^T \left(\frac{a}{c}\right)(\xi) \widetilde{X}_k^T \left[ \varphi_\varepsilon\left(\Theta(\eta,\cdot)\right) \Gamma^T(\xi_0;\Theta(\eta,\cdot)) \right] (\xi)\, \omega\left(\xi\right) d\xi \right) d\eta$$
$$\equiv A_\varepsilon + B_\varepsilon + C_\varepsilon.$$

$$A_\varepsilon \to \int_{\widetilde{B}} bf(\eta) \left( \int_{\widetilde{B}} \widetilde{\mathcal{L}}_0^T \left( \frac{a(\xi)}{c(\xi)} \right) \Gamma^T(\xi_0;\Theta(\eta,\xi)) \omega\left(\xi\right) d\xi \right) d\eta$$
$$= \sum_{h,k=1}^{q} \widetilde{a}_{hk}\left(\xi_0\right) \int_{\widetilde{B}} f(\eta) P_{hk}^T\left(\xi_0\right) \omega\left(\eta\right) d\eta + \int_{\widetilde{B}} f(\eta) P_0^T\left(\xi_0\right) \omega\left(\eta\right) d\eta$$
$$= \sum_{h,k=1}^{q} \widetilde{a}_{hk}\left(\xi_0\right) \int_{\widetilde{B}} \omega\left(\xi\right) P_{hk}\left(\xi_0\right) f(\xi) d\xi + \int_{\widetilde{B}} \omega\left(\xi\right) P_0\left(\xi_0\right) f(\xi) d\xi$$

where

$$P_{hk}\left(\xi_0\right) f(\xi) = \widetilde{X}_h^T \widetilde{X}_k^T \left(\frac{a}{c}\right)(\xi) \left( \int_{\widetilde{B}} \Gamma^T(\xi_0;\Theta(\eta,\xi)) b\left(\eta\right) f(\eta) d\eta \right) d\xi$$
$$P_0\left(\xi_0\right) f(\xi) = \widetilde{X}_0^T \left(\frac{a}{c}\right)(\xi) \left( \int_{\widetilde{B}} \Gamma^T(\xi_0;\Theta(\eta,\xi)) b\left(\eta\right) f(\eta) d\eta \right) d\xi$$

are frozen operators of type 2, modeled on $\Gamma^T$.

$$B_\varepsilon = \int_{\widetilde{B}} bf(\eta) \left\{ \int_{\widetilde{B}} \frac{a(\xi)}{c(\xi)} \left[ \mathcal{L}_0^{*T} \left( \varphi_\varepsilon \Gamma^T(\xi_0;\cdot) \right) \right] \left(\Theta(\eta,\xi)\right) \omega\left(\xi\right) d\xi \right.$$
$$\left. + \left( \int_{\widetilde{B}} \frac{a(\xi)}{c(\xi)} \left( \sum_{i,j} \widetilde{a}_{ij}\left(\xi_0\right) \left[ Y_i R_j^\eta + R_i^\eta Y_j + R_i^\eta R_j^\eta \right] + R_0^\eta \right) \left( \varphi_\varepsilon \Gamma^T(\xi_0;\cdot) \right) \left(\Theta(\eta,\xi)\right) \omega\left(\xi\right) d\xi \right) \right\} d\eta$$
$$= \int_{\widetilde{B}} bf(\eta) \left( \int_{\|u\|<R} \mathcal{L}_0^{*T} \left( \varphi_\varepsilon\left(u\right) \Gamma^T(\xi_0;u) \right) (a\omega) \left( \Theta(\eta,\cdot)^{-1}\left(u\right) \right) \left(1 + O\left(\|u\|\right)\right) du \right) d\eta$$
$$+ \int_{\widetilde{B}} bf(\eta) \int_{\|u\|<R} \left( \sum_{i,j} \widetilde{a}_{ij}\left(\xi_0\right) \left[ Y_i R_j^\eta + R_i^\eta Y_j + R_i^\eta R_j^\eta \right] + R_0^\eta \right) \left( \varphi_\varepsilon\left(u\right) \Gamma^T(\xi_0;u) \right) \cdot$$
$$\cdot\, (a\omega) \left( \Theta(\eta,\cdot)^{-1}\left(u\right) \right) \left(1 + O\left(\|u\|\right)\right) du d\eta$$
$$= D_\varepsilon + E_\varepsilon.$$

To study $D_\varepsilon$, let $(a\omega)_\eta(u) \equiv (a\omega)\left(\Theta(\eta,\cdot)^{-1}(u)\right)$. Then

$$
\begin{aligned}
D_\varepsilon &= \int_{\widetilde{B}} bf(\eta) \int_{\|u\|<R} \varphi_\varepsilon(u)\,\Gamma^T(\xi_0;u)\mathcal{L}_0^*(a\omega)_\eta(u)\,dud\eta + \\
&+ \int_{\widetilde{B}} bf(\eta) \int_{\|u\|<R} \mathcal{L}_0^{*T}\left[\varphi_\varepsilon\Gamma^T(\xi_0;\cdot)\right](u)\,(a\omega)_\eta(u)\,O(\|u\|)\,dud\eta \\
&\equiv D_\varepsilon^1 + D_\varepsilon^2.
\end{aligned}
$$

$$
\begin{aligned}
D_\varepsilon^1 &\to \int_{\widetilde{B}} bf(\eta) \int_{\|u\|<R} \Gamma^T(\xi_0;u)\mathcal{L}_0^*(a\omega)_\eta(u)\,dud\eta = -\int_{\widetilde{B}} bf(\eta)\,(a\omega)_\eta(0)\,d\eta \\
&= -\int_{\widetilde{B}} bf(\eta)\,(a\omega)(\eta)\,d\eta = -\int_{\widetilde{B}} (af\omega)(\eta)d\eta.
\end{aligned}
$$

$$
\begin{aligned}
D_\varepsilon^2 &= \int_{\widetilde{B}} bf(\eta) \int_{\|u\|<R} \left(\mathcal{L}_0^{*T}\varphi_\varepsilon\right)(u)\,\Gamma^T(\xi_0;u)\,(a\omega)_\eta(u)\,O(\|u\|)\,dud\eta \\
&+ \int_{\widetilde{B}} bf(\eta) \int_{\|u\|<R} 2\sum_{i,j} \widetilde{a}_{ij}(\xi_0)\,(Y_i\varphi_\varepsilon)(u)\,Y_j\Gamma^T(\xi_0;u)\,(a\omega)_\eta(u)\,O(\|u\|)\,dud\eta.
\end{aligned}
$$

A dilation argument as in the proof of Lemma 4.13 then gives

$$D_\varepsilon^2 \to 0.$$

Moreover,

$$
\begin{aligned}
E_\varepsilon &\to \int_{\widetilde{B}} bf(\eta) \int_{\|u\|<R} \left( \sum_{i,j} \widetilde{a}_{ij}(\xi_0)\left[Y_iR_j^\eta + R_i^\eta Y_j + R_i^\eta R_j^\eta\right] + R_0^\eta \right) \Gamma^T(\xi_0;u)\cdot \\
&\cdot (a\omega)\left(\Theta(\eta,\cdot)^{-1}(u)\right)(1+O(\|u\|))\,dud\eta
\end{aligned}
$$

coming back to the original variables $\xi$ in the inner integral

$$
\begin{aligned}
&= \int_{\widetilde{B}} bf(\eta) \int_{\widetilde{B}} \left( \sum_{i,j} \widetilde{a}_{ij}(\xi_0)\left[Y_iR_j^\eta + R_i^\eta Y_j + R_i^\eta R_j^\eta\right]\Gamma^T(\xi_0;\cdot) + \right. \\
&\left. + R_0^\eta \Gamma^T(\xi_0;\cdot)\,(\Theta(\eta,\xi))\,\frac{(a\omega)(\xi)}{c(\xi)} \right) d\xi d\eta \\
&= \int_{\widetilde{B}} f(\eta) \left[ \sum_{i,j} \widetilde{a}_{ij}(\xi_0)\,S'_{ij}(\xi_0) + S'_0(\xi_0) \right] \omega(\eta)\,d\eta \\
&= \int_{\widetilde{B}} \omega(\eta) \left[ \sum_{i,j} \widetilde{a}_{ij}(\xi_0)\,S'^T_{ij}(\xi_0) + S'^T_0(\xi_0) \right] f(\eta)d\eta
\end{aligned}
$$

where $S'_{ij}(\xi_0), S'_0(\xi_0)$ are transposed of frozen operators of type one modeled on $\Gamma^T$, hence are frozen operators of type one, modeled on $\Gamma$ (see Proposition 4.10); therefore $S'^T_{ij}(\xi_0), S'^T_0(\xi_0)$ are frozen operators of type one, modeled on $\Gamma^T$. Analogously, one can check that

$$C_\varepsilon \to \sum_{h,k=1}^{q} \widetilde{a}_{hk}(\xi_0) \int_{\widetilde{B}} \omega(\xi) S''_{hk}(\xi_0)\ f(\xi) d\xi$$

where $S''_{hk}(\xi_0)$ are frozen operators of type one, modeled on $\Gamma^T$.

Hence we have proved that

$$\begin{aligned}\widetilde{\mathcal{L}}_0^T P^*(\xi_0) f &= P_1(\xi_0) f - af + \left[\sum_{i,j} \widetilde{a}_{ij}(\xi_0) S'^T_{ij}(\xi_0) + S'^T_0(\xi_0)\right] f \\ &= -af + \left[\sum_{i,j} \widetilde{a}_{ij}(\xi_0) S^*_{ij}(\xi_0) + S^*_0(\xi_0)\right] f\end{aligned}$$

since $S'^T_0(\xi_0) + P_1(\xi_0)$ is a frozen operator of type 1, and simplifying our notation with $S_{ij}$ in place of $S'^T_{ij}$. Note that $S^*_{ij}(\xi_0), S^*_0(\xi_0)$ are frozen operators of type 1, modeled on $\Gamma^T$. This proves the first identity in the statement of the theorem, apart from an immaterial change of sign in the definition of $P^*(\xi_0)$.

Next, let us transpose this identity, getting

$$P^{*T}(\xi_0)\widetilde{\mathcal{L}}_0 f(\xi) = \left(\sum_{ij} \widetilde{a}_{ij}(\xi_0) S^{*T}_{ij}(\xi_0) + S^{*T}_0(\xi_0)\right) f(\xi) - (af)(\xi).$$

Note that

$$\begin{aligned}P^{*T}(\xi_0) f(\xi) &= b(\xi) \int_{\widetilde{B}} \frac{a(\eta)}{c(\eta)} \Gamma^T(\xi_0; \Theta(\xi,\eta)) f(\eta)\, d\eta \\ &= b(\xi) \int_{\widetilde{B}} \frac{a(\eta)}{c(\eta)} \Gamma(\xi_0; \Theta(\eta,\xi)) f(\eta)\, d\eta,\end{aligned}$$

which is a frozen operator of type two, modeled on $\Gamma$. On the other hand, $S^{*T}_{ij}(\xi_0), S^{*T}_0(\xi_0)$ are transposed of frozen operators of type 1 modeled on $\Gamma^T$, therefore are frozen operators of type 1, modeled on $\Gamma$. This concludes the proof. ■

**Theorem 4.19 (Representation of $\widetilde{X}_m \widetilde{X}_l u$ by frozen operators)** *Given* $a \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi}, R\right)\right), \xi_0 \in \widetilde{B}\left(\overline{\xi}, R\right)$, *for any* $m, l = 1, 2, ..., q$, *there exist frozen oper-*

*ators over the ball* $\widetilde{B}\left(\overline{\xi},R\right)$, *such that for any* $u\in C_0^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$

$$\begin{aligned}\widetilde{X}_m\widetilde{X}_l\left(au\right)&=T_{lm}\left(\xi_0\right)\widetilde{\mathcal{L}}_0u+\sum_{k=1}^{q}T_{lm,k}\left(\xi_0\right)\widetilde{X}_ku+T_{lm}^0\left(\xi_0\right)u\\&+\sum_{i,j=1}^{q}\widetilde{a}_{ij}\left(\xi_0\right)\left\{\sum_{k=1}^{q}T_{lm,k}^{ij}\left(\xi_0\right)\widetilde{X}_ku+\sum_{h,k=1}^{q}\widetilde{a}_{hk}\left(\xi_0\right)T_{lm,h}^{\prime ij}\left(\xi_0\right)\widetilde{X}_ku+\right.\\&+\left.S_{lm}^{ij}\left(\xi_0\right)\widetilde{\mathcal{L}}_0u+T_{lm}^{ij}\left(\xi_0\right)u\right\}\end{aligned}\tag{4.28}$$

*(All the* $T_{\dots}\left(\xi_0\right)$ *are frozen operators of type 0,* $S_{lm}^{ij}\left(\xi_0\right)$ *are of type 1). Also,*

$$\begin{aligned}&\widetilde{X}_m\widetilde{X}_l\left(au\right)=T_{lm}\left(\xi_0\right)\widetilde{\mathcal{L}}u+T_{lm}\left(\xi_0\right)\left(\sum_{i,j=1}^{q}\left[\widetilde{a}_{ij}(\xi_0)-\widetilde{a}_{ij}(\cdot)\right]\widetilde{X}_i\widetilde{X}_ju\right)+\\&+\sum_{k=1}^{q}T_{lm,k}\left(\xi_0\right)\widetilde{X}_ku+T_{lm}^0\left(\xi_0\right)u+\\&+\sum_{i,j=1}^{q}\widetilde{a}_{ij}\left(\xi_0\right)\left\{\sum_{k=1}^{q}T_{lm,k}^{ij}\left(\xi_0\right)\widetilde{X}_ku+\sum_{h,k=1}^{q}\widetilde{a}_{hk}\left(\xi_0\right)T_{lm,h}^{\prime ij}\left(\xi_0\right)\widetilde{X}_ku+S_{lm}^{ij}\left(\xi_0\right)\widetilde{\mathcal{L}}u\right.\\&\left.+S_{lm}^{ij}\left(\xi_0\right)\left(\sum_{i,j=1}^{q}\left[\widetilde{a}_{ij}(\xi_0)-\widetilde{a}_{ij}(\cdot)\right]\widetilde{X}_i\widetilde{X}_ju\right)+T_{lm}^{ij}\left(\xi_0\right)u\right\}.\end{aligned}\tag{4.29}$$

**Remark 4.20** *The representation formulas of the above theorem have a cumbersome aspect, due to the presence of the coeffi cients*$\widetilde{a}_{ij}(\xi_0)$ *which appear several times as multiplicative factors. Anyway, if we agree to leave implicitly understood in the symbol of frozen operators the possible multiplication by the coeffi cients*$\widetilde{a}_{ij}$, *our formulas assume the following more compact form:*

$$\widetilde{X}_m\widetilde{X}_l\left(au\right)=T_{lm}\left(\xi_0\right)\widetilde{\mathcal{L}}_0u+\sum_{k=1}^{q}T_k^{lm}\left(\xi_0\right)\widetilde{X}_ku+T_{lm}^0\left(\xi_0\right)u$$

*and*

$$\begin{aligned}\widetilde{X}_m\widetilde{X}_l\left(au\right)&=T_{lm}\left(\xi_0\right)\widetilde{\mathcal{L}}u+T_{lm}\left(\xi_0\right)\left(\sum_{i,j=1}^{q}\left[\widetilde{a}_{ij}(\xi_0)-\widetilde{a}_{ij}(\cdot)\right]\widetilde{X}_i\widetilde{X}_ju\right)+\\&+\sum_{k=1}^{q}T_k^{lm}\left(\xi_0\right)\widetilde{X}_ku+T_{lm}^0\left(\xi_0\right)u.\end{aligned}$$

*In the proof of a priori estimates, when we will take* $C^\alpha$ *or* $L^p$ *norms of both sides of these identities, the multiplicative factors* $\widetilde{a}_{hj}$ *will be simply bounded by taking, respectively, the* $C^\alpha$ *or the* $L^\infty$ *norms of the* $\widetilde{a}_{hj}$*; hence leaving these factors implicitly understood is harmless.*

**Proof.** For $u \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$, let us start with the identity

$$au = P\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + \sum_{i,j=1}^{q} \widetilde{a}_{ij}\left(\xi_0\right) S_{ij}\left(\xi_0\right) u + S_0\left(\xi_0\right) u$$

(see Theorem 4.18); taking one derivative $\widetilde{X}_l$ $(l = 1, 2, ..., q)$ and applying Theorem 4.11 and Theorem 4.15, we get

$$\begin{aligned}
\widetilde{X}_l\left(au\right) &= \widetilde{X}_l P\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + \sum_{i,j=1}^{q} \widetilde{a}_{ij}\left(\xi_0\right)\widetilde{X}_l S_{ij}\left(\xi_0\right) u + \widetilde{X}_l S_0\left(\xi_0\right) u \\
&= S_l\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + \sum_{i,j=1}^{q} \widetilde{a}_{ij}\left(\xi_0\right)\left\{\sum_{k=1}^{q} S_{l,k}^{ij}\left(\xi_0\right)\widetilde{X}_k u + \right. \\
&\left. + \sum_{h,k=1}^{q} \widetilde{a}_{hk}\left(\xi_0\right) S_{l,h}^{\prime ij}\left(\xi_0\right)\widetilde{X}_k u + P_l^{ij}\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + S_l^{ij}\left(\xi_0\right) u\right\} + \\
&+ \sum_{k=1}^{q} S_k^l\left(\xi_0\right)\widetilde{X}_k u + \sum_{h,k=1}^{q} \widetilde{a}_{hk}\left(\xi_0\right) S^{hl}\left(\xi_0\right)\widetilde{X}_k u + S_0^l\left(\xi_0\right) u + P^l\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u \\
&= S_l'\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + \sum_{k=1}^{q} S_{l,k}\left(\xi_0\right)\widetilde{X}_k u + S_l^0\left(\xi_0\right) u \\
&+ \sum_{i,j=1}^{q} \widetilde{a}_{ij}\left(\xi_0\right)\left\{\sum_{k=1}^{q} S_{l,k}^{ij}\left(\xi_0\right)\widetilde{X}_k u + \sum_{h,k=1}^{q} \widetilde{a}_{hk}\left(\xi_0\right) S_{l,h}^{\prime ij}\left(\xi_0\right)\widetilde{X}_k u + \right. \\
&\left. + P_l^{ij}\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + S_l^{ij}\left(\xi_0\right) u\right\}
\end{aligned}$$

where all the $S_{\dots}\left(\xi_0\right)$ are frozen operators of type 1 and $P_l^{ij}\left(\xi_0\right)$ is of type 2.

Next, we perform another derivative $\widetilde{X}_m$ of both sides, getting

$$\begin{aligned}
\widetilde{X}_m\widetilde{X}_l\left(au\right) &= T_{lm}\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + \sum_{k=1}^{q} T_{lm,k}\left(\xi_0\right)\widetilde{X}_k u + T_{lm}^0\left(\xi_0\right) u \\
&+ \sum_{i,j=1}^{q} \widetilde{a}_{ij}\left(\xi_0\right)\left\{\sum_{k=1}^{q} T_{lm,k}^{ij}\left(\xi_0\right)\widetilde{X}_k u + \sum_{h,k=1}^{q} \widetilde{a}_{hk}\left(\xi_0\right) T_{lm,h}^{\prime ij}\left(\xi_0\right)\widetilde{X}_k u + \right. \\
&\left. + S_{lm}^{ij}\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + T_{lm}^{ij}\left(\xi_0\right) u\right\}
\end{aligned}$$

where all the $T_{\dots}\left(\xi_0\right)$ are frozen operators of type 0, and $S_l^{ij}\left(\xi_0\right)$ is of type 1.

This is exactly (4.28). Finally, inserting in this equation the identity

$$\begin{aligned}\widetilde{\mathcal{L}}_0 u(\xi) &= \widetilde{\mathcal{L}} u(\xi) + (\widetilde{\mathcal{L}}_0 - \widetilde{\mathcal{L}}) u(\xi) \\ &= \widetilde{\mathcal{L}} u(\xi) + \sum_{i,j=1}^{q} \left[\widetilde{a}_{ij}(\xi_0) - \widetilde{a}_{ij}(\xi)\right] \widetilde{X}_i \widetilde{X}_j u(\xi)\end{aligned} \tag{4.30}$$

we find (4.29), and the theorem is proved. ■

The above theorem is suited to the proof of $C^{\alpha}_{\widetilde{X}}$ estimates for $\widetilde{X}_i \widetilde{X}_j u$. In order to prove $L^p$ estimate for $\widetilde{X}_i \widetilde{X}_j u$ we need the following variation:

**Theorem 4.21 (Representation of $\widetilde{X}_m \widetilde{X}_l u$ by variable operators)** *Given $a \in C_0^{\infty}\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$, for any $m, l = 1, 2, ..., q$, there exist variable operators over the ball $\widetilde{B}\left(\overline{\xi}, R\right)$, such that for any $u \in C_0^{\infty}\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$*

$$\begin{aligned}\widetilde{X}_m \widetilde{X}_l (au) &= T_{lm} \widetilde{\mathcal{L}} u + \sum_{i,j=1}^{q} \left[\widetilde{a}_{ij}, T_{lm}\right] \widetilde{X}_i \widetilde{X}_j u + \sum_{k=1}^{q} T_{lm,k} \widetilde{X}_k u + T^0_{lm} u + \\ &+ \sum_{i,j=1}^{q} \widetilde{a}_{ij} \left\{ \sum_{k=1}^{q} T^{ij}_{lm,k} \widetilde{X}_k u + \sum_{h,k=1}^{q} \widetilde{a}_{hk} T'^{ij}_{lm,h} \widetilde{X}_k u + S^{ij}_{lm} \widetilde{\mathcal{L}} u + \right. \\ &\left. + \sum_{i,j=1}^{q} \left[\widetilde{a}_{ij}, S^{ij}_{lm}\right] \widetilde{X}_i \widetilde{X}_j u + T^{ij}_{lm} u \right\}.\end{aligned} \tag{4.31}$$

*Here all the $T_{...}$ are variable operators of type 0, $S^{ij}_{lm}$ is of type 1, $[a, T]$ denotes the commutator of the multiplication for $a$ with the operator $T$, and $\widetilde{a}_{ij}$ are the coeffi cients of the operator$\widetilde{\mathcal{L}}$ (which are no longer frozen at $\xi_0$).*

**Remark 4.22** *The above representation formula can be written in a shorter way as*

$$\widetilde{X}_m \widetilde{X}_l (au) = T_{lm} \widetilde{\mathcal{L}} u + \sum_{i,j=1}^{q} \left[\widetilde{a}_{ij}, T_{lm}\right] \widetilde{X}_i \widetilde{X}_j u + \sum_{k=1}^{q} T_{lm,k} \widetilde{X}_k u + T^0_{lm} u$$

*if we leave understood in the symbol of variable operator the possible multiplication by the coeffi cients$\widetilde{a}_{ij}$. See the previous remark.*

**Proof.** Let us rewrite (4.29) as

$$\widetilde{X}_m\widetilde{X}_l\,(au)\,(\xi) = \left(T_{lm}\,(\xi_0)\,\widetilde{\mathcal{L}}u\right)(\xi) + T_{lm}\,(\xi_0)\left(\sum_{i,j=1}^{q}\left[\widetilde{a}_{ij}(\xi_0)-\widetilde{a}_{ij}(\cdot)\right]\,\widetilde{X}_i\widetilde{X}_ju\right)(\xi)+$$
$$+\sum_{k=1}^{q}\left(T_{lm,k}\,(\xi_0)\,\widetilde{X}_ku\right)(\xi) + \left(T^0_{lm}\,(\xi_0)\,u\right)(\xi)+$$
$$+\sum_{i,j=1}^{q}\widetilde{a}_{ij}\,(\xi_0)\left\{\sum_{k=1}^{q}\left(T^{ij}_{lm,k}\,(\xi_0)\,\widetilde{X}_ku\right)(\xi)\right.$$
$$+\sum_{h,k=1}^{q}\widetilde{a}_{hk}\,(\xi_0)\left(T'^{ij}_{lm,h}\,(\xi_0)\,\widetilde{X}_ku\right)(\xi) + \left(S^{ij}_{lm}\,(\xi_0)\,\widetilde{\mathcal{L}}u\right)(\xi)+$$
$$\left.+S^{ij}_{lm}\,(\xi_0)\left(\sum_{i,j=1}^{q}\left[\widetilde{a}_{ij}(\xi_0)-\widetilde{a}_{ij}(\cdot)\right]\,\widetilde{X}_i\widetilde{X}_ju\right)(\xi) + \left(T^{ij}_{lm}\,(\xi_0)\,u\right)(\xi)\right\}.$$

for any $\xi \in \widetilde{B}\left(\overline{\xi}, R\right)$. Letting now $\xi_0 = \xi$ we get

$$\widetilde{X}_m\widetilde{X}_l\,(au)\,(\xi) = \left(T_{lm}\widetilde{\mathcal{L}}u\right)(\xi) + T_{lm}\left(\sum_{i,j=1}^{q}\left[\widetilde{a}_{ij}(\xi)-\widetilde{a}_{ij}(\cdot)\right]\,\widetilde{X}_i\widetilde{X}_ju\right)(\xi)+$$
$$+\sum_{k=1}^{q}\left(T_{lm,k}\widetilde{X}_ku\right)(\xi) + \left(T^0_{lm}u\right)(\xi) + \sum_{i,j=1}^{q}\widetilde{a}_{ij}\,(\xi)\left\{\sum_{k=1}^{q}\left(T^{ij}_{lm,k}\widetilde{X}_ku\right)(\xi)\right.$$
$$+\sum_{h,k=1}^{q}\widetilde{a}_{hk}\,(\xi)\left(T'^{ij}_{lm,h}\widetilde{X}_ku\right)(\xi) + \left(S^{ij}_{lm}\widetilde{\mathcal{L}}u\right)(\xi)+$$
$$\left.+S^{ij}_{lm}\left(\sum_{i,j=1}^{q}\left[\widetilde{a}_{ij}(\xi)-\widetilde{a}_{ij}(\cdot)\right]\,\widetilde{X}_i\widetilde{X}_ju\right)(\xi) + \left(T^{ij}_{lm}u\right)(\xi)\right\}.$$

where all the $T_{\ldots}$ are *variable* operators of type zero over $\widetilde{B}\left(\overline{\xi}, R\right)$, and $S^{ij}_{lm}$ are variable operators of type 1. Note that

$$T\left(\left[\widetilde{a}_{ij}(\xi)-\widetilde{a}_{ij}(\cdot)\right]\,\widetilde{X}_i\widetilde{X}_ju\right)(\xi)$$

is exactly the commutator $[\widetilde{a}_{ij}, T]$ applied to $\widetilde{X}_i\widetilde{X}_ju$. Hence the theorem is proved. ■

# 5 Singular integral estimates for operators of type zero

The proof of *a priori* estimates on the derivatives $\widetilde{X}_i\widetilde{X}_ju$ will follow, as will be explained in detail in § 6.1 and § 7.1, combining the representation formulas proved in § 4.3 with suitable $C^\alpha$ or $L^p$ estimates for "operators of type zero". To be more precise, the results we need are the $C^\alpha_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$ continuity of a *frozen operator of type zero* and the $L^p\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$ continuity of a *variable operator of type zero*, together with the $L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$ estimate for the commutator of a variable operator of type zero with the multiplication with a $VMO$ function, implying that the $L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$ norm of the commutator vanishes as $r \to 0$. All these results will be derived in the present section, as an application of abstract results proved in [8] in the context of locally homogeneous spaces, which have been recalled in § 3.3. To apply them, we need to check that our kernels of type zero satisfy suitable properties. Moreover, to study *variable* operators of type zero, we also have to resort to the classical technique of expansion in series of spherical harmonics, dating back to Calderón-Zygmund [9], and already applied in the framework of vector fields in [2], [3]. This study will be split in two subsection, the first devoted to frozen operators on $C^\alpha$, the second to variable operators on $L^p$.

## 5.1 $C^\alpha$ continuity of frozen operators of type 0

The goal of this section is the proof of the following:

**Theorem 5.1** *Let $\widetilde{B}\left(\overline{\xi}, R\right)$ be as before, $\xi_0 \in \widetilde{B}\left(\overline{\xi}, R\right)$, and let $T(\xi_0)$ be a frozen operator of type $\lambda \geq 0$ over $\widetilde{B}\left(\overline{\xi}, R\right)$. Then there exists $c > 0$ such that for any $r < R$ and $u \in C^\alpha_{\widetilde{X},0}(\widetilde{B}\left(\overline{\xi}, r\right))$,*

$$\|T(\xi_0)u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},r))} \leq c\,\|u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},r))} \tag{5.1}$$

*where $c$ depends on $R, \left\{\widetilde{X}_i\right\}, \alpha$ and $\mu$.*

Recall that $\mu$ is the "ellipticity constant" appearing in Assumption (H) (see § 2).

To prove this, we will apply Theorems 3.11 and 3.14 about the $C^\alpha$ continuity of singular and fractional integrals in spaces of locally homogeneous type, taking

$$\Omega_k = \widetilde{B}\left(\overline{\xi}, \frac{kR}{k+1}\right) \text{ for } k = 1, 2, 3... \tag{5.2}$$

By Definition 4.5, our frozen kernel of type zero can be written as:

$$\begin{aligned} k(\xi_0; \xi, \eta) &= k'(\xi_0; \xi, \eta) + k''(\xi_0; \xi, \eta) \\ &= \left\{\sum_{i=1}^{H} a_i(\xi) b_i(\eta) D_i \Gamma(\xi_0; \cdot) + a_0(\xi) b_0(\eta) D_0 \Gamma(\xi_0; \cdot)\right\}(\Theta(\eta, \xi)) \\ &+ \left\{\sum_{i=1}^{H} a_i'(\xi) b_i'(\eta) D_i' \Gamma^T(\xi_0; \cdot) + a_0'(\xi) b_0'(\eta) D_0' \Gamma^T(\xi_0; \cdot)\right\}(\Theta(\eta, \xi)) \end{aligned}$$

for some positive integer $H$, where $a_i, b_i, a_i', b_i' \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$ $(i = 0, 1, \ldots H)$, $D_i$ and $D_i'$ are differential operators such that: for $i = 1, \ldots, H$ , $D_i$ and $D_i'$ are homogeneous of degree $\leq 2$ (so that $D_i\Gamma(\xi_0; \cdot)$ and $D_i'\Gamma^T(\xi_0; \cdot)$ are homogeneous function of degree $\geq -Q$), $D_0$ and $D_0'$ are differential operators such that $D_0\Gamma(\xi_0; \cdot)$ and $D_0'\Gamma^T(\xi_0; \cdot)$ have bounded first order (Euclidean) derivatives (we will briefly say that $D_0\Gamma(\xi_0; \cdot)$ and $D_0'\Gamma^T(\xi_0; \cdot)$ are regular).

We will prove our Theorem for the operator with kernel $k'$, the proof for $k''$ being completely analogous. Let us split $k'$ as

$$\begin{aligned} k'(\xi_0; \xi, \eta) &= a_1(\xi)b_1(\eta)D_1\Gamma(\xi_0; \Theta(\eta, \xi)) \\ &+ \left\{ \sum_{i=2}^{H_m} a_i(\xi)b_i(\eta)D_i\Gamma(\xi_0; \cdot) + a_0(\xi)b_0(\eta)D_0\Gamma(\xi_0; \cdot) \right\} (\Theta(\eta, \xi)) \\ &\equiv k_S(\xi, \eta) + k_F(\xi, \eta) \end{aligned}$$

where $D_1\Gamma(\xi_0; u)$ is homogeneous of degree $-Q$ while all the kernels $D_i\Gamma(\xi_0; u)$ are homogeneous of some degree $\geq 1 - Q$ and $D_0\Gamma(\xi_0; u)$ is regular. Recall that all these kernels may have also an explicit (smooth) dependence on $\xi, \eta$; we will write $D_i^{\xi,\eta}\Gamma(\xi_0; \Theta(\eta, \xi))$ to point out this fact, when it will be important.

We want to apply Theorem 3.11 (about singular integrals) to the kernel $k_S$ and Theorem 3.14 (about fractional integrals) to each term of the kernel $k_F$.

We start with the following result, very similar to [3, Proposition 2.17]:

**Proposition 5.2** *Let $W^{\xi,\eta}(\cdot)$ be a function defined on the homogeneous group $\mathbb{G}$, smooth outside the origin and homogeneous of degree $\ell - Q$ for some nonnegative integer $\ell$, smoothly depending on the parameters $\xi, \eta \in \widetilde{B}\left(\overline{\xi}, R\right)$, and let*

$$K(\xi, \eta) = W^{\xi,\eta}(\Theta(\eta, \xi))$$

*be defined for $\xi, \eta \in \widetilde{B}\left(\overline{\xi}, R\right)$. Then $K$ satisfies*

*(i) the growth condition: there exists a constant $c$ such that*

$$|K(\xi, \eta)| \leq c \cdot \sup_{\|u\|=1} \left|W^{\xi,\eta}(u)\right| \cdot d_{\widetilde{X}}(\xi, \eta)^{\ell - Q};$$

*(ii) the mean value inequality: there exists a constant $c > 0$, such that for every $\xi_0, \xi, \eta$ with $d_{\widetilde{X}}(\xi_0, \eta) \geq 2d_{\widetilde{X}}(\xi_0, \xi)$,*

$$|K(\xi_0, \eta) - K(\xi, \eta)| + |K(\eta, \xi_0) - K(\eta, \xi)| \leq C \frac{d_{\widetilde{X}}(\xi_0, \xi)}{d_{\widetilde{X}}(\xi_0, \eta)^{Q+1-\ell}} \tag{5.3}$$

*where the constant $C$ has the form*

$$c \cdot \sup_{\|u\|=1, \xi, \eta \in \widetilde{B}\left(\overline{\xi}, R\right)} \left\{ \left|\nabla_u W^{\xi,\eta}(u)\right| + \left|\nabla_\xi W^{\xi,\eta}(u)\right| + \left|\nabla_\eta W^{\xi,\eta}(u)\right| \right\}$$

*(iii) the cancellation property: if $\ell = 0$ and $W$ satisfies the vanishing property*

$$\int_{r<\|u\|<R} W^{\xi,\eta}(u)du = 0 \text{ for every } R > r > 0, \text{ any } \xi, \eta \in \widetilde{B}\left(\overline{\xi}, R\right) \tag{5.4}$$

*then for any positive integer $k$, for every $\varepsilon_2 > \varepsilon_1 > 0$ and $\xi \in \Omega_k$ (see (5.2)) such that $\widetilde{B}(\xi, \varepsilon_2) \subset \Omega_{k+1}$*

$$\left| \int_{\Omega_{k+1}, \varepsilon_1 < \rho(\xi,\eta) < \varepsilon_2} K(\xi, \eta)\, d\eta \right| + \left| \int_{\Omega_{k+1}, \varepsilon_1 < \rho(\xi,\eta) < \varepsilon_2} K(\eta, \xi)\, d\eta \right| \leq C \cdot (\varepsilon_2 - \varepsilon_1) \tag{5.5}$$

*where the constant $C$ has the form*

$$c \cdot \sup_{\|u\|=1, \xi, \eta \in \widetilde{B}(\overline{\xi}, R)} \left\{ \left| W^{\xi,\eta}(u) \right| + \left| \nabla_\xi W^{\xi,\eta}(u) \right| + \left| \nabla_\eta W^{\xi,\eta}(u) \right| \right\}.$$

**Proof.** Point (i) is trivial, by the homogeneity of $W$, and the equivalence between $d_{\widetilde{X}}$ and $\rho$ (see Lemma 3.9).

In order to prove (ii), fix $\xi_0$, $\eta$, and let $r = \frac{1}{2}\rho(\eta, \xi_0)$. Condition $\rho(\eta, \xi_0) > 2\rho(\xi, \xi_0)$ means that $\xi$ is a point ranging in $\widetilde{B}_r(\xi_0)$. Applying (3.28) to the function

$$f(\xi) = K(\xi, \eta)$$

we can write

$$|f(\xi) - f(\xi_0)| \leq c d_{\widetilde{X}}(\xi, \xi_0) \cdot$$
$$\cdot \left( \sum_{i=1}^{q} \sup_{\zeta \in \widetilde{B}\left(\xi_0, \frac{3}{4} d_{\widetilde{X}}(\xi_0, \eta)\right)} \left| \widetilde{X}_i f(\zeta) \right| + d_{\widetilde{X}}(\xi, \xi_0) \sup_{\zeta \in \widetilde{B}\left(\xi_0, \frac{3}{4} d_{\widetilde{X}}(\xi_0, \eta)\right)} \left| \widetilde{X}_0 f(\zeta) \right| \right).$$

Noting that, for $\zeta \in \widetilde{B}\left(\xi_0, \frac{3}{4} d_{\widetilde{X}}(\xi_0, \eta)\right)$,

$$\begin{aligned} \left| \widetilde{X}_i K(\cdot, \eta)(\zeta) \right| &= \left| \widetilde{X}_i \left( W^{\zeta,\eta}(\Theta(\cdot, \eta)) \right)(\zeta) + \left( \widetilde{X}_i W^{\cdot,\eta}(\Theta(\zeta, \eta)) \right)(\zeta) \right| \\ &\leq \left| (Y_i W + R_i^\eta W)(\Theta(\eta, \zeta)) \right| + \left| \left( \widetilde{X}_i W^{\cdot,\eta}(\Theta(\zeta, \eta)) \right)(\zeta) \right| \end{aligned}$$

and recalling that, by Remark 4.6, $\nabla_\zeta W^{\zeta,\eta}(u)$ has the same $u$ homogeneity as $W^{\zeta,\eta}(u)$, we get

$$\begin{aligned} \left| \widetilde{X}_i K(\cdot, \eta)(\zeta) \right| &\leq \sup_{\|u\|=1, \zeta, \eta \in \widetilde{B}(\overline{\xi}, R)} \left| \nabla_u W^{\xi,\eta}(u) \right| \frac{c}{\rho(\zeta, \eta)^{Q-\ell+1}} + \\ &\quad + \sup_{\|u\|=1, \zeta, \eta \in \widetilde{B}(\overline{\xi}, R)} \left| \nabla_\zeta W^{\zeta,\eta}(u) \right| \frac{c}{\rho(\zeta, \eta)^{Q-\ell}} \\ &\leq \sup_{\|u\|=1, \zeta, \eta \in \widetilde{B}(\overline{\xi}, R)} \left\{ \left| \nabla_u W^{\zeta,\eta}(u) \right| + \left| \nabla_\zeta W^{\zeta,\eta}(u) \right| \right\} \frac{c}{d_{\widetilde{X}}(\xi_0, \eta)^{Q-\ell+1}}. \end{aligned}$$

Analogously

$$\left| \widetilde{X}_0 K(\cdot, \eta)(\zeta) \right| \leq \sup_{\|u\|=1, \zeta, \eta \in \widetilde{B}(\overline{\xi}, R)} \left\{ \left| \nabla_u W^{\zeta,\eta}(u) \right| + \left| \nabla_\zeta W^{\zeta,\eta}(u) \right| \right\} \frac{c}{d_{\widetilde{X}}(\xi_0, \eta)^{Q-\ell+2}},$$

hence

$$\begin{aligned}|K(\xi,\eta)-K(\xi_0,\eta)| &\leq C d_{\widetilde{X}}(\xi,\xi_0)\left(\frac{1}{d_{\widetilde{X}}(\xi_0,\eta)^{Q-\ell+1}}+\frac{d_{\widetilde{X}}(\xi,\xi_0)}{d_{\widetilde{X}}(\xi_0,\eta)^{Q-\ell+2}}\right)\\ &\leq C\frac{d_{\widetilde{X}}(\xi,\xi_0)}{d_{\widetilde{X}}(\xi_0,\eta)^{Q-\ell+1}}\end{aligned}$$

with

$$C=c\sup_{\|u\|=1,\zeta,\eta\in\widetilde{B}(\overline{\xi},R)}\left\{\left|\nabla_u W^{\zeta,\eta}(u)\right|+\left|\nabla_\zeta W^{\zeta,\eta}(u)\right|\right\}.$$

To get the analogous bound for $|K(\eta,\xi_0)-K(\eta,\xi)|$, it is enough to apply the previous estimate to the function

$$\widetilde{K}(\xi,\eta)=\widetilde{W}^{\xi,\eta}(\Theta(\eta,\xi)) \text{ with } \widetilde{W}^{\xi,\eta}(u)=W^{\eta,\xi}(u^{-1}).$$

This completes the proof of (ii).

To prove (iii), we first ignore the dependence on the parameters $\xi,\eta$, and then we will show how to modify our argument to keep it into account. Let us write:

$$\int_{\Omega_{k+1},\varepsilon_1<\rho(\xi,\eta)<\varepsilon_2} W(\Theta(\eta,\xi))\,d\eta=$$

by the change of variables $u=\Theta(\eta,\xi)$, Theorem 3.3 (b) gives

$$=c(\xi)\int_{\varepsilon_1<\|u\|<\varepsilon_2} W(u)\,(1+\omega(\xi,u))\,du=$$

by the vanishing property of $W$,

$$=c(\xi)\int_{\varepsilon_1<\|u\|<\varepsilon_2} W(u)\omega(\xi,u)\,du.$$

Then

$$\begin{aligned}\left|\int_{\Omega_{k+1},\varepsilon_1<\rho(\xi,\eta)<\varepsilon_2} W(\Theta(\eta,\xi))\,d\eta\right| &\leq c\cdot\int_{\varepsilon_1<\|u\|<\varepsilon_2}|W(u)|\,\|u\|\,du\\ &\leq c\cdot\sup_{\|u\|=1}|W|\cdot\int_{\varepsilon_1<\|u\|<\varepsilon_2}\|u\|^{1-Q}\,du\\ &\leq c\cdot\sup_{\|u\|=1}|W|\cdot(\varepsilon_2-\varepsilon_1).\end{aligned}$$

Analogously one proves the bound on $W(\Theta(\xi,\eta))$. Now, to keep track of the possible dependence of $W$ on the parameters $\xi,\eta$, let us write:

$$\begin{aligned}&\int_{\Omega_{k+1},\varepsilon_1<\rho(\xi,\eta)<\varepsilon_2} W^{\xi,\eta}(\Theta(\eta,\xi))\,d\eta=\int_{\Omega_{k+1},\varepsilon_1<\rho(\xi,\eta)<\varepsilon_2} W^{\xi,\xi}(\Theta(\eta,\xi))\,d\eta+\\ &+\int_{\Omega_{k+1},\varepsilon_1<\rho(\xi,\eta)<\varepsilon_2}\left[W^{\xi,\eta}(\Theta(\eta,\xi))-W^{\xi,\xi}(\Theta(\eta,\xi))\right]d\eta\\ &\equiv I+II.\end{aligned}$$

The term $I$ can be bounded as above, while

$$\left|W^{\xi,\eta}(u)-W^{\xi,\xi}(u)\right|\leq|\xi-\eta|\left|\nabla_{\eta}W^{\xi,\eta'}(u)\right|$$

for some point $\eta'$ near $\xi$ and $\eta$. Recalling again that the function $\nabla_{\eta}W^{\xi,\eta'}(\cdot)$ has the same homogeneity as $W^{\xi,\eta'}(\cdot)$, while

$$|\xi-\eta|\leq cd_{\widetilde{X}}(\xi,\eta)\leq c\rho(\xi,\eta),$$

we have

$$|II|\leq c\sup_{\|u\|=1,\xi,\eta\in\widetilde{B}\left(\overline{\xi},R\right)}\left|\nabla_{\eta}W^{\xi,\eta}(u)\right|\int_{\Omega_{k+1},\varepsilon_1<\|u\|<\varepsilon_2}\|u\|^{1-Q}\,du$$

and the same reasoning as above applies. This proves the bound on $\left|\int K(\xi,\eta)\,d\eta\right|$ in (5.5). The proof of the bound on $\left|\int K(\eta,\xi)\,d\eta\right|$ is analogous, since the vanishing property (5.4) also implies the same for $\int_{r<\|u\|<R}W^{\xi,\eta}(u^{-1})du$. ■

The above Proposition implies that $D_1\Gamma(\xi_0;\Theta(\eta,\xi))$ satisfies the standard estimates, cancellation property and convergence condition stated in § 3.3. Note that (5.5) implies both the cancellation property and the convergence condition.

In order to apply to the kernel $k_S(\xi,\eta)$ Thm. 3.11 we still need to prove that the singular integral $T$ with kernel $k_S(\xi,\eta)$ satisfies the condition $T(1)\in C^{\gamma}_{\widetilde{X}}$. (see condition (3.18) in Theorem 3.11).

This result is more delicate than the previous condition, and is contained in the following:

**Proposition 5.3** *Let*

$$\widetilde{h}(\xi)\equiv\lim_{\varepsilon\to0}\int_{\rho(\xi,\eta)>\varepsilon}\widetilde{K}(\xi,\eta)d\eta$$

*with*

$$\widetilde{K}(\xi,\eta)=a_1(\xi)b_1(\eta)D_1^{\xi,\eta}\Gamma(\xi_0;\Theta(\eta,\xi)),$$

$D_1^{\xi,\eta}\Gamma(\xi_0;\cdot)$ *homogeneous of degree* $-Q$ *and satisfying the vanishing property*

$$\int_{r<\|u\|<R}D_1^{\xi,\eta}\Gamma(\xi_0;u)du=0\text{ for every }R>r>0,\text{ any }\xi,\eta\in\widetilde{B}\left(\overline{\xi},R\right).$$

*Then* $\widetilde{h}\in C^{\gamma}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi},R\right)\right)$ *for any* $\gamma\in(0,1)$.

**Proof.** Since $a_1,b_1$ are compactly supported in $\widetilde{B}\left(\overline{\xi},R\right)$, we can choose a radial cutoff function

$$\phi(\xi,\eta)=f(\rho(\xi,\eta)),$$

with

$$f(\|u\|)=1\text{ for }\|u\|\leq R,\ f(\|u\|)=0\text{ for }\|u\|\geq2R,$$

so that $\widetilde{K}(\xi,\eta) = \widetilde{K}(\xi,\eta)\phi(\xi,\eta)$. To begin with, let us prove the assertion without taking into consideration the dependence of $D_1^{\xi,\eta}\Gamma(\xi_0;u)$ on $\xi,\eta$. Then

$$\begin{aligned}\widetilde{h}(\xi) &= a_1(\xi)b_1(\xi)\lim_{\varepsilon\to 0}\int_{\rho(\xi,\eta)>\varepsilon}\phi(\xi,\eta)\,D_1\Gamma(\xi_0;\Theta(\eta,\xi))d\eta+\\ &+a_1(\xi)\int\phi(\xi,\eta)\,D_1\Gamma(\xi_0;\Theta(\eta,\xi))\,[b_1(\eta)-b_1(\xi)]\,d\eta\\ &\equiv I(\xi)+II(\xi).\end{aligned}$$

Now,

$$\begin{aligned}I(\xi) &= a_1(\xi)b_1(\xi)\,c(\xi)\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon}f(\|u\|)\,D_1\Gamma(\xi_0;u)\,(1+\omega(\xi,u))\,du\\ &= a_1(\xi)b_1(\xi)\,c(\xi)\int f(\|u\|)\,D_1\Gamma(\xi_0;u)\omega(\xi,u)\,du,\end{aligned}$$

by the vanishing property, with $\omega$ smoothly depending on $\xi$ and uniformly bounded by $c\|u\|$. Hence $I(\xi)$ is Lipschitz continuous, in particular Hölder continuous of any exponent $\gamma\in(0,1)$. Moreover,

$$\begin{aligned}II(\xi) &= a_1(\xi)\int_{\widetilde{B}(\overline{\xi},R)}\kappa(\xi,\eta)\,d\eta\text{ with}\\ \kappa(\xi,\eta) &= \phi(\xi,\eta)\,D_1\Gamma(\xi_0;\Theta(\eta,\xi))\,[b_1(\eta)-b_1(\xi)].\end{aligned}$$

It is not diffi cult to check that the kernel$\kappa(\xi,\eta)$ satisfies the standard estimates of *fractional* integrals (3.14) and (3.15) in § 3.3 for any $\nu\in(0,1)$ (actually, for $\nu=1$). Hence the operator with kernel $\kappa$ is continuous on $C^\gamma\left(\widetilde{B}(\overline{\xi},R)\right)$; in particular, it maps the function 1 into $C^\gamma\left(\widetilde{B}(\overline{\xi},R)\right)$, which proves that $II(\xi)$ is Hölder continuous.

To conclude the proof, we have to show how to take into account the possible dependence of $D_1^{\xi,\eta}\Gamma(\xi_0;u)$ on $\xi,\eta$. Let us start with the $\eta$ dependence.

$$\begin{aligned}\widetilde{h}(\xi) &= a_1(\xi)b_1(\xi)\lim_{\varepsilon\to 0}\int_{\rho(\xi,\eta)>\varepsilon}\phi(\xi,\eta)\,D_1^\eta\Gamma(\xi_0;\Theta(\eta,\xi))d\eta+\\ &+a_1(\xi)\int\phi(\xi,\eta)\,D_1^\eta\Gamma(\xi_0;\Theta(\eta,\xi))\,[b_1(\eta)-b_1(\xi)]\,d\eta\\ &\equiv I'(\xi)+II'(\xi).\end{aligned}$$

The term $II'(\xi)$ can be handled as the term $II(\xi)$ above. As to $I'(\xi)$,

$$\begin{aligned}I'(\xi) &= a_1(\xi)b_1(\xi)\,c(\xi)\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon}f(\|u\|)\,D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0;u)\,(1+\omega(\xi,u))\,du\\ &= a_1(\xi)b_1(\xi)\,c(\xi)\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon}f(\|u\|)\,D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0;u)du+\\ &+a_1(\xi)b_1(\xi)\,c(\xi)\int f(\|u\|)\,D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0;u)\omega(\xi,u)\,du.\end{aligned}$$

The second term can be handled as above, while the first requires some care. By the vanishing property of the kernel $D_1^{\zeta}\Gamma(\xi_0;u)$ for any fixed $\zeta$ we can write

$$\begin{aligned}&\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon} f\left(\|u\|\right) D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0;u)du\\&=\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon} f\left(\|u\|\right)\left[D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0;u)-D_1^{\xi}\Gamma(\xi_0;u)\right]du.\end{aligned}$$

On the other hand,

$$D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0;u)=D_1^{\xi}\Gamma(\xi_0;u)+D_0^{\xi}\Gamma(\xi_0;u)$$

where $D_0^{\xi}$ is a vector field of local weight $\leq 0$, smoothly depending on $\xi$. Hence

$$\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon} f\left(\|u\|\right) D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0;u)du=\int f\left(\|u\|\right) D_0^{\xi}\Gamma(\xi_0;u)du,$$

which can be handled as the term $I\left(\xi\right)$ above.

Dependence on the variable $\xi$ can be taken into account as follows. If

$$\begin{aligned}\widetilde{h}\left(\xi\right)&=a_1(\xi)b_1\left(\xi\right)\lim_{\varepsilon\to 0}\int_{\rho(\xi,\eta)>\varepsilon}\phi\left(\xi,\eta\right) D_1^{\xi,\eta}\Gamma(\xi_0;\Theta(\eta,\xi))d\eta\\&\equiv\lim_{\varepsilon\to 0}\int F_{\varepsilon}\left(\xi,\xi,\eta\right)\text{ with}\\F_{\varepsilon}\left(\zeta,\xi,\eta\right)&=a_1(\xi)b_1\left(\xi\right)\chi_{\rho(\xi,\eta)>\varepsilon}\left(\eta\right)\phi\left(\xi,\eta\right) D_1^{\zeta,\eta}\Gamma(\xi_0;\Theta(\eta,\xi))d\eta\end{aligned}$$

then

$$\begin{aligned}\widetilde{h}\left(\xi_1\right)-\widetilde{h}\left(\xi_2\right)&=\lim_{\varepsilon\to 0}\int\left[F_{\varepsilon}\left(\xi_1,\xi_1,\eta\right)-F_{\varepsilon}\left(\xi_2,\xi_1,\eta\right)\right]d\eta+\\&+\lim_{\varepsilon\to 0}\int\left[F_{\varepsilon}\left(\xi_2,\xi_1,\eta\right)-F_{\varepsilon}\left(\xi_2,\xi_2,\eta\right)\right]d\eta\\&\equiv A\left(\xi_1,\xi_2\right)+B\left(\xi_1,\xi_2\right).\end{aligned}$$

Now,

$$\left|A\left(\xi_1,\xi_2\right)\right|\leq c\rho\left(\xi_1,\xi_2\right)$$

by the smoothness of $\xi\mapsto D_1^{\xi,\eta}\Gamma(\xi_0;u)$. As to $B\left(\xi_1,\xi_2\right)$, it is enough to apply the previous reasoning to the kernel $D_1^{\zeta,\eta}\Gamma(\xi_0;\Theta(\eta,\xi))$, for any fixed $\zeta$, to conclude that

$$\left|\lim_{\varepsilon\to 0}\int\left[F\left(\zeta,\xi_1,\eta\right)-F\left(\zeta,\xi_2,\eta\right)\right]d\eta\right|\leq c\rho\left(\xi_1,\xi_2\right)^{\gamma}$$

for some constant uniformly bounded in $\zeta$, and then apply this inequality taking $\zeta=\xi_2$. This completes the proof. ■

We are now ready for the

**Conclusion of the proof of Theorem 5.1.** Recall that a frozen operator of type zero is written as

$$T(\xi_0)f(\xi) = P.V. \int_{\widetilde{B}} k(\xi_0;\xi,\eta)\, f(\eta)\, d\eta + \alpha\,(\xi_0,\xi)\, f\,(\xi)\,,$$

where $\alpha$ is a bounded measurable function, smooth in $\xi$. The multiplicative part

$$f\,(\xi) \longmapsto \alpha\,(\xi_0,\xi)\, f\,(\xi)$$

clearly maps $C^\alpha$ in $C^\alpha$, since $\alpha\,(\xi_0,\cdot)$ is smooth, with operator norm bounded by some constant depending on the vector fields and the ellipticity constant $\mu$, by Theorem 4.3.

Let us now consider the integral part. With the notation introduced at the beginning of this section, let us consider first

$$k_S\,(\xi,\eta) = a_1(\xi)b_1(\eta)D_1^{\xi,\eta}\Gamma(\xi_0;\Theta(\eta,\xi)),$$

with $D_1^{\xi,\eta}\Gamma\,(\xi_0;u)$ homogeneous of degree $-Q$ and satisfying the vanishing property (5.4). By Proposition 5.2, $k_S\,(\xi,\eta)$ satisfies conditions (i), (ii), (iii) in § 3.3, with constants bounded by

$$c \sup_{\|u\|=1} \left\{ \left| D^2\Gamma\,(\xi_0,u)\right| + \left| D^3\Gamma\,(\xi_0,u)\right| \right\} \tag{5.6}$$

where the symbols $D^2, D^3$ denote standard derivatives of orders $2,3$, respectively, with respect to $u$, and the constant $c$ depends on the vector fields but not on the point $\xi_0$. By Proposition 5.3, condition (3.18) is also satisfied by $k_S\,(\xi,\eta)\,,$ with $C^\gamma$ norm bounded by a quantity of the kind (5.6). Hence by Theorem 3.11 and Remark 3.12, the operator with kernel $k_S\,(\xi,\eta)$ satisfies the assertion of the theorem we are proving, with a constant bounded by a quantity like (5.6). In turn, by Theorem 4.3 this quantity can be bounded by a constant depending on the vector fields and the ellipticity constant $\mu$ of the matrix $a_{ij}\,(x)\,.$

Let us now come to the kernel

$$k_F\,(\xi,\eta) = \left\{ \sum_{i=2}^{H} a_i(\xi)b_i(\eta)D_i^{\xi,\eta}\Gamma\,(\xi_0;\cdot) + a_0(\xi)b_0(\eta)D_0^{\xi,\eta}\Gamma\,(\xi_0;\cdot) \right\} (\Theta(\eta,\xi))$$

where each function $D_i^{\xi,\eta}\Gamma\,(\xi_0;u)$ $(i = 2,3,...,H)$ is homogeneous of some degree $\geq 1-Q$, while $D_0^{\xi,\eta}\Gamma\,(\xi_0;u)$ is bounded and smooth.

By Proposition 5.2, each kernel

$$a_i(\xi)b_i(\eta)D_i^{\xi,\eta}\Gamma\,(\xi_0;\Theta(\eta,\xi))$$

satisfies the standard estimates (i) in § 3.3 for some $\nu > 0$, hence we can apply Theorem 3.14 to the integral operators defined by these kernels, and conclude as above. Finally, the integral operator with regular kernel clearly is $C^\gamma$ continuous. So we are done. ■

## 5.2 $L^p$ continuity of variable operators of type $0$ and their commutators

In this subsection we are going to prove the following:

**Theorem 5.4** *Let $T$ be a variable operator of type $0$ (see § 4.2) over the ball $\widetilde{B}\left(\overline{\xi}, R\right)$, and $p \in (1, \infty)$. Then:*

*(i) there exists $c > 0$, depending on $p$, $R$, $\left\{\widetilde{X}_i\right\}_{i=0}^{q}$, $\mu$, such that*

$$\|Tu\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq c\,\|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}$$

*for every $u \in L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$ and $r \leq R$;*

*(ii) for every $a \in VMO_{X,loc}(\Omega)$, any $\varepsilon > 0$, there exists $r \leq R$ such that for every $u \in L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$,*

$$\|T(\widetilde{a}u) - \widetilde{a}\cdot Tu\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq \varepsilon\,\|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \tag{5.7}$$

*where $\widetilde{a}(x,h) = a(x)$. The number $r$ depends on $p$, $R$, $\left\{\widetilde{X}_i\right\}_{i=0}^{q}$, $\mu$, $\eta^*_{a,\Omega',\Omega}$, and $\varepsilon$ (see 3.4.3 for the definition of $VMO_{X,loc}(\Omega)$ and $\eta^*_{a,\Omega',\Omega}$).*

A basic difference with the context of the previous section is that here we are considering *variable* kernels and operators of type zero. To reduce the study of these operators to that of constant operators of type zero we will make use of the classical technique of expansion in series of spherical harmonics, as already done in [3].

**Proof.** This proof is similar to that of [3, Thm. 2.11]. Recall that a variable operator of type zero is written as

$$Tf(\xi) = P.V.\int_{\widetilde{B}} k(\xi;\xi,\eta)\,f(\eta)\,d\eta + \alpha(\xi,\xi)\,f(\xi),$$

where $\alpha(\xi_0,\xi)$ is a bounded measurable function in $\xi_0$, smooth in $\xi$. The multiplicative part

$$f(\xi) \longmapsto \alpha(\xi,\xi)\,f(\xi)$$

clearly maps $L^p$ into $L^p$, with operator norm bounded by some constant depending on the vector fields and the ellipticity constant $\mu$, by Theorem 4.3. Moreover, this part does not affect the commutator of $T$.

As to the integral part of $T$, let us split the variable kernel as

$$k(\xi;\xi,\eta) = k'(\xi;\xi,\eta) + k''(\xi;\xi,\eta).$$

Like in the previous section, it is enough to prove our result for the kernel $k'$. Let us expand it as

$$\begin{aligned} k'(\xi;\xi,\eta) &= \sum_{i=1}^{H} a_i(\xi)b_i(\eta)D_i^{\xi,\eta}\Gamma\left(\xi;\Theta(\eta,\xi)\right) + a_0(\xi)b_0(\eta)D_0^{\xi,\eta}\Gamma(\xi;\Theta(\eta,\xi)) \\ &\equiv k_S(\xi;\xi,\eta) + k_B(\xi;\xi,\eta) \end{aligned}$$

where the kernels $D_i^{\xi,\eta}\Gamma(\xi;u)$ (for $i=1,2,3,...,H$) are homogeneous of some degree $\geq -Q$, $D_i^{\xi,\eta}\Gamma(\xi;u)$ satisfies the cancellation property, and $D_0^{\xi,\eta}\Gamma(\xi;u)$ is bounded in $u$ and smooth in $\xi,\eta$. The kernels $k_S, k_B$ are "singular" and "bounded", respectively.

The operator with kernel $k_B$ is obviously $L^p$ continuous. Moreover, it satisfies the commutator estimate (5.7) by Theorem 3.18, since

$$|k_B(\xi;\xi,\eta)| \leq c a_0(\xi) b_0(\eta)$$

and the constant function 1 obviously satisfies the standard estimates (3.14), (3.15) with $\nu = 1$.

To handle the kernel $k_S$ we expand each of its terms in series of spherical harmonics, exactly like in [3, Section 2.4]:

$$D_i^{\xi,\eta}\Gamma(\xi;u) = \sum_{m=0}^{\infty}\sum_{k=1}^{g_m} c_{i,km}^{\xi,\eta}(\xi)\, K_{i,km}(u)$$

where $K_{i,km}(u)$ are homogeneous kernels which on the sphere $\|u\|=1$ coincide with the spherical harmonics, and $c_{i,km}^{\xi,\eta}(\cdot)$ the corresponding Fourier coefficients.

Let us first prove the assertion without taking into account the dependence of the coefficients $c_{i,km}^{\xi,\eta}(\xi)$ on $\eta$. Then the operator with kernel $k_S$ can be written as:

$$Sf(\xi) = \sum_{m=0}^{\infty}\sum_{k=1}^{g_m} c_{i,km}^{\xi}(\xi)\, S_{i,km}f(\xi) \tag{5.8}$$

with

$$S_{i,km}f(\xi) = a_i(\xi)\int_{\widetilde{B}} b_i(\eta) K_{i,km}(\Theta(\eta,\xi))\, f(\eta)\, d\eta.$$

The number $g_m$ in (5.8) is the dimension of the space of spherical harmonics of degree $m$ in $\mathbb{R}^N$; it is known that

$$g_m \leq c(N)\cdot m^{N-2} \quad \text{for every } m = 1,2,\ldots \tag{5.9}$$

For every $p\in(1,\infty)$ we can write:

$$\|Sf\|_{L^p(\widetilde{B}(\overline{\xi},r))} \leq \sum_{m=0}^{\infty}\sum_{k=1}^{g_m} \left\|c_{i,km}^{\cdot}(\cdot)\right\|_{L^\infty(\widetilde{B}(\overline{\xi},r))} \|S_{i,km}f\|_{L^p(\widetilde{B}(\overline{\xi},r))}$$

and

$$\|S(\widetilde{a}f) - \widetilde{a}\cdot Sf\|_{L^p(\widetilde{B}(\overline{\xi},r))} \leq$$
$$\leq \sum_{m=0}^{\infty}\sum_{k=1}^{g_m} \left\|c_{i,km}^{\cdot}(\cdot)\right\|_{L^\infty(\widetilde{B}(\overline{\xi},r))} \|S_{i,km}(\widetilde{a}f) - \widetilde{a}\cdot S_{i,km}f\|_{L^p(\widetilde{B}(\overline{\xi},r))}.$$

Now each $S_{i,km}$ is a frozen operator of type $\lambda \geq 0$, and the same arguments of the previous section show that the kernel of $S_{i,km}$ satisfies the assumptions (i),(ii),(iii) in § 3.3 with constants bounded by

$$c \cdot \sup_{\|u\|=1} |\nabla_u K_{km}(u)|,$$

(with $c$ depending on the vector fields); in turn, by known properties of spherical harmonics we have

$$\sup_{\|u\|=1} |\nabla_u K_{km}(u)| \leq c\,(N)\, m^{N/2},$$

so that, by Theorem 3.11 and Theorem 3.13 we conclude as in [3, p. 807],

$$\|S_{i,km} f\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq c \cdot m^{N/2} \|f\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \text{ for } i = 1, 2, ..., H.$$

where we have also taken into account Remark 3.19.

Analogously, applying Theorem 3.16 and Theorem 3.17 we have the commutator estimate:

$$\|S_{i,km}(\widetilde{a}f) - \widetilde{a} \cdot S_{i,km} f\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq \varepsilon \cdot m^{N/2} \|f\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \text{ for } i = 1, 2, ..., H,$$

for any $\varepsilon > 0$, provided $r$ is small enough, depending on $\varepsilon$ and $\eta^*_{\widetilde{a},\Omega_{k+2},\Omega_{k+3}}$ (see (5.2) and Definition 3.15 for the meaning of symbols). By Proposition 3.35, then, the constant $r$ depends on the function $a$ only through the local VMO modulus $\eta^*_{a,\Omega',\Omega}$.

Next, again by known properties of spherical harmonics, we can say that for any positive integer $h$ there exists $c_h$ such that

$$\left|c^{\zeta}_{i,km}(\xi)\right| \leq c_h \cdot m^{-2h} \sup_{\|u\|=1, |\beta|=2h} \left| \left(\frac{\partial}{\partial u}\right)^{\beta} D^{\zeta}_i \Gamma(\xi; u)\right|.$$

By the uniform estimates contained in Theorem 4.3, the last expression is bounded by $Cm^{-2h}$, for some constant $C$ depending on $h$, the vector fields, and the ellipticity constant $\mu$. Taking into account also (5.9) and choosing $h$ large enough we conclude

$$\|Sf\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq \sum_{m=0}^{\infty} C g_m m^{-2h} m^{N/2} \|f\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} = c \|f\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}$$

and

$$\|S(\widetilde{a}f) - \widetilde{a} \cdot Sf\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq c\varepsilon \|f\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}$$

for any $\varepsilon > 0$, provided $r$ is small enough.

We are left to show how the previous argument needs to be modified to take into account the possible dependence of $D^{\xi,\eta}_i \Gamma(\xi; u)$ (and then of $c^{\xi,\eta}_{i,km}(\xi)$) on $\eta$. Let us expand:

$$D^{\zeta,\Theta(\cdot,\zeta)^{-1}(u)}_i \Gamma(\xi; u) = \sum_{m=0}^{\infty} \sum_{k=1}^{g_m} c^{\zeta}_{i,km}(\xi)\, K_{i,km}(u)$$

so that

$$D_i^{\zeta,\eta}\Gamma(\xi;\Theta(\eta,\zeta))=\sum_{m=0}^{\infty}\sum_{k=1}^{g_m}c_{i,km}^{\zeta}(\xi)\,K_{i,km}(\Theta(\eta,\zeta))\,.$$

The kernels $K_{i,km}$ are the same as above, hence the estimates on the operators $S_{i,km}$ and their commutators remain unchanged. As to the coeffi cients $c_{i,km}^{\zeta}(\xi)$, we now have to write, for any positive integer $h$ and some constant $c_h$,

$$\left|c_{i,km}^{\zeta}(\xi)\right|\leq c_h\cdot m^{-2h}\sup_{\|u\|=1,|\beta|=2h}\left|\left(\frac{\partial}{\partial u}\right)^{\beta}\left(D_i^{\zeta,\Theta(\cdot,\zeta)^{-1}(u)}\Gamma(\xi;u)\right)\right|.$$

Now, from the identity

$$\frac{\partial}{\partial u_j}\left(D_i^{\zeta,\Theta(\cdot,\zeta)^{-1}(u)}\Gamma(\xi;u)\right)=\frac{\partial}{\partial u_j}\left(D_i^{\zeta,\eta}\Gamma(\xi;u)\right)_{/\eta=\Theta(\cdot,\zeta)^{-1}(u)}+$$
$$+\sum_m\frac{\partial}{\partial\eta_m}\left(D_i^{\zeta,\eta}\Gamma(\xi;u)\right)\frac{\partial}{\partial u_j}\left(\Theta(\cdot,\zeta)^{-1}(u)\right)_m$$

it is easy to see that we can still get a uniform bound of the kind

$$\left|c_{i,km}^{\zeta}(\xi)\right|\leq C\cdot m^{-2h}$$

with $C$ depending on $h$, the vector fields and the ellipticity constant $\mu$. So we are done. ■

# 6 Schauder estimates

We are now in position to apply all the machinery presented in the previous sections to prove our main results, that is $C^{\alpha}$ and $L^p$ estimates on $X_iX_ju$ in terms of $u$ and $\mathcal{L}u$. We will prove $C^{\alpha}$ estimates (that is Theorem 2.1) in this section, and $L^p$ estimates (that is Theorem 2.2) in § 7.

We keep assuming $a_0(x)\equiv 1$, which is not restrictive in view of Remark 2.3.

Let us recall the setting described at the end of § 3.3. For a fixed subdomain $\Omega'\Subset\Omega\subset\mathbb{R}^n$ and a fixed point $\overline{x}\in\Omega'$, let us consider a lifted ball $\widetilde{B}\left(\overline{\xi},R\right)\subset\mathbb{R}^N$ (with $\overline{\xi}=(\overline{x},0)$) where the lifted vector fields $\widetilde{X}_i$ are defined and satisfy Hörmander's condition, the map $\Theta$ is defined and satisfies the properties stated in § 3.1.

According to the procedure followed in [4, § 5], the proof of $C_X^{\alpha}$ a-priori estimates for second order derivatives will proceed in three steps: first, in the space of lifted variables, for test functions supported in a ball $\widetilde{B}\left(\overline{\xi},r\right)$ with $r$ small enough; then for any function in $C_{\widetilde{X}}^{2,\alpha}\left(\widetilde{B}\left(\overline{\xi},r\right)\right)$ (not necessarily vanishing at the boundary); then for any function in $C_X^{2,\alpha}\left(B\left(\overline{x},r\right)\right)$, that is in the original space. The three steps will be performed in separate subsections. The theory of singular integrals in locally homogeneous spaces will play its main role in the first step, considering the space

$$\left(\widetilde{\Omega}, \left\{\widetilde{\Omega}_k\right\}_{k=1}^{\infty}, d_{\widetilde{X}}, d\xi\right)$$

where

$$\widetilde{\Omega} = \widetilde{B}\left(\overline{\xi}, R\right); \widetilde{\Omega}_k = \widetilde{B}\left(\overline{\xi}, \frac{kR}{k+1}\right) \text{ for } k = 1, 2, 3, ...$$

## 6.1 Schauder estimates for functions with small support

The first step in the proof of Schauder estimates is contained in the following theorem, which is the main result in this subsection.

**Theorem 6.1** *Let $\widetilde{B}\left(\overline{\xi}, R\right)$ be as before. There exist $R_0 < R$ and $c > 0$ such that for every $u \in C^{2,\alpha}_{\widetilde{X},0}\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)$,*

$$\|u\|_{C^{2,\alpha}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)} \leq c\left\{\left\|\widetilde{\mathcal{L}}u\right\|_{C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)} + \|u\|_{L^{\infty}\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)}\right\}$$

*where $c$ and $R_0$ depend on $R, \left\{\widetilde{X}_i\right\}, \alpha, \mu,$ and $\|\widetilde{a}_{ij}\|_{C^{\alpha}\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)}$.*

**Proof.** This theorem relies on the representation formulas proved in § 4.2 and Theorem 5.1 about singular integrals on $C^{\alpha}$, in § 5.1. The proof is similar to that of [4, Thm. 5.2]. We start from the representation formula (4.29), choosing $r < R$ such that $\widetilde{B}_r \equiv \widetilde{B}\left(\overline{\xi}, r\right)$ be contained in the set where $a \equiv 1$. Taking $C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$ norm of both sides of (4.29) and applying Theorem 5.1 and (3.31) in Proposition 3.27, we can write, for any $u \in C^{2,\alpha}_{\widetilde{X},0}\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$ and $m, l = 1, 2, ..., q$

$$\begin{aligned}\left\|\widetilde{X}_m\widetilde{X}_l u\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \leq c\Bigg\{&\left\|\widetilde{\mathcal{L}}u\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} + \sum_{i,j=1}^{q}\left\|\left[\widetilde{a}_{ij}(\xi_0) - \widetilde{a}_{ij}\left(\cdot\right)\right]\widetilde{X}_i\widetilde{X}_j u\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \\ &+ \sum_{k=1}^{q}\left\|\widetilde{X}_k u\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} + \|u\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)}\Bigg\}\end{aligned}$$

for some $c$ depending on $R, \left\{\widetilde{X}_i\right\}, \alpha, \mu$.

To handle the terms involving $\widetilde{X}_i\widetilde{X}_j u$ in the right-hand side of the last inequality, we now exploit the fact that, for $u \in C^{2,\alpha}_{\widetilde{X},0}(\widetilde{B}_r)$, both $\widetilde{X}_i\widetilde{X}_j u$ and $\left[\widetilde{a}_{ij}(\xi_0) - \widetilde{a}_{ij}\left(\cdot\right)\right]$ vanish at a point of $\widetilde{B}_r$; then (3.32) in Proposition 3.27 implies

$$\left|\left[\widetilde{a}_{ij}(\xi_0) - \widetilde{a}_{ij}\left(\cdot\right)\right]\widetilde{X}_i\widetilde{X}_j u\right|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \leq c r^{\alpha}\left|\widetilde{a}_{ij}\right|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \cdot \left|\widetilde{X}_i\widetilde{X}_j u\right|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)},$$

while obviously

$$\left\| [\widetilde{a}_{ij}(\xi_0) - \widetilde{a}_{ij}(\cdot)] \; \widetilde{X}_i \widetilde{X}_j u \right\|_{L^\infty(\widetilde{B}_r)} \le c r^\alpha \left| \widetilde{a}_{ij} \right|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \cdot \left\| \widetilde{X}_i \widetilde{X}_j u \right\|_{L^\infty(\widetilde{B}_r)}.$$

This allows, for $r$ small enough, to get

$$\left\| \widetilde{X}_m \widetilde{X}_l u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \le c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \sum_{k=1}^q \left\| \widetilde{X}_k u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \|u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \right\} \quad (6.1)$$

for some $c$ also depending on $\left| \widetilde{a}_{ij} \right|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)}$. From the equation (4.1) we also read

$$\left\| \widetilde{X}_0 u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \le \left\| \widetilde{\mathcal{L}} u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + c \sum_{k,h=1}^q \left\| \widetilde{X}_k \widetilde{X}_h u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)}. \quad (6.2)$$

By (6.1) and (6.2) we get, for $r$ small enough,

$$\|u\|_{C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \le c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \|u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \sum_{k=1}^q \left\| \widetilde{X}_k u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \right\}. \quad (6.3)$$

Next, we want to get rid of the term $\left\| \widetilde{X}_k u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)}$ in the last inequality. Taking only one derivative in the parametrix formula (4.27) we have

$$\widetilde{X}_l(u) = S(\xi_0) \left( \widetilde{\mathcal{L}} u + \sum_{i,j=1}^q [\widetilde{a}_{ij}(\xi_0) - \widetilde{a}_{ij}(\xi)] \; \widetilde{X}_i \widetilde{X}_j u \right) + T(\xi_0) u,$$

where $S(\xi_0), T(\xi_0)$ are frozen operators of type $1, 0$, respectively. Taking $C^\alpha_{\widetilde{X}}(\widetilde{B}_r)$ norms of both sides and applying Theorem 5.1, we can write

$$\left\| \widetilde{X}_l u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \le c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \left\| \sum_{i,j=1}^q [\widetilde{a}_{ij}(\xi_0) - \widetilde{a}_{ij}(\xi)] \; \widetilde{X}_i \widetilde{X}_j u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \|u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \right\}$$

and reasoning as above,

$$\left\| \widetilde{X}_l u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \le c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + r^\alpha \left\| \widetilde{X}_i \widetilde{X}_j u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \|u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \right\} \quad (6.4)$$

Inserting (6.4) in (6.3), for $r$ small enough we get

$$\|u\|_{C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \le c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} + \|u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_r)} \right\}. \quad (6.5)$$

Finally, we want to replace the term $\|u\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)}$ with $\|u\|_{L^{\infty}(\widetilde{B}_r)}$ in the last inequality. To do this, we apply (3.30) in Proposition 3.27 and write

$$\|u\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \leq \|u\|_{L^{\infty}(\widetilde{B}_r)} + cr^{1-\alpha}\left(\sum_{i=1}^{q}\left\|\widetilde{X}_i u\right\|_{L^{\infty}(\widetilde{B}_r)} + r\left\|\widetilde{X}_0 u\right\|_{L^{\infty}(\widetilde{B}_r)}\right).$$

substituting this in (6.5), for $r$ small enough the term

$$\left(\sum_{i=1}^{q}\left\|\widetilde{X}_i u\right\|_{L^{\infty}(\widetilde{B}_r)} + r\left\|\widetilde{X}_0 u\right\|_{L^{\infty}(\widetilde{B}_r)}\right)$$

can be taken to the left-hand side, to get

$$\|u\|_{C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}_r)} \leq c\left\{\left\|\widetilde{\mathcal{L}}u\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}_r)} + \|u\|_{L^{\infty}(\widetilde{B}_r)}\right\},$$

so we are done. ■

## 6.2 Schauder estimates for nonvanishing functions

The second step in the proof of Schauder estimates consists in establishing *a priori* estimates for functions non necessarily compactly supported:

**Theorem 6.2** *There exist $r_0 < R_0$ and $c, \beta > 0$ (with $R_0$ as in Theorem 6.1) such that, for every $u \in C^{2,\alpha}_{\widetilde{X}}\left(\widetilde{B}(\overline{\xi}, r_0)\right)$, $0 < t < s < r_0$,*

$$\|u\|_{C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},t))} \leq \frac{c}{(s-t)^{\beta}}\left\{\left\|\widetilde{\mathcal{L}}u\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},s))} + \|u\|_{L^{\infty}(\widetilde{B}(\overline{\xi},s))}\right\},$$

*where $r_0, c$ depend on $R$, $\left\{\widetilde{X}_i\right\}_{i=1}^{q}$, $\alpha, \mu, \|\widetilde{a}_{ij}\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))}$; $\beta$ depends on $\left\{\widetilde{X}_i\right\}_{i=0}^{q}$ and $\alpha$.*

As in [4], this result relies on interpolation inequalities for $C^{k,\alpha}_{\widetilde{X}}$ norms and the use of suitable cutoff function. The following result can be proved as [4, Lemma 6.2], by the results in Proposition 3.27.

**Lemma 6.3 (cutoff functions)** *For any $0 < \rho < r$, $\xi \in \widetilde{B}\left(\overline{\xi}, R\right)$ there exists $\varphi \in C_0^{\infty}\left(\mathbb{R}^N\right)$ with the following properties:*

**i)** $0 \leqslant \varphi \leqslant 1$, $\varphi \equiv 1$ *on* $\widetilde{B}(\xi, \rho)$ *and* $\operatorname{sprt}\varphi \subseteq \widetilde{B}(\xi, r)$*;*

**ii)** *for* $i, j = 1, 2, ..., q$,

$$\left|\widetilde{X}_i\varphi\right| \leqslant \frac{c}{r-\rho} \tag{6.6}$$

$$\left|\widetilde{X}_0\varphi\right|, \left|\widetilde{X}_i\widetilde{X}_j\varphi\right| \leqslant \frac{c}{(r-\rho)^2}$$

**iii)** *For any $f \in C^{\alpha}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, R\right)\right)$, and $r-\rho$ small enough,*

$$\left\| f\,\widetilde{X}_i \varphi\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))} \leqslant \frac{c}{(r-\rho)^2}\,\|f\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))}$$

$$\left\| f\,\widetilde{X}_0 \varphi\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))}, \left\| f\widetilde{X}_i\widetilde{X}_j \varphi\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))} \leqslant \frac{c}{(r-\rho)^3}\,\|f\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))}. \tag{6.7}$$

We will write

$$\widetilde{B}_{\rho}(\xi) \prec \varphi \prec \widetilde{B}_r(\xi)$$

to indicate that $\varphi$ satisfies all the previous properties.

Next, let us state the following:

**Proposition 6.4 (Interpolation inequality for test functions)** *Let*

$$H = \sum_{i=1}^{q} \widetilde{X}_i^2 + \widetilde{X}_0$$

*and let $\widetilde{B}\left(\overline{\xi}, R\right)$ be as before. Then for every $\alpha \in (0,1)$, there exist constants $\gamma \geq 1$ and $c > 0$, depending on $\alpha, R$ and $\left\{\widetilde{X}_i\right\}$, such that for every $\varepsilon \in (0,1)$ and every $f \in C_0^{\infty}(\widetilde{B}\left(\overline{\xi}, R/2\right))$,*

$$\left\|\widetilde{X}_l f\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R/2))} \leq \varepsilon \|Hf\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R/2))} + \frac{c}{\varepsilon^{\gamma}}\|f\|_{L^{\infty}(\widetilde{B}(\overline{\xi},R/2))} \tag{6.8}$$

*for $l = 1, 2, \cdots, q$; moreover, we have*

$$\|Df\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R/2))} \leq \varepsilon \left\|\widetilde{\mathcal{L}} f\right\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R/2))} + \frac{c}{\varepsilon^{\gamma}}\|f\|_{L^{\infty}(\widetilde{B}(\overline{\xi},R/2))}, \tag{6.9}$$

*where $D$ is any vector field of local degree $\leq 1$.*

To prove Proposition 6.4, we need the following

**Lemma 6.5** *Let $P(\xi_0)$ be a frozen operator of type $\lambda \geq 1$ over $\widetilde{B}\left(\overline{\xi}, R\right)$ and $\alpha \in (0,1)$. Then there exist positive constants $\gamma > 1$ and $c$, depending on $\alpha, \mu$ and $\left\{\widetilde{X}_i\right\}$, such that for every $f \in C_0^{\infty}(\widetilde{B}\left(\overline{\xi}, R\right))$ and $\varepsilon \in (0,1)$*

$$\|PHf\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))} \leq \varepsilon \|Hf\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))} + \frac{c}{\varepsilon^{\gamma}}\|f\|_{L^{\infty}(\widetilde{B}(\overline{\xi},R))}. \tag{6.10}$$

**Remark 6.6** *As will be clear from the proof, (6.10) still holds if $H$ is replaced by any differential operator of weight two, like $\widetilde{X}_i\widetilde{X}_j$ or $\widetilde{X}_0$.*

**Proof of the Lemma.** This proof is adapted from [4, Lemma 7.2]. Let

$$PHf\,(\xi) = \int_{\widetilde{B}\left(\overline{\xi},R\right)} k(\xi,\eta)Hf\,(\eta)\,d\eta,$$

where $k$ is a frozen kernel of type $\lambda \geq 1$, and let $\varphi_\varepsilon$ be a cutoff function such that $\widetilde{B}_{\varepsilon/2}(\xi) \prec \varphi_\varepsilon \prec \widetilde{B}_\varepsilon(\xi)$, for $\varepsilon \in (0,1)$. We split $PH$ as follows: for $\xi \in \widetilde{B}\left(\overline{\xi},R\right)$

$$\begin{aligned} PHf(\xi) &= \int_{\widetilde{B}\left(\overline{\xi},R\right),\rho(\xi,\eta)>\frac{\varepsilon}{2}} k(\xi,\eta)[1-\varphi_\varepsilon(\eta)]Hf(\eta)d\eta \\ &+ \int_{\widetilde{B}\left(\overline{\xi},R\right),\rho(\xi,\eta)\leq\varepsilon} k(\xi,\eta)Hf(\eta)\varphi_\varepsilon(\eta)d\eta \\ &= I(\xi) + II(\xi). \end{aligned}$$

Then

$$I(\xi) = \int_{\widetilde{B}\left(\overline{\xi},R\right),\rho(\xi,\eta)>\frac{\varepsilon}{2}} H^T\left(k(\xi,\cdot)[1-\varphi_\varepsilon(\cdot)]\right)(\eta)\,f(\eta)d\eta.$$

Let $h^\varepsilon(\xi,\eta) = H^T\left(k(\xi,\cdot)[1-\varphi_\varepsilon(\cdot)]\right)(\eta)$. Since $k$ is a frozen kernel of type $\lambda$, there exist $c>0, \gamma>1$, such that

$$|h^\varepsilon(\xi,\eta)| + \left|\widetilde{X}_0 h^\varepsilon(\xi,\eta)\right| + \left|\sum \widetilde{X}_i h^\varepsilon(\xi,\eta)\right| \leq c\varepsilon^{-\gamma}.$$

By definition of frozen kernel, the function $\xi \longmapsto h^\varepsilon(\xi,\eta)$ is compactly supported in $\widetilde{B}\left(\overline{\xi},R\right)$ for any $\eta \in \widetilde{B}\left(\overline{\xi},R\right)$, hence by (3.29) in Proposition 3.27, it follows that

$$|h^\varepsilon(\xi_1,\eta) - h^\varepsilon(\xi_2,\eta)| \leq c_R d_{\widetilde{X}}(\xi_1,\xi_2)\varepsilon^{-\gamma} \leq c_R\rho(\xi_1,\xi_2)\varepsilon^{-\gamma}$$

for any $\xi_1,\xi_2 \in \widetilde{B}\left(\overline{\xi},R\right)$, and therefore

$$\begin{aligned} |I(\xi_1) - I(\xi_2)| &\leq \int |h^\varepsilon(\xi_1,\eta) - h^\varepsilon(\xi_2,\eta)|\,|f(\eta)|\,d\eta \\ &\leq c_R\varepsilon^{-\gamma}\rho(\xi_1,\xi_2)\left|\widetilde{B}_R\right|\|f\|_{L^\infty(\widetilde{B}_R)}. \end{aligned}$$

Also, since

$$|I(\xi)| \leq \int_{\widetilde{B}\left(\overline{\xi},R\right),\rho(\xi,\eta)>\frac{\varepsilon}{2}} c\varepsilon^{-\gamma}\,|f(\eta)|\,d\eta \leq c\varepsilon^{-\gamma}\left|\widetilde{B}_R\right|\|f\|_{L^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)},$$

we obtain

$$\|I(\xi)\|_{C^\alpha_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi},R\right)\right)} \leq c\varepsilon^{-\gamma}\|f\|_{L^\infty\left(\widetilde{B}\left(\overline{\xi},R\right)\right)} \text{ for any } \alpha \in (0,1).$$

Now, let us consider $II(\xi)$, and let

$$k_\varepsilon(\xi,\eta) = k(\xi,\eta)\varphi_\varepsilon(\eta).$$

By the properties of frozen kernels of type 1, keeping into account the support of $k_\varepsilon$ and applying again (3.29) in Proposition 3.27, we can say that for any fixed $\delta \in (0,1)$, the kernel $k_\varepsilon(\xi,\eta)$ satisfies the following standard estimates of fractional integral kernels (see § 3.3):

$$|k_\varepsilon(\xi,\eta)| \le c\rho(\xi,\eta)^{1-Q} \le c\varepsilon^\delta \rho(\xi,\eta)^{1-\delta-Q}, \tag{6.11}$$

$$|k_\varepsilon(\xi,\eta) - k_\varepsilon(\xi_1,\eta)| \le c\frac{\rho(\xi,\xi_1)}{\rho(\xi,\eta)^Q} \le c\varepsilon^\delta \rho(\xi,\eta)^{-\delta-Q}\rho(\xi,\xi_1) \tag{6.12}$$

for $\rho(\xi,\eta) \ge 2\rho(\xi,\xi_1)$. Therefore, by Theorem 3.14 and Remark 3.19 in § 3.3,

$$\|II\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))} \le c\varepsilon^\delta \|Hf\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))}$$

for any $\alpha < 1-\delta$. We conclude that for every $\alpha \in (0,1)$ there exist $\delta$, $\gamma, c > 0$ such that

$$\|PHf\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))} \le \varepsilon^\delta \|Hf\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},R))} + \frac{c}{\varepsilon^\gamma}\|f\|_{L^\infty(\widetilde{B}(\overline{\xi},R))},$$

which implies the lemma. ■

**Proof of Proposition 6.4.** By Theorem 4.18, we can write

$$af = PHf(\xi) + Sf,$$

where $P, S$ are frozen operators of type 2 and 1, respectively, over $\widetilde{B}\left(\overline{\xi},R\right)$. More precisely, they should be called "constant kernels of type 2 and 1", since they satisfy the definition of frozen kernels with the matrix $\{\widetilde{a}_{ij}(\xi_0)\}$ replaced by the identity matrix.

If we assume $a = 1$ on $\widetilde{B}\left(\overline{\xi},R/2\right)$, then, for $f \in C_0^\infty(\widetilde{B}\left(\overline{\xi},R/2\right))$ we obtain

$$f = PHf(\xi) + Sf \tag{6.13}$$

and therefore, by Theorem 4.11,

$$\widetilde{X}_i f = S_1 Hf(\xi) + Tf, \tag{6.14}$$

where $S_1, T$ are frozen operators of type 1 and 0, respectively. Substituting (6.13) in (6.14) yields

$$\widetilde{X}_i f = S_1 Hf(\xi) + TPHf + TSf$$

and therefore, by Theorem 5.1 and Lemma 6.5

$$\begin{aligned}\left\|\widetilde{X}_i f\right\|_\alpha &\le \|S_1 Hf\|_\alpha + \|TPHf\|_\alpha + \|TSf\|_\alpha \\ &\le \|S_1 Hf\|_\alpha + c\left\{\|PHf\|_\alpha + \|Sf\|_\alpha\right\} \\ &\le c\left\{\varepsilon\|Hf\|_\alpha + \varepsilon^{-\gamma}\|f\|_\infty + \|Sf\|_\alpha\right\}\end{aligned} \tag{6.15}$$

where all the norms are taken over $\widetilde{B}\left(\overline{\xi}, R/2\right)$. We end the proof by showing that for an operator $S$ of type 1,

$$\|Sf\|_\alpha \leq c\,\|f\|_{L^\infty},$$

which by (6.15) will complete the proof of the first inequality in the proposition.

Indeed, if

$$Sf(\xi) = \int_{\widetilde{B}_R} k(\xi,\eta)f(\eta)d\eta.$$

We have

$$|Sf(\xi_1) - Sf(\xi_2)| \leq \|f\|_{L^\infty(\widetilde{B}_R)} \int_{\widetilde{B}\left(\overline{\xi},R\right)} |k(\xi_1,\eta) - k(\xi_2,\eta)|\, d\eta. \tag{6.16}$$

Moreover,

$$\begin{aligned}\int_{\widetilde{B}_R} |k(\xi_1,\eta) - k(\xi_2,\eta)|\, d\eta &= \int_{\widetilde{B}\left(\overline{\xi},R\right),\rho(\xi_1,\eta)>M\rho(\xi_1,\xi_2)} |k(\xi_1,\eta) - k(\xi_2,\eta)|\, d\eta \\ &+ \int_{\widetilde{B}\left(\overline{\xi},R\right),\rho(\xi_1,\eta)\leq M\rho(\xi_1,\xi_2)} |k(\xi_1,\eta) - k(\xi_2,\eta)|\, d\eta \\ &\equiv I + II.\end{aligned}$$

By (6.12),

$$\begin{aligned} I &\leq \int_{\rho(\xi_1,\eta)>M\rho(\xi_1,\xi_2)} \frac{c}{\rho(\xi_1,\eta)^{Q-1}} \frac{\rho(\xi_1,\xi_2)}{\rho(\xi_1,\eta)} d\eta \\ &= \rho(\xi_1,\xi_2)^\alpha \int_{\rho(\xi_1,\eta)>M\rho(\xi_1,\xi_2)} \frac{\rho(\xi_1,\eta)^{1-\alpha}}{\rho(\xi_1,\eta)^Q} \frac{\rho(\xi_1,\xi_2)^{1-\alpha}}{\rho(\xi_1,\eta)^{1-\alpha}} d\eta \\ &\leq c\rho(\xi_1,\xi_2)^\alpha \int_{\widetilde{B}_R} \frac{\rho(\xi_1,\eta)^{1-\alpha}}{\rho(\xi_1,\eta)^Q} d\eta \\ &\leq c\rho(\xi_1,\xi_2)^\alpha R^{1-\alpha}, \end{aligned}$$

where in the last inequality we have used the following standard computation (which will be useful also other times):

$$\int_{\widetilde{B}\left(\overline{\xi},R\right),\rho(\xi_1,\eta)<r} \frac{d\eta}{\rho(\xi_1,\eta)^{Q-\beta}} \leq cr^\beta \text{ for any } \xi_1 \in \widetilde{B}\left(\overline{\xi}, R\right) \tag{6.17}$$

As to $II$, by (6.11),

$$II \leq \int_{\rho(\xi_1,\eta)\leq M\rho(\xi_1,\xi_2)} |k(\xi_1,\eta)|\, d\eta + \int_{\rho(\xi_1,\eta)\leq M\rho(\xi_1,\xi_2)} |k(\xi_2,\eta)|\, d\eta$$

since there exists $M_1 > 0$ such that if $\rho(\xi_1,\eta) \leq M\rho(\xi_1,\xi_2)$ then $\rho(\xi_2,\eta) \leq M_1\rho(\xi_1,\xi_2)$,

$$\leq c\left\{\int_{\rho(\xi_1,\eta)\leq M\rho(\xi_1,\xi_2)} \frac{1}{\rho(\xi_1,\eta)^{Q-1}} d\eta + \int_{\rho(\xi_2,\eta)\leq M_1\rho(\xi_1,\xi_2)} \frac{1}{\rho(\xi_2,\eta)^{Q-1}} d\eta\right\}$$

again by (6.17)

$$\leq c\rho\left(\xi_1,\xi_2\right) \leq c\rho\left(\xi_1,\xi_2\right)^{\alpha} R^{1-\alpha}.$$

Hence for every $\alpha \in (0,1)$,

$$\int_{\widetilde{B}_R} |k(\xi_1,\eta) - k(\xi_2,\eta)|\, d\eta \leq c_\alpha \rho(\xi_1,\xi_2)^\alpha R^{1-\alpha}$$

and, by (6.16),

$$|Sf|_\alpha \leq c\,\|f\|_{L^\infty}.$$

Moreover,

$$\begin{aligned} |Sf(\xi)| &\leq \int_{\widetilde{B}_R} |k(\xi,\eta) f(\eta)|\, d\eta \\ &\leq \|f\|_{L^\infty} \int_{\rho(\xi,\eta)\leq cR} \frac{c}{\rho(\xi,\eta)^{Q-1}} d\eta \leq cR\,\|f\|_{L^\infty}, \end{aligned}$$

hence

$$\|Sf\|_\alpha \leq c\,\|f\|_{L^\infty}.$$

This completes the proof of (6.8). A similar argument gives (6.9). ■

**Theorem 6.7 (Interpolation inequality)** *There exist positive constants $c,\gamma$ and $r_1 < R$ such that for any $u \in C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}\left(\overline{\xi},r_1\right))$, $0<\rho<r_1$, $0<\delta<1/3$,*

$$\left\|\widetilde{D}u\right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},\rho))} \leq \delta \sum_{i=1}^{q} \left\|\widetilde{D}^2 u\right\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\overline{\xi},r_1))} + \frac{c}{\delta^\gamma \left(r_1-\rho\right)^{2\gamma}} \|u\|_{L^\infty(\widetilde{B}(\overline{\xi},r_1))}$$

*where*

$$\left\|\widetilde{D}u\right\| \equiv \sum_{i=1}^{q} \left\|\widetilde{X}_i u\right\| \ \text{ and } \ \left\|\widetilde{D}^2 u\right\| \equiv \sum_{i,j=1}^{q} \left\|\widetilde{X}_i \widetilde{X}_i u\right\| + \left\|\widetilde{X}_0 u\right\|.$$

*The constants $c, r_1, \gamma$ depend on $\alpha, \left\{\widetilde{X}_i\right\}$; $\gamma$ is as in Proposition 6.4.*

**Proof.** The proof can be carried out exactly as in [4, Proposition 7.4], exploiting the properties of cutoff functions (Lemma 6.3), the interpolation inequality for test functions (Proposition 6.4) and (3.30) in Proposition 3.27. ■

We are now in position for the main goal of this subsection:
**Proof of Theorem 6.2.** This proof can now be carried out exactly like in [4, Theorem 5.3], exploiting: Schauder estimates for functions with small support (Theorem 6.1), the properties of Hölder continuous functions contained in (3.30), (3.31), (3.34), the properties of cutoff functions (Lemma 6.3) and the interpolation inequalities contained in Theorem 6.7 and (6.9). ■

## 6.3 Schauder estimates in the original variables

Letṡ now prove Theorem 2.1. We finally come back to our original context, which we are going to recall. We have a bounded domain $\Omega$ where our vector fields and coeffi cients are defined, and a fixed subdomain $\Omega' \Subset \Omega$. Fix $\overline{x} \in \Omega'$ and $R$ such that in $B(\overline{x}, R) \subset \Omega$ all the construction of the previous two subsections (lifting to $\widetilde{B}\left(\overline{\xi}, R\right)$ and so on) can be performed. Let $r_0$ be as in Theorem 6.2. To begin with, we want to prove Schauder estimates for functions $u \in C_X^{2,\alpha}(B(\overline{x}, r_0))$. By Theorem 3.28 we know that the function $\widetilde{u}(x,h) = u(x)$ belongs to $C_{\widetilde{X}}^{2,\alpha}\left(B\left(\overline{\xi}, r_0\right)\right)$, so we can apply to $\widetilde{u}$ Schauder estimates contained in Theorem 6.2. Combining this fact with the two estimates in Theorem 3.28 and choosing $t, s$ such that

$$r_0 > s > t > 0 \text{ and } s - t = r_0 - s,$$

we get, for some exponent $\omega > 2$

$$\begin{aligned} \|u\|_{C_X^{2,\alpha}(B(\overline{x},s))} &\leq \frac{c}{(s-t)^2} \|\widetilde{u}\|_{C_{\widetilde{X}}^{2,\alpha}\left(\widetilde{B}\left(\overline{\xi},t\right)\right)} \\ &\leq \frac{c}{(r_0-t)^\omega} \left( \left\| \widetilde{\mathcal{L}}\widetilde{u} \right\|_{C_{\widetilde{X}}^{\alpha}\left(\widetilde{B}\left(\overline{\xi},r_0\right)\right)} + \|\widetilde{u}\|_{L^\infty\left(\widetilde{B}\left(\overline{\xi},r_0\right)\right)} \right) \\ &\leq \frac{c}{(r_0-s)^\omega} \left( \|\mathcal{L}u\|_{C_X^\alpha(B(\overline{x},r_0))} + \|u\|_{L^\infty(B(\overline{x},r_0))} \right) \end{aligned} \tag{6.18}$$

since $\widetilde{\mathcal{L}}\widetilde{u} = \widetilde{(\mathcal{L}u)}$.

Next, let us choose a family of balls $B(x_i, r_0)$ in $\Omega$ such that

$$\Omega' \subset \bigcup_{i=1}^k B(x_i, r_0/2) \subset \bigcup_{i=1}^k B(x_i, r_0) \subset \Omega.$$

Then by Proposition 3.27 (v) and (6.18), with $s = r_0/2$,

$$\begin{aligned} \|u\|_{C_X^{2,\alpha}(\Omega')} \leq \|u\|_{C_X^{2,\alpha}(\cup B(x_i,r_0/2))} &\leq c \sum_{i=1}^k \|u\|_{C_X^{2,\alpha}(B(x_i,r_0))} \\ &\leq c \sum_{i=1}^k \left\{ \|\mathcal{L}u\|_{C_X^\alpha(B(x_i,r_0))} + \|u\|_{L^\infty(B(x_i,r_0))} \right\} \\ &\leq c \left\{ \|\mathcal{L}u\|_{C_X^\alpha(\Omega)} + \|u\|_{L^\infty(\Omega)} \right\} \end{aligned}$$

with $c$ also depending on $r_0$. Finally, let us note that the constant $c$ depends on the coeffi cients$a_{ij}$ through the norms

$$\|\widetilde{a}_{ij}\|_{C_{\widetilde{X}}^\alpha\left(\widetilde{B}\left(\overline{\xi},R_0\right)\right)},$$

which in turn are bounded by the norms

$$\|a_{ij}\|_{C_X^\alpha(B(\overline{x},R_0))}$$

(by Proposition 3.28), and hence by $\|a_{ij}\|_{C^{\alpha}_{X}(\Omega)}$ (or more precisely, by $\|a_{ij}\|_{C^{\alpha}_{X}(\Omega'')}$ for some $\Omega''$ such that $\Omega' \Subset \Omega'' \Subset \Omega$). This completes the proof of Theorem 2.1.

# 7 $L^p$ estimates

The logical structure of this section, as well as the general setting, is very similar to that of the previous one. Here, following as close as possible the strategy of [3], the proof will be still divided into three steps: $L^p$ a-priori estimates in the space of lifted variables, for test functions supported in a ball $\widetilde{B}\left(\overline{\xi}, r\right)$ with $r$ small enough; then for any function in $S^{2,p}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$ (not necessarily vanishing at the boundary); then for any function in $S^{2,p}_{X}\left(B\left(\overline{x}, r\right)\right)$, that is in the original space.

Again, it is not restrictive to assume $a_0 = 1$.

The basic difference with the setting of Schauder estimates consists in the fact that here we start with representation formulas where the "frozen" point has been finally unfrozen; therefore now singular integrals with *variable* kernels are involved, together with their commutators with $VMO$ functions. This makes the singular integral part of the theory more involved.

## 7.1 $L^p$ estimates for functions with small support

**Theorem 7.1** *Let $\widetilde{B}\left(\overline{\xi}, R\right)$ be as in the previous section, and $p \in (1, \infty)$. There exists $R_0 < R$ such that for every $u \in C_0^{\infty}\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)$,*

$$\|u\|_{S^{2,p}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)} \leq c\left\{\left\|\widetilde{\mathcal{L}}u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)} + \|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, R_0\right)\right)}\right\} \tag{7.1}$$

*for some constant $c$ depending on $\left\{\widetilde{X}_i\right\}_{i=0}^{q}, p, \mu, R$; the number $R_0$ also depends on the local $VMO$ moduli $\eta^{*}_{a_{ij},\Omega',\Omega}$.*

**Proof.** This theorem relies on the representation formula proved in Theorem 4.21 and on the results about singular integrals and commutators contained in Theorem 5.4.

Let $u \in C_0^{\infty}\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$ with $r < R$. Let us write the representation formula of Theorem 4.21 choosing the cutoff function $a$ such that $a = 1$ in $\widetilde{B}\left(\overline{\xi}, r\right)$. Taking $L^p$ norms of both sides of the formula we get (see also Remark 4.22), for $p \in (1, \infty)$, any $m, l = 1, 2, ..., q$,

$$\left\|\widetilde{X}_m \widetilde{X}_l u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)} \leq \left\|T_{lm}\widetilde{\mathcal{L}}u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)} + \sum_{i,j=1}^{q}\left\|\left[\widetilde{a}_{ij}, T_{lm}\right]\, \widetilde{X}_i \widetilde{X}_j u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)} +$$
$$+ \sum_{k=1}^{q}\left\|T_{lm,k}\widetilde{X}_k u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)} + \left\|T^0_{lm} u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)}$$

where all the $T_{lm}, T_{lm,k}, T_{lm}^0$ are *variable* operators of type 0 over $\widetilde{B}\left(\overline{\xi}, 2r\right)$.

We now apply Theorem 5.4: there exists $c$ (depending on $R$) and for every fixed $\varepsilon > 0$ there exists $r < R$ such that for every $u \in C_0^\infty\left(\widetilde{B}\left(\overline{\xi}, r\right)\right)$,

$$\left\|\widetilde{X}_m\widetilde{X}_l u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq c\left\{\left\|\widetilde{\mathcal{L}}u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} + \varepsilon\sum_{i,j=1}^q\left\|\widetilde{X}_i\widetilde{X}_j u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}\right. \tag{7.2}$$
$$\left. + \left\|u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} + \sum_k\left\|\widetilde{X}_k u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}\right\}.$$

Now, let us come back to (4.27) and take only one derivative $\widetilde{X}_l$ (for $l = 1, ..., q$) of both sides; we find:

$$\begin{aligned}\widetilde{X}_l u &= \widetilde{X}_l P\left(\xi_0\right)\widetilde{\mathcal{L}}_0 u + \widetilde{X}_l S\left(\xi_0\right) u \\ &= S_l\left(\xi_0\right)\widetilde{\mathcal{L}}_0 + T\left(\xi_0\right) u\end{aligned} \tag{7.3}$$

where $S_l\left(\xi_0\right), T\left(\xi_0\right)$ are frozen operators of type 1 and 0, respectively. By (4.30) we have

$$\widetilde{X}_l u = S_{1,l}\widetilde{\mathcal{L}}u + \sum_{i,j=1}^q\left[S_{1,l}, \widetilde{a}_{ij}\right]\widetilde{X}_i\widetilde{X}_j u + Tu,$$

where $S_{1,l}$, $T$ are variable operators of type 1 and 0, respectively (in particular, both can be seen as operators of type 0). By Theorem 5.4,

$$\left\|\widetilde{X}_l u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq c\left\|\widetilde{\mathcal{L}}u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} + \varepsilon\sum_{i,j=1}^q\left\|\widetilde{X}_i\widetilde{X}_j u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} + c\left\|u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}. \tag{7.4}$$

Finally, from the equation we can bound

$$\left\|\widetilde{X}_0 u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq c\sum_{i,j=1}^q\left\|\widetilde{X}_i\widetilde{X}_j u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} + \left\|\widetilde{\mathcal{L}}u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}. \tag{7.5}$$

Combining (7.2), (7.4) and (7.5) we have, for $r$ small enough,

$$\left\|u\right\|_{S_{\widetilde{X}}^{2,p}\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} \leq c\left(\left\|\widetilde{\mathcal{L}}u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} + \left\|u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}\right) \tag{7.6}$$

and the theorem is proved. ■

## 7.2 $L^p$ estimates for nonvanishing functions

The main result in this subsection is the following:

**Theorem 7.2** *Let $\widetilde{B}\left(\overline{\xi},R\right)$ be as before. There exists $r_0 < R$ and for any $r \leq r_0$ there exists $c > 0$ such that for any $u \in S^{2,p}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi},r\right)\right)$ we have*

$$\|u\|_{S^{2,p}_{\widetilde{X}}\left(\widetilde{B}\left(\overline{\xi},r/2\right)\right)} \leq c\left\{\left\|\widetilde{\mathcal{L}}u\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)} + \|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi},r\right)\right)}\right\}.$$

*The constants $c, r_0$ depend on $\left\{\widetilde{X}_i\right\}_{i=0}^{q}, p, \mu, R,$and $\eta^*_{a_{ij},\Omega',\Omega}$; $c$ also depends on $r$.*

Analogously to what seen in § 6.2, the proof of the above theorem relies on interpolation inequalities for Sobolev norms and the use of cutoff function.

Regarding cutoff functions, we need the following statement:

**Lemma 7.3 (Radial cutoff functions)** *For any $\sigma \in (\frac{1}{2},1)$, $r > 0$ and $\xi \in \widetilde{B}\left(\overline{\xi},r\right)$, there exists $\varphi \in C_0^\infty\left(\mathbb{R}^N\right)$ with the following properties:*

*i) $\widetilde{B}_{\sigma r}(\xi) \prec \varphi \prec \widetilde{B}_{\sigma' r}(\xi)$ with $\sigma' = (1+\sigma)/2$ (this means that $\varphi = 1$ in $\widetilde{B}_{\sigma r}(\xi)$ and it is supported in $\widetilde{B}_{\sigma' r}(\xi)$);*

*ii) for $i,j = 1,\ldots,q$, we have*

$$\left|\widetilde{X}_i\varphi\right| \leq \frac{c}{(1-\sigma)r}; \left|\widetilde{X}_0\varphi\right|, \left|\widetilde{X}_i\widetilde{X}_j\varphi\right| \leq \frac{c}{(1-\sigma)^2r^2}. \tag{7.7}$$

The above lemma, very similar to [3, Lemma 3.3], is actually contained in Lemma 6.3, but we have preferred to state it explicitly because it is formulated in a slightly different notation, suitable to our application to $L^p$ estimates.

**Theorem 7.4 (Interpolation inequality for Sobolev norms)** *Let $\widetilde{B}(\overline{\xi},R)$ be as before. For every $p \in (1,\infty)$ there exists $c > 0$ and $r_1 < R$ such that for every $0 < \varepsilon \leq 4r_1$, $u \in C_0^\infty\left(\widetilde{B}(\overline{\xi},r_1)\right)$, then*

$$\left\|\widetilde{X}_i u\right\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)} \leq \varepsilon\|Hu\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)} + \frac{c}{\varepsilon}\|u\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)} \tag{7.8}$$

*for every $i = 1,\ldots,q$, where $H \equiv \sum_{i=1}^q \widetilde{X}_i^2 + \widetilde{X}_0$.*

**Proof.** The proof of this proposition is adapted from [3, Thm. 3.6].

Let $r_1$ be a small number to be fixed later. Like in the proof of Theorem 6.4 we can write, for any $u \in C_0^\infty\left(\widetilde{B}(\overline{\xi},r_1)\right)$ and $\xi \in \widetilde{B}(\overline{\xi},r_1)$,

$$\widetilde{X}_i u(\xi) = SHu(\xi) + Tu(\xi),$$

where $S,T$ are constant operators of type 1 and 0, respectively, over $\widetilde{B}(\overline{\xi},2r_1)$, provided $2r_1 < R$. (See the proof of Proposition 6.4 for the explanation of the term "constant operators of type $\lambda$"). Since

$$\|Tu\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)} \leq c\|u\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)},$$

the result will follow if we prove that

$$\|SHu\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)} \leq \varepsilon \|Hu\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)} + \frac{c}{\varepsilon} \|u\|_{L^p\left(\widetilde{B}(\overline{\xi},r_1)\right)}. \tag{7.9}$$

Let $k(\xi,\eta)$ be the kernel of $S$ and, for any fixed $\xi \in \widetilde{B}(\overline{\xi},r_1)$, $\varphi_\varepsilon$ a cutoff function (as in lemma 7.3) with $\widetilde{B}_{\frac{\varepsilon}{2}}(\xi) \prec \varphi_\varepsilon \prec \widetilde{B}_\varepsilon(\xi)$. Let us split:

$$\begin{aligned} SHu(\xi) = &\int_{\widetilde{B}(\overline{\xi},r_1),\rho(\xi,\eta)>\frac{\varepsilon}{2}} k(\xi,\eta)[1-\varphi_\varepsilon(\eta)]Hu(\eta)d\eta + \\ &+ \int_{\widetilde{B}(\overline{\xi},r_1),\rho(\xi,\eta)\leq\varepsilon} k(\xi,\eta)Hu(\eta)\varphi_\varepsilon(\eta)d\eta = I(\xi) + II(\xi). \end{aligned}$$

Then

$$\begin{aligned} |I(\xi)| &= \left| \int_{\widetilde{B}(\overline{\xi},r_1),\rho(\xi,\eta)>\frac{\varepsilon}{2}} H^T\left(k(\xi,\cdot)[1-\varphi_\varepsilon(\cdot)]\right)(\eta)\, u(\eta)d\eta \right| \\ &\leq \int_{\widetilde{B}(\overline{\xi},r_1),\rho(\xi,\eta)>\frac{\varepsilon}{2}} \left\{ \left|[1-\varphi_\varepsilon]H^T k(\xi,\cdot)\right| + \right. \\ &\left. + c\sum \left|\widetilde{X}_i[1-\varphi_\varepsilon]\cdot \widetilde{X}_j k(\xi,\cdot)\right| + \left|k(\xi,\cdot)\, H^T[1-\varphi_\varepsilon]\right| (\eta)\, |u(\eta)|\, d\eta \right. \\ &\equiv A(\xi) + B(\xi) + C(\xi). \end{aligned}$$

Recall that, for $i,j = 1,2,...,q$,

$$\begin{aligned} |k(\xi,\eta)| &\leq \frac{c}{d(\xi,\eta)^{Q-1}}; \\ \left|\widetilde{X}_i k(\xi,\eta)\right| &\leq \frac{c}{d(\xi,\eta)^{Q}}; \\ \left|H^T k(\xi,\cdot)(\eta)\right| &\leq \frac{c}{d(\xi,\eta)^{Q+1}}; \\ \left|\widetilde{X}_i(1-\varphi_\varepsilon)(\eta)\right| &\leq \frac{c}{\varepsilon}, \left|H^T(1-\varphi_\varepsilon)(\eta)\right| \leq \frac{c}{\varepsilon^2} \end{aligned}$$

and the derivatives of $(1-\varphi_\varepsilon)$ are supported in the annulus $\frac{\varepsilon}{2} \leq d(\xi,\eta) \leq \varepsilon$. Since $\xi,\eta \in \widetilde{B}(\overline{\xi},r_1)$, we have $d(\xi,\eta) < 2r_1$. Hence letting $k_0$ be the integer such that $2^{k_0-1}\varepsilon < 2r_1 \leq 2^{k_0}\varepsilon$ we have

$$\begin{aligned} |A(\xi)| &\leq c\sum_{k=0}^{k_0} \int_{2^{k-1}\varepsilon<\rho(\xi,\eta)\leq 2^k\varepsilon} \frac{c}{d(\xi,\eta)^{Q+1}} |u(\eta)|\, d\eta \\ &\leq c\sum_{k=0}^{k_0} \frac{1}{2^{k-1}\varepsilon} \frac{1}{(\varepsilon 2^{k-1})^Q} \int_{\rho(\xi,\eta)\leq 2^k\varepsilon} |u(\eta)|\, d\eta \\ &\leq \frac{c}{\varepsilon}\cdot \sup_{r\leq 4r_1} \frac{1}{\left|\widetilde{B}(\xi,r)\right|} \int_{\widetilde{B}(\xi,r)} |u(\eta)|\, d\eta. \end{aligned} \tag{7.10}$$

We now have to recall the definition of the local maximal function $\mathcal{M}$ (see § 3.3). With the notation of Theorem 3.21, we have $4r_1 = r_n = \frac{2}{5}\varepsilon_n$, hence $\varepsilon_n = 10r_1$ and, for $\xi \in \widetilde{B}\left(\overline{\xi}, r_1\right)$, we have $\widetilde{B}\left(\xi, \varepsilon_n\right) \subset \widetilde{B}\left(\overline{\xi}, 11r_1\right)$. Therefore by (7.10) we can write

$$|A(\xi)| \leq \frac{c}{\varepsilon} \cdot \mathcal{M}_{\widetilde{B}\left(\overline{\xi}, r_1\right), \widetilde{B}\left(\overline{\xi}, 11r_1\right)} u(\xi)$$

and by Theorem 3.21, we have

$$\|A\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)} \leq \frac{c}{\varepsilon} \|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, 11r_1\right)\right)} = \frac{c}{\varepsilon} \|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)},$$

since $u \in C_0^\infty\left(\widetilde{B}(\overline{\xi}, r_1)\right)$, provided $11r_1 < R$. Also

$$\begin{aligned} |B(\xi)| &\leq c \int_{\frac{\varepsilon}{2} < \rho(\xi,\eta) \leq \varepsilon,} \frac{1}{\varepsilon} \cdot \frac{1}{d(\xi,\eta)^Q} |u(\eta)|\, d\eta \\ &\leq \frac{c}{\varepsilon^{Q+1}} \int_{\rho(\xi,\eta)\leq\varepsilon} |u(\eta)|\, d\eta \\ &\leq \frac{c}{\varepsilon} \cdot \sup_{r\leq\varepsilon} \frac{1}{\left|\widetilde{B}(\xi, r)\right|} \int_{\widetilde{B}(\xi,r)} |u(\eta)|\, d\eta \\ &\leq \frac{c}{\varepsilon} \cdot \mathcal{M}_{\widetilde{B}\left(\overline{\xi}, r_1\right), \widetilde{B}\left(\overline{\xi}, 11r_1\right)} u(\xi) \end{aligned}$$

provided $\varepsilon < 4r_1$. As before we have

$$\|B\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)} \leq \frac{c}{\varepsilon} \|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)}.$$

Finally,

$$\begin{aligned} |C(\xi)| &\leq c \int_{\frac{\varepsilon}{2} < \rho(\xi,\eta) \leq \varepsilon} \frac{1}{\varepsilon^2} \cdot \frac{1}{d(\xi,\eta)^{Q-1}} |u(\eta)|\, \eta dy \\ &\leq \frac{c}{\varepsilon^{Q+1}} \int_{\rho(\xi,\eta)\leq\varepsilon} |u(\eta)|\, d\eta \end{aligned}$$

as for the term $B(\xi)$, therefore

$$\|I\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)} \leq \frac{c}{\varepsilon} \|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)}.$$

Let us bound $II$:

$$|II(\xi)| \leq c \int_{\rho(\xi,\eta)\leq\varepsilon} \frac{|Hu(\eta)|}{\rho(\xi,\eta)^{Q-1}} d\eta.$$

Then, a computation similar to that of $C(\xi)$ gives

$$|II(\xi)| \leq c\varepsilon \mathcal{M}_{\widetilde{B}\left(\overline{\xi}, r_1\right), \widetilde{B}\left(\overline{\xi}, 11r_1\right)} u(\xi)$$

and

$$\|II\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)} \leq c\varepsilon \|u\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_1\right)\right)},$$

provided $\varepsilon < 4r_1$. So we are done. ■

**Theorem 7.5** *For any $u \in S^{2,p}_{\widetilde{X}}(\widetilde{B}(\overline{\xi}, r))$, $p \in [1, \infty)$, $0 < r < r_1$ (where $r_1$ is the number in Theorem 7.4), define the following quantities:*

$$\Phi_k = \sup_{1/2<\sigma<1} \left( (1-\sigma)^k r^k \left\| \widetilde{D}^k u \right\|_{L^p\left(\widetilde{B}_{r\sigma}\right)} \right) \quad \textit{for } k = 0, 1, 2.$$

*Then for any $\delta > 0$ (small enough)*

$$\Phi_1 \leq \delta\, \Phi_2 + \frac{c}{\delta} \Phi_0.$$

**Proof.** This result follows exactly as in [2, Thm. 21] exploiting the interpolation result for compactly supported functions (Theorem 7.4), cutoff functions (Lemma 7.3) and Proposition 3.32. ■

We can now come to the

**Proof of Theorem 7.2.** This proof is similar to that of theorem [2, Thm. 3]. Due to the different context, we include a complete proof for convenience of the reader.

Pick $r_0 = \min(R_0, r_1)$ where $R_0, r_1$ are the numbers appearing in Theorems 7.1, 7.4, respectively. For $r \leq r_0$, let $u \in S^{2,p}_{\widetilde{X}}\left(\widetilde{B}(\overline{\xi}, r)\right)$. Let $\varphi$ be a cutoff function as in lemma 7.3,

$$\widetilde{B}\left(\overline{\xi}, \sigma r\right) \prec \varphi \prec \widetilde{B}\left(\overline{\xi}, \sigma' r\right).$$

By Theorem 7.1, $\varphi u \in S^{2,p}_{\widetilde{X},0}\left(\widetilde{B}(\overline{\xi}, r)\right)$; then, by density, we can apply Theorem 7.1 to $\varphi u$:

$$\|\varphi u\|_{S^{2,p}\left(\widetilde{B}(\overline{\xi},r)\right)} \leq c \left\{ \left\| \widetilde{\mathcal{L}}(\varphi u) \right\|_{L^p\left(\widetilde{B}(\overline{\xi},r)\right)} + \|\varphi u\|_{L^p\left(\widetilde{B}(\overline{\xi},r)\right)} \right\}.$$

For $1 \leq i, j \leq q$, from the above inequality we get

$$\begin{aligned}
\left\| \widetilde{X}_i \widetilde{X}_j u \right\|_{L^p(\widetilde{B}_{\sigma r})} &\leq c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \|u\|_{L^p(\widetilde{B}_{\sigma' r})} + \right. \\
&\quad \left. + \frac{1}{(1-\sigma) r} \left\| \widetilde{D} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \frac{1}{(1-\sigma)^2 r^2} \|u\|_{L^p(\widetilde{B}_{\sigma' r})} \right\} \\
&\leq c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \frac{1}{(1-\sigma) r} \left\| \widetilde{D} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} \right. \\
&\quad \left. + \frac{1}{(1-\sigma)^2 r^2} \|u\|_{L^p(\widetilde{B}_{\sigma' r})} \right\}
\end{aligned}$$

where, as before, we let

$$\left\| \widetilde{D} u \right\| \equiv \sum_{i=1}^{q} \left\| \widetilde{X}_i u \right\| \text{ and } \left\| \widetilde{D}^2 u \right\| \equiv \sum_{i,j=1}^{q} \left\| \widetilde{X}_i \widetilde{X}_i u \right\| + \left\| \widetilde{X}_0 u \right\|.$$

Multiplying both sides for $(1-\sigma)^2 r^2$ we get

$$(1-\sigma)^2 r^2 \parallel \widetilde{X}_i \widetilde{X}_j u \parallel_{L^p(\widetilde{B}_{\sigma r})} \leq c \left\{ (1-\sigma)^2 r^2 \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \right. \tag{7.11}$$

$$\left. +(1-\sigma) r \left( \left\| \widetilde{D} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} \right) + \|u\|_{L^p(\widetilde{B}_{\sigma' r})} \right\}.$$

Next, we compute $(1-\sigma)^2 r^2 \left\| \widetilde{X}_0 u \right\|_{L^p(\widetilde{B}_{\sigma r})}$:

$$(1-\sigma)^2 r^2 \left\| \widetilde{X}_0 u \right\|_{L^p(\widetilde{B}_{\sigma r})} = (1-\sigma)^2 r^2 \left\| \widetilde{\mathcal{L}} u - \sum_{i,j=1}^{q} \widetilde{a}_{ij} \widetilde{X}_i \widetilde{X}_j u \right\|_{L^p(\widetilde{B}_{\sigma r})} \tag{7.12}$$

$$\leq c(1-\sigma)^2 r^2 \left( \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_{\sigma r})} + \left\| \widetilde{X}_i \widetilde{X}_j u \right\|_{L^p(\widetilde{B}_{\sigma r})} \right).$$

Combining (7.11) and (7.12), we have

$$(1-\sigma)^2 r^2 \left\| \widetilde{D}^2 u \right\|_{L^p(\widetilde{B}_{\sigma r})} \leq c \left\{ (1-\sigma)^2 r^2 \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \right. \tag{7.13}$$

$$\left. +(1-\sigma) r \left\| \widetilde{D} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \|u\|_{L^p(\widetilde{B}_{\sigma' r})} \right\}.$$

Adding $(1-\sigma) r \|Du\|_{L^p(\widetilde{B}_{\sigma r})}$ to both sides of (7.13),

$$(1-\sigma) r \left\| \widetilde{D} u \right\|_{L^p(\widetilde{B}_{\sigma r})} + (1-\sigma)^2 r^2 \left\| \widetilde{D}^2 u \right\|_{L^p(\widetilde{B}_{\sigma r})} \tag{7.14}$$

$$\leq c \left\{ (1-\sigma)^2 r^2 \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + (1-\sigma) r \left\| \widetilde{D} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \|u\|_{L^p(\widetilde{B}_{\sigma' r})} \right\},$$

by Theorem 7.5,

$$\leq c \left\{ (1-\sigma)^2 r^2 \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_{\sigma' r})} + \left( \delta \Phi_2 + \frac{c}{\delta} \Phi_0 \right) + \|u\|_{L^p(\widetilde{B}_{\sigma' r})} \right\}.$$

Choosing $\delta$ small enough, we have

$$\Phi_2 + \Phi_1 \leq c \left\{ r^2 \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}_r)} + \|u\|_{L^p(\widetilde{B}_r)} \right\},$$

then

$$r^2 \left\| \widetilde{D}^2 u \right\|_{L^p(\widetilde{B}(\overline{\xi}, r/2))} + r \left\| \widetilde{D} u \right\|_{L^p(\widetilde{B}(\overline{\xi}, r/2))} \leq c \left\{ r^2 \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}(\overline{\xi}, r))} + \|u\|_{L^p(\widetilde{B}(\overline{\xi}, r))} \right\},$$

hence

$$\|u\|_{S^{2,p}_{\widetilde{X}}(\widetilde{B}(\overline{\xi}, r/2))} \leq c \left\{ \left\| \widetilde{\mathcal{L}} u \right\|_{L^p(\widetilde{B}(\overline{\xi}, r))} + \|u\|_{L^p(\widetilde{B}(\overline{\xi}, r))} \right\},$$

which is the desired result. ■

## 7.3 $L^p$ estimates in the original variables

Letś now prove Theorem 2.2, which follows from Theorem 7.2 in a way which is analogous to that followed in § 6.3 to prove Schauder estimates.

Fix $\overline{x} \in \Omega' \Subset \Omega$ and $R$ such that in $B(\overline{x}, R) \subset \Omega$ all the construction of the previous two subsections (lifting to $\widetilde{B}\left(\overline{\xi}, R\right)$ and so on) can be performed. Let $r_0 < R$ as in Theorem 7.2, and let $u \in S_X^{2,p}(B(\overline{x}, r_0))$. By Theorem 3.33 we know that the function $\widetilde{u}(x,h) = u(x)$ belongs to $S_{\widetilde{X}}^{2,p}\left(B\left(\overline{\xi}, r_0\right)\right)$, so we can apply to $\widetilde{u}$ the $L^p$ estimates contained in Theorem 7.2. Combining this fact with the two estimates in Theorem 3.33 we get

$$\begin{aligned}
\|u\|_{S_X^{2,\alpha}(B(\overline{x},\delta_0 r_0/2))} &\le c\,\|\widetilde{u}\|_{S_{\widetilde{X}}^{2,\alpha}\left(\widetilde{B}\left(\overline{\xi}, r_0/2\right)\right)} \\
&\le c\left(\left\|\widetilde{\mathcal{L}}\widetilde{u}\right\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_0\right)\right)} + \|\widetilde{u}\|_{L^p\left(\widetilde{B}\left(\overline{\xi}, r_0\right)\right)}\right) \\
&\le c\left(\|\mathcal{L}u\|_{L^p(B(\overline{x},r_0))} + \|u\|_{L^p(B(\overline{x},r_0))}\right)
\end{aligned}$$

since $\widetilde{\mathcal{L}}\widetilde{u} = \widetilde{(\mathcal{L}u)}$.

Next, let us choose a family of balls $B(x_i, r_0)$ in $\Omega$ such that

$$\Omega' \subset \bigcup_{i=1}^{k} B(x_i, \delta_0 r_0/2) \subset \bigcup_{i=1}^{k} B(x_i, r_0) \subset \Omega.$$

Therefore

$$\begin{aligned}
\|u\|_{S_X^{2,p}(\Omega')} \le \|u\|_{S_X^{2,p}(\cup B(x_i,\delta_0 r_0/2))} &\le \sum_{i=1}^{k} \|u\|_{S_X^{2,p}(B(x_i,\delta_0 r_0/2))} \\
&\le c\sum_{i=1}^{k}\left\{\|\mathcal{L}u\|_{L^p(B(x_i,r_0))} + \|u\|_{L^p(B(x_i,r_0))}\right\} \\
&\le c\left\{\|\mathcal{L}u\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)}\right\}
\end{aligned}$$

with $c$ also depending on $r_0$.

# References


[1] M. Bramanti: Singular integrals in nonhomogeneous spaces: $L^2$and $L^p$continuity from Hölder estimates. Revista Matematica Iberoamericana 26 (2010), no. 1, 347-366.

[2] M. Bramanti, L. Brandolini: $L^p$-estimates for uniformly hypoelliptic operators with discontinuous coeffi cients on homogeneous groups. Rend. Sem. Mat. dellUniv. e del Politec. di Torino. Vol. 58, 4 (2000), 389-433.

[3] M. Bramanti-L. Brandolini: $L^p$-estimates for nonvariational hypoelliptic operators with VMO coeffi cients. Trans. Amer. Math. Soc. 352 (2000), no. 2, 781-822.

[4] M. Bramanti, L. Brandolini: Schauder estimates for parabolic nondivergence operator of Hörmander type. J. Differential Equations 234 (2007) 177-245.

[5] M. Bramanti, L. Brandolini, M. Pedroni: Basic properties of nonsmooth Hörmanderś vector fields and Poincaréś inequality. Submitted preprint, 2008. arXiv:0809.2872

[6] M. Bramanti, L. Brandolini, M. Pedroni: On the lifting and approximation theorem for nonsmooth vector fields. To appear on Indiana University Mathematics Journal.

[7] M. Bramanti, M. C. Cerutti, M. Manfredini: $L^p$-estimates for some ultraparabolic operators with discontinuous coeffi cients. Journal of Math. Anal. and Appl., 200, 332-354 (1996).

[8] M. Bramanti, M. Zhu: Local real analysis in locally doubling spaces. Submitted Preprint, 2011.

[9] A. P. Calderón-A. Zygmund: Singular integral operators and differential equations, Am. Journ. of Math., 79 (1957), 901-921.

[10] S. Campanato, Proprietà di Hölderianità di alcune classi di funzioni, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) 17 (1963) 175–188.

[11] F. Chiarenza, M. Frasca, P. Longo: Interior $W^{2,p}$-estimates for nondivergence elliptic equations with discontinuous coeffi cients. Ricerche di Mat. XL (1991), 149-168.

[12] F. Chiarenza, M. Frasca, P. Longo: $W^{2,p}$-solvability of the Dirichlet problem for non divergence elliptic equations with VMO coeffi cients. Trans. of Am. Math. Soc., 336 (1993), n. 1, 841-853.

[13] R. Coifman, R. Rochberg, G. Weiss: Factorization theorems for Hardy spaces in several variables. Annals of Math., 103 (1976) 611-635.

[14] R. Coifman, G. Weiss: Analyse harmonique non-commutative sur certains espaces homogènes. Lecture Notes in Mathematics, n. 242. Springer-Verlag, Berlin-Heidelberg-New York, 1971.

[15] M. Di Francesco, S. Polidoro: Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. Adv. Differential Equations 11 (2006), no. 11, 1261–1320.

[16] G. B. Folland: Subelliptic estimates and function spaces on nilpotent Lie groups, Arkiv för Math. 13, (1975), 161-207.

[17] G. B. Folland, E. M. Stein: Estimates for the $\overline{\partial}_b$complex and analysis on the Heisenberg group, Comm. on Pure and Appl. Math. XXVII (1974), 429-522.

[18] B. Franchi, E. Lanconelli: Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coeffi cients. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 10 (1983), no. 4, 523–541.

[19] C. E. Gutiérrez, E. Lanconelli: Schauder estimates for sub-elliptic equations. J. Evol. Equ. 9 (2009), no. 4, 707–726.

[20] L. Hörmander: Hypoelliptic second order differential equations. Acta Mathematica, 119 (1967), 147-171.

[21] E. Lanconelli, A. Pascucci, S. Polidoro: Linear and nonlinear ultraparabolic equations of Kolmogorov type arising in diffusion theory and in finance. Nonlinear problems in mathematical physics and related topics, II, 243–265, Int. Math. Ser. (N. Y.), 2, Kluwer/Plenum, New York, 2002.

[22] E. Lanconelli, S. Polidoro: On a class of hypoelliptic evolution operators. Partial differential equations, II (Turin, 1993). Rend. Sem. Mat. Univ. Politec. Torino 52 (1994), no. 1, 29–63.

[23] A. Nagel, E. M. Stein, S. Wainger: Balls and metrics defined by vector fields I: Basic properties. Acta Math., 155 (1985), 130-147.

[24] F. Nazarov, S. Treil, A. Volberg: The $Tb$-theorem on non-homogeneous spaces. Acta Math. 190 (2003), no. 2, 151–239.

[25] L. P. Rothschild-E. M. Stein: Hypoelliptic differential operators and nilpotent groups. Acta Math., 137 (1976), 247-320.

[26] A. Sanchez-Calle: Fundamental solutions and geometry of sum of squares of vector fields. Inv. Math., 78 (1984), 143-160.

[27] E. M. Stein: Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals. Princeton Univ. Press. Princeton, New Jersey, 1993.

[28] X. Tolsa: A proof of the weak $(1,1)$-inequality for singular integrals with non doubling measures based on a Calderón-Zygmund decomposition. Publ. Mat. 45 (2001), no. 1, 163–174.

Marco Bramanti
Dipartimento di Matematica, Politecnico di Milano
Via Bonardi 9. 20133 Milano. ITALY
marco.bramanti@polimi.it

Maochun Zhu
Department of Applied Mathematics, Northwestern Polytechincal University
127 West Youyi Road. 710072 Xiàn. P. R. CHINA
zhumaochun2006@126.com