Contents of Volume 142, Number 2

STUDIA MATHEMATICA

The journal publishes original papers in English, French, German and Russian, mainly in functional analysis, abstract methods of mathematical analysis and probability theory. Usually 3 issues constitute a volume.

Detailed information for authors is given on the inside back cover. Manuscripts and correspondence concerning editorial work should be addressed to

STUDIA MATHEMATICA
Śniadeckich 8, P.O. Box 137, 00-950 Warszawa, Poland, fax 48-22-6293997
E-mail: studia@impan.gov.pl

Subscription information (2000): Vols. 138–143 (18 issues); $33.50 per issue.

Correspondence concerning subscription, exchange and back numbers should be addressed to

Institute of Mathematics, Polish Academy of Sciences
Publications Department
Śniadeckich 8, P.O. Box 137, 00-950 Warszawa, Poland, fax 48-22-6293997
E-mail: publ@impan.gov.pl

Published by the Institute of Mathematics, Polish Academy of Sciences
Typeset using TEX at the Institute
Printed and bound by

drukarnia
herman & herman
SPÓŁKA CYWILNA
02-240 WARSZAWA UL. JAKOBINÓW 23
tel. (0-22) 846-05-10, 28, 66; tel/fax: (0-22) 849-06-40

PRINTED IN POLAND

ISSN 0039-3223

High order representation formulas and embedding theorems on stratified groups and generalizations

by

GUOZHEN LU (Detroit, MI)
and RICHARD L. WHEEDEN (New Brunswick, NJ)

Abstract. We derive various integral representation formulas for a function minus a polynomial in terms of vector field gradients of the function of appropriately high order. Our results hold in the general setting of metric spaces, including those associated with Carnot–Carathéodory vector fields, under the assumption that a suitable L^1 to L^1 Poincaré inequality holds. Of particular interest are the representation formulas in Euclidean space and stratified groups, where polynomials exist and L^1 to L^1 Poincaré inequalities involving high order derivatives are known to hold. We apply the formulas to derive embedding theorems and potential type inequalities involving high order derivatives.

1. Introduction. The main goal of this article is to prove the existence of representation formulas for functions as (fractional) integral transforms of their high order vector field gradients. We prove the formulas assuming there is a suitable L^1 to L^1 Poincaré inequality for two doubling measures. As special examples, we obtain the existence of such formulas in both Euclidean spaces and stratified groups, where Poincaré inequalities are known to hold for several choices of polynomials. We also give some applications of the representation formulas to various estimates of potential type.

It is well known that the following pointwise estimate holds for a smooth, real-valued function $f(x)$ defined on a ball B in N-dimensional Euclidean space $\mathbb{R}^N$:

$$|f(x) - f_B| \le C \int_B \frac{|\nabla f(y)|}{|x-y|^{N-1}}\, dy, \qquad x \in B,$$

2000 *Mathematics Subject Classification*: Primary 46E35, 31B10.

Key words and phrases: representation formulas, Poincaré inequalities, embedding theorems, polynomials, stratified groups, vector fields, doubling measures.

The authors were partially supported by NSF Grants DMS96-22996 (G. Z. Lu) and 95-00799 (R. L. Wheeden), and by NATO CR Grant 972144 (G. Z. Lu and R. L. Wheeden).

where ∇f denotes the gradient of f, f_B is the average $|B|^{-1}\int_B f(y)\,dy$, $|B|$ is Lebesgue measure of B, and C is a constant which is independent of f, x, B. A proof of this estimate can be found for example in [GT, Lemma 7.16].

In recent years, various analogues of this formula for more general systems of first order vector fields $Xf = (X_1 f, \ldots, X_k f)$ have been derived in the form

$$(1.1)\qquad |f(x) - f_B| \le C \int_{\tau B} |Xf(y)| \frac{\varrho(x,y)}{|B(x, \varrho(x,y))|}\,dy$$

for $x \in B$ and $\tau > 1$, where B is a ball with respect to a metric ϱ which is naturally associated with the vector fields, $B(x,r)$ is the ball with center x and radius r, and τB is the ball concentric with B of radius τ times that of B. As is customary, we shall refer to such estimates as *first order representation formulas.*

Formula (1.1) on a stratified group $\mathbb{G}$ was proved in [L1] for the left invariant vector fields $X_1, \ldots, X_k$ which generate the Lie algebra of $\mathbb{G}$. In $\mathbb{R}^N$, (1.1) was first derived for general Hörmander vector fields in [FLW1]. In this case, $\varrho(x,y)$ denotes the associated metric (see e.g. [FP]). Alternate proofs of (1.1) for Hörmander vector fields are given in [CDG] and [FLW2]. Such formulas were derived in [F] and [FGW] for nonsmooth Grushin vector fields of the types in [FL], [F] and [FGW]. In [FLW2] and [FW], it was shown that an L^1 to L^1 Poincaré inequality implies (1.1) in fairly general spaces of homogeneous type; the reverse doubling condition required in [FLW2] was relaxed in [FW] for spaces in which geodesics exist. A first order representation formula has also been obtained in [LW1] for the product of multiple spaces when each component space is associated with a family of vector fields of the types above. More recently, in [LW2], it was shown that (1.1) holds with $\tau = 1$ for Carnot–Carathéodory vector fields, again assuming an L^1 to L^1 Poincaré estimate, and it was also shown that the ball B can be replaced by any Boman chain domain Ω. Similar first order estimates have also been derived recently in [HK].

In this paper, we study formulas involving high order vector field gradients (see Theorems A, B and C). As particular examples, we obtain analogues of (1.1) in Euclidean spaces and stratified groups with f_B replaced by appropriate polynomials and with the first order derivatives replaced by higher order derivatives (see Corollary E). More precisely, on a stratified group $\mathbb{G}$ with generators $X_1, \ldots, X_k$, we will show that for any positive integer m, any smooth function f and any ball B, there exists a polynomial $P_m(B,f)$ of degree less than m such that

$$(1.2)\qquad \begin{aligned} &|f(x) - P_m(B,f)(x)| \\ &\quad\le C \int_B |X^m f(y)| \frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy + C\frac{r(B)^m}{|B|}\int_B |X^m f(y)|\,dy \end{aligned}$$

for $x \in B$, where $r(B)$ is the radius of B and $|X^m f| = (\sum_{|\alpha|=m} |X^\alpha f|^2)^{1/2}$ with $X^\alpha = X_{i_1}^{\alpha_1} \ldots X_{i_k}^{\alpha_k}$ for nonnegative integers α_i, $1 \le i_j \le k$ for $1 \le j \le k$, and $|\alpha| = \alpha_1 + \ldots + \alpha_k$. We will show that such a formula holds for various choices of polynomials in Euclidean spaces and stratified groups. Moreover, if Q is the homogeneous dimension of $\mathbb{G}$ (see the definition below) and $0 < m \le Q$, then we have the more refined formula

$$(1.3)\qquad |f(x) - P_m(B,f)(x)| \le C \int_B |X^m f(y)| \frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy, \qquad x \in B.$$

In Corollary F, we will show that (1.3) remains valid when $m \le Q$ if B is replaced by a Boman chain domain Ω (or even a "weak Boman chain domain") in $\mathbb{G}$.

If $X_i = \partial/\partial x_i$, $1 \le i \le N$, then $\mathbb{R}^N$ is identical to the corresponding group $\mathbb{G}$ and has homogeneous dimension $Q = N$. Thus $\mathbb{R}^N$ can be regarded as a special case of the group setting. In this special case, by a result of S. L. Sobolev (see [AH], p. 215–217), inequality (1.3) holds for all m, i.e., (1.2) holds without the second term on the right for all m, provided $P_m(B,f)$ is a particular polynomial related to the Taylor expansion of f. A similar formula as well as some others are derived in [BH] (initially for cubes instead of balls, but also for sets that are starlike with respect to a ball), and (1.3) is even shown to hold everywhere, without an exceptional set of measure zero, for suitable Sobolev functions. On the other hand, as we shall see even in the usual Euclidean case, formula (1.2) has some flexibility in regard to the choice of a polynomial since it is valid for any polynomial for which there is an L^1 to L^1 Poincaré inequality for f. There are different possible choices of such polynomials in Euclidean spaces and stratified groups. It is the presence of the second term on the right side of (1.2) which allows our method to work for arbitrarily large m, and we do not know if (1.2) is generally valid without this term when m is large. The extra term causes no problem in the applications we consider.

Adding the second term on the right side in (1.2) also allows us to derive analogous formulas in any metric space of homogeneous type without assuming that the underlying measure satisfies the reverse doubling condition of order m (see the definition below). We only need to assume that a suitable L^1 to L^1 Poincaré inequality holds and that geodesics exist. Removing the assumption of reverse doubling of order m is important when we deal with representation formulas involving higher order vector field gradients because even Lebesgue measure in $\mathbb{R}^N$ fails to satisfy reverse doubling of order m

when $m > N$. Another advantage of dropping the reverse doubling hypothesis (H2) is that we can then define high order Sobolev spaces in metric spaces by only assuming that the underlying measure μ is doubling. This has been recently shown in [LLW] and generalizes some of the results in [H], [FLW2] and [FHK] for first order Sobolev spaces. It is fairly natural to assume the reverse doubling condition of order 1, since, by [FW], in those metric spaces that correspond to Lipschitz continuous Carnot–Carathéodory vector fields, Lebesgue measure automatically satisfies the reverse doubling condition of order 1. However, even when $m = 1$, the formula holds in a metric space of homogeneous type without assuming the reverse doubling condition of any specific order (see Theorem A*).

A basic ingredient in proving the representation formulas is the existence of an appropriate chain of balls. If the space has geodesics, a suitable chain was constructed in [FW] and used to relax the reverse doubling assumption made in [FLW2] from order $1+\varepsilon$ to order 1. No geodesic property was assumed in [FLW2]. The chain we will use below again requires the geodesic (segment) property, but it has extra properties compared to the chain in [FW]; in particular, it lies entirely in the particular ball B under consideration, rather than lying in an enlarged ball τB, $\tau > 1$. This allows proving the representation formula without having to integrate over an enlarged ball on the right side there. A chain of balls with largely the same properties was discovered independently in [HK] and used to derive the improved version (i.e., the version without an enlarged ball) of the formula in [FW], as in [LW2]. The extra properties of the chain also lead to improvements of some of the results in [LW2], namely, to relaxation of the Boman chain restriction to a weaker one. We state the weaker restrictions in §2 after recalling the definition of the usual Boman chain condition (see Definition 2.3 in §2).

As one application of (1.2) and (1.3), we will derive weighted Poincaré inequalities of high order by combining with the results in [SW] and assuming a balance condition similar to the one in [CW]. We will also derive Moser–Trudinger exponential estimates, Lipschitz estimates and embedding theorems on Campanato–Morrey spaces in terms of higher order derivatives. As an application of the representation formula for weak Boman domains, we can derive exponential and Lipschitz estimates for such domains directly, rather than deducing them from ones for balls.

The organization of the paper is as follows. In Section 2, we give some preliminaries, including a discussion of the necessary properties of polynomials, and state the representation theorems. In Section 3, we construct the chain of balls used to prove the representation formulas. Section 4 contains the proofs of the representation theorems. Some applications of these formulas are given in §5.

2. Preliminaries and statements of representation formulas in metric spaces. Let $(\mathcal{S},\varrho)$ be a metric space with metric ϱ, so that for all $x,y,z\in\mathcal{S}$, ϱ satisfies $\varrho(x,y)\le\varrho(x,z)+\varrho(z,y)$. A measure μ is called a *doubling measure* if there is a constant A such that

$$\mu(B(x,2r))\le A\,\mu(B(x,r)),\qquad x\in\mathcal{S},\ r>0,$$

where by definition $B(x,r)=\{y\in\mathcal{S}:\varrho(x,y)<r\}$, and $\mu(B(x,r))$ denotes the μ-measure of $B(x,r)$. As usual, we refer to $B(x,r)$ as the ball with center x and radius r, and if B is a ball, we write x_B for its center, $r(B)$ for its radius and cB for the ball of radius $cr(B)$ having the same center as B. We always assume that $(\mathcal{S},\varrho)$ is locally compact.

Before we proceed, we need to define the notion of polynomial functions. Let m be a positive integer, Ω be a domain in $(\mathcal{S},\varrho)$, and ν,μ be doubling measures. The two main properties that we require of a polynomial function $P(x)$, $x\in\Omega$, are:

(P1) There is a constant $C_1>0$ such that for every metric ball $D\subset\Omega$,

$$\operatorname*{ess\,sup}_{x\in D}|P(x)|\le\frac{C_1}{\nu(D)}\int_D|P(y)|\,d\nu(y),$$

where the essential supremum is taken with respect to ν.

(P2) If D is a metric ball in Ω and E is a subball of D with $\nu(E)>\gamma\nu(D)$, $\gamma>0$, then

$$\|P\|_{L^\infty_\nu(E)}\ge C_2(\gamma)\|P\|_{L^\infty_\nu(D)}.$$

If $P(x)$ is an ordinary polynomial of degree $m-1$ in Euclidean space, with the usual Euclidean distance, and if ν,μ are both Lebesgue measure, then (P1) and (P2) of course hold with constants $C_1, C_2(\gamma)$ depending additionally only on m. For general $(\mathcal{S},\varrho)$, the role of the degree will be played by the order of the L^1 to L^1 Poincaré inequality that we assume is valid. More precisely, given m and Ω, let f be a function on Ω for which there are functions g and $P_m(B,f)$ such that the L^1 to L^1 Poincaré inequality

$$\frac{1}{\nu(B)}\int_B|f(x)-P_m(B,f)(x)|\,d\nu(x)\le C\frac{r(B)^m}{\mu(B)}\int_B|g(x)|\,d\mu(x)$$

holds for every ball $B\subset\Omega$, with C and g independent of B. The functions $P_m(B,f)$ may depend on ν, μ and g in addition to m, B and f.

Definition 2.1. Given m, f and Ω, we say that functions $P_m(f,B)$ are polynomial functions *associated* with m, balls $B\subset\Omega$, f,g,ν and μ if the Poincaré inequality above holds and if (P1), (P2) hold with $P=P_m(B,f)$ and also with $P=P_m(B_1,f)-P_m(B_2,f)$ for every $B\subset\Omega$ and $B_1\subset B_2\subset\Omega$, with constants $C, C_1, C_2(\gamma)$ that are uniform in B, B_1, B_2.

We denote such polynomials by $P_m(B,f,g,\nu,\mu)$ but usually write them simply as $P_m(B,f)$. In practice, the constants $C_1, C_2(\gamma)$ are also independent of f, but we do not need to make this assumption.

REMARK 2.2. For a stratified group $\mathcal{S}=\mathbb{G}$, polynomials $P_m(B,f)$ with these properties exist, and in fact we have several choices of them: see the comments made before the statements of Corollaries D and F in this section.

DEFINITION 2.3. A domain (i.e., an open connected set) Ω in $\mathcal{S}$ is said to satisfy the *Boman chain condition* of type σ, M, or to be a member of $\mathcal{F}(\sigma,M)$, if there exist constants $\sigma>1$, $M>0$, and a family $\mathcal{F}$ of metric balls $B\subset\Omega$ such that

(i) $\Omega=\bigcup_{B\in\mathcal{F}}B$.
(ii) $\sum_{B\in\mathcal{F}}\chi_{\sigma B}(x)\le M\,\chi_\Omega(x)$ for all $x\in\mathcal{S}$.
(iii) There is a "central ball" $B_0\in\mathcal{F}$ such that for each ball $B\in\mathcal{F}$, there is a positive integer $k=k(B)$ and a chain $\{B_j\}_{j=0}^k$ of balls for which $B_k=B$ and each $B_j\cap B_{j+1}$ contains a ball D_j with $B_j\cup B_{j+1}\subset MD_j$.
(iv) $B\subset MB_j$ for all $j=0,\dots,k(B)$.

If we replace the hypothesis that $\sigma>1$ by $\sigma=1$, we say that Ω satisfies the *weak Boman chain condition.*

In fact, we will only need to assume that the weak Boman chain condition holds in order to prove representation formulas on domains other than balls. This weakens the requirements made in [LW2] in case $m=1$.

The following four hypotheses will enter the results below, but not all four are needed in every theorem. As always, $(\mathcal{S},\varrho)$ is a metric space. Let μ and ν be doubling measures with respect to metric balls, and let Ω be a domain in $\mathcal{S}$.

(H1) f is a function as in Definition 2.1.

(H2) The measure μ in Definition 2.1 satisfies a *reverse doubling condition* of order m, i.e., there is a constant $C>0$ such that if B and $\widetilde{B}$ are balls with centers in Ω and with $B\subset\widetilde{B}$, then

$$\mu(\widetilde{B})\ge C\left(\frac{r(\widetilde{B})}{r(B)}\right)^m\mu(B).$$

(H3) $(\mathcal{S},\varrho)$ has the *segment* (or *geodesic*) *property*, i.e., for all $x,y\in\mathcal{S}$, there is a continuous curve γ connecting x and y such that $\varrho(\gamma(t),\gamma(s))=|t-s|$.

It is easy to see that if the segment property holds, then for every ball $B\subset\Omega$ and all $x\in B$, there is a continuous curve $\gamma=\gamma_{x_Bx}(t)$, $0\le t\le1$, in B with $\gamma(0)=x_B$, $\gamma(1)=x$ and $\varrho(x_B,z)=\varrho(x_B,y)+\varrho(y,z)$ for all $y,z\in\gamma$ with $y=\gamma(s)$, $z=\gamma(t)$, $0\le s\le t\le1$. The fact that the curve $\gamma=\gamma_{x_Bx}$ lies in B is a corollary of the assumed additivity of ϱ along γ. To see this, we choose $z=x$ in the additivity statement; then any $y\in\gamma$ also lies in B since

$$r(B)>\varrho(x_B,x)=\varrho(x_B,y)+\varrho(y,x)\ge\varrho(x_B,y).$$

(H4) Ω is a weak Boman chain domain.

Let us now state our main results.

THEOREM A. *Let ν, μ be doubling measures on a metric space $(\mathcal{S},\varrho)$. Let B_0 be a ball and suppose that* (H1) *and* (H3) *hold with $\Omega=B_0$. Then for ν-a.e. $x\in B_0$,*

$$|f(x)-P_m(B_0,f)(x)|\le C\int_{B_0}|g(y)|\frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\,d\mu(y)+C\frac{r(B_0)^m}{\mu(B_0)}\int_{B_0}|g|\,d\mu$$

where C depends only on ν, μ and the constants in (H1).

In the special case $m=1$, the polynomial functions are just constants, and then Theorem A improves the results in [FLW2] and [FW] by dropping the reverse doubling assumption made there, although we need to have the second term on the right and geodesics. We state this result separately.

THEOREM A*. *Let ν, μ be doubling measures on a metric space $(\mathcal{S},\varrho)$. Let B_0 be a ball and suppose that the Poincaré inequality*

$$\frac{1}{\nu(B)}\int_B|f-f_B|\,d\nu\le C\frac{r(B)}{\mu(B)}\int_B|g|\,d\mu$$

holds for every ball $B\subset B_0$, with $f_B=\fint_B f\,d\nu$. Suppose also that (H3) *holds in B_0. Then for ν-a.e. $x\in B_0$,*

$$|f(x)-f_{B_0}|\le C\int_{B_0}|g(y)|\frac{\varrho(x,y)}{\mu(B(x,\varrho(x,y)))}\,d\mu(y)+C\frac{r(B_0)}{\mu(B_0)}\int_{B_0}|g|\,d\mu$$

where C depends only on ν,μ and the constant in the Poincaré inequality.

Other versions of this result (some more refined than others) are given in [FLW1,2], [CDG], [FW], [LW2] and [HK]. The version in [HK], which is very similar to Theorem A* and was derived independently, does not have the second term on the right side, but the average f_{B_0} on the left is replaced by the average over a smaller ball of size comparable to B_0.

By also imposing the reverse doubling condition (H2) in Theorem A, we get

THEOREM B. *Let ν, μ be doubling measures on a metric space $(\mathcal{S},\varrho)$. If B_0 is a ball and* (H1), (H2) *and* (H3) *hold with $\Omega=B_0$, then for ν-a.e.*

$x \in B_0$,

$$|f(x) - P_m(B_0, f)(x)| \le C \int_{B_0} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x, \varrho(x,y)))}\, d\mu(y),$$

where C depends only on ν, μ and the constants in (H1), (H2).

The next result extends Theorem B to any weak Boman chain domain Ω.

THEOREM C. *Suppose that ν and μ are doubling measures on a metric space $(\mathcal{S}, \varrho)$ and that hypotheses* (H1)–(H4) *hold for a domain $\Omega \subset \mathcal{S}$. Then for ν-a.e. $x \in \Omega$,*

$$|f(x) - P_m(B_0, f)(x)| \le C \int_{\Omega} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x, \varrho(x,y)))}\, d\mu(y),$$

where B_0 is the central ball in Ω, $P_m(B_0, f)$ is the polynomial associated with m, B_0, f, g, ν and μ of (H1), *and C depends only on ν, μ and the constants of* (H1), (H2) *and* (H4).

Under the segment hypothesis (H3), any metric ball is a Boman domain (see [FGW], [L2]), and thus Theorem B is a special case of Theorem C.

The proof of Theorem A relies on the construction of a chain of metric balls, assuming the segment hypothesis (H3). We state this construction separately as

THEOREM D. *Let $(\mathcal{S}, \varrho)$ be a metric space in which the segment property* (H3) *holds. Let B_0 be a ball in $\mathcal{S}$. For given $x \in B_0$, there exists a chain $\{B_k\}_{k \ge 1}$ of balls with the following properties*:

(1) *$B_k \subset B_0$ and $\varrho(B_k, x) \to 0$ as $k \to \infty$.*
(2) *$r(B_1) \approx r(B_0)$ and $r(B_k) \to 0$ as $k \to \infty$.*
(3) *If $y \in B_k$, then $\varrho(y, x) \approx r(B_k)$.*
(4) *$B_k \cap B_{k-1}$ contains a ball S_k with $r(S_k) \approx r(B_k) \approx r(B_{k+1})$.*
(5) *If $j < k$, then $B_k \subset cB_j$.*
(6) *$\{B_k\}_{k \ge 1}$ has bounded overlaps, i.e., $\sum_k \chi_{B_k}(y) \le c$ for all y.*

The constants of equivalence in (2), (3) *and* (4) *and the constants c in* (5) *and* (6) *are independent of x, k, j and B_0, but the chain $\{B_k\}$ depends on x.*

REMARKS. (i) As the proof of Theorem D will show, B_k can be chosen with $r(B_k) \approx 2^{-k} r(B_0)$ and constants of equivalence that are independent of k, x and B_0.

(ii) A chain of balls with similar properties is constructed in [HK], although there are some differences. In [HK], the chain corresponding to x contains only a finite number of balls, ending with a ball which is arbitrarily close to x.

The hypotheses in the theorems above are valid when μ and ν are Lebesgue measure on a stratified group $\mathbb{G}$, such as the Heisenberg group. To state our results in this context, we need to briefly recall some notions about stratified groups.

Let $\mathcal{G}$ be a finite-dimensional, stratified, nilpotent Lie algebra. Assume

$$\mathcal{G} = \bigoplus_{i=1}^{s} V_i,$$

and $[V_i, V_j] \subset V_{i+j}$ for $i + j \le s$, $[V_i, V_j] = 0$ for $i + j > s$. Let $X_1, \ldots, X_k$ be a basis for V_1 and suppose that $X_1, \ldots, X_k$ generate $\mathcal{G}$ as a Lie algebra. Thus we can choose a basis $\{X_{ij}\}$, $1 \le j \le s$, $1 \le i \le k_j$, for V_j consisting of commutators of length j. In particular, $X_{i1} = X_i$, $i = 1, \ldots, k$ and $k = k_1$.

Let $\mathbb{G}$ be the simply connected Lie group associated with $\mathcal{G}$. Since the exponential mapping is a global diffeomorphism from $\mathcal{G}$ to $\mathbb{G}$, for each $g \in \mathbb{G}$, there is $x = (x_{ij}) \in \mathbb{R}^N$, $1 \le i \le k_j$, $1 \le j \le s$, $N = \sum_{j=1}^{s} k_j$, such that

$$g = \exp\Big(\sum x_{ij} X_{ij}\Big).$$

The homogeneous norm function $|\cdot|$ on $\mathbb{G}$ is defined by

$$|g| = \Big(\sum |x_{ij}|^{2s!/j}\Big)^{1/(2s!)}.$$

Also, the integer $Q = \sum_{j=1}^{s} j k_j$ is called the *homogeneous dimension* of $\mathbb{G}$, and Q is usually greater than $\dim \mathbb{G} = N$.

We now define polynomials on $\mathbb{G}$ by following Folland–Stein (see [FS]). Let $X_1, \ldots, X_k$ be the generators of the Lie algebra $\mathcal{G}$, and $X_1, \ldots, X_k, \ldots \ldots, X_N$ be a basis of $\mathcal{G}$. We define $d(X_j) = d_j$ as the length of X_j as a commutator, and arrange the order so that $1 \le d_1 \le \ldots \le d_N$. Thus it is easy to see $d_j = 1$ for $j = 1, \ldots, k$. Let $\xi_1, \ldots, \xi_N$ be the dual basis for $\mathcal{G}^*$, and let $\eta_i = \xi_i \circ \exp^{-1}$. Thus $\eta_1, \ldots, \eta_N$ are a system of global coordinates on $\mathbb{G}$. A function P on $\mathbb{G}$ is called a polynomial on $\mathbb{G}$ if $P \circ \exp$ is a polynomial on $\mathcal{G}$. By this definition, $\eta_1, \ldots, \eta_N$ are polynomials on $\mathbb{G}$ and generate the algebra of polynomials on $\mathbb{G}$. Thus every polynomial on $\mathbb{G}$ can be written uniquely as

$$P = \sum_I a_I \eta^I, \quad \eta^I = \eta_1^{i_1} \ldots \eta_N^{i_N}, \ a_I \in \mathbb{R},$$

where all but finitely many of the coefficients a_I vanish. Clearly, η^I is homogeneous of degree $d(I) = \sum_{j=1}^{N} i_j d(i_j)$. If $P = \sum_I a_I \eta^I$, then we define the *homogeneous degree* (or *order*) of P to be $\max\{d(I) : a_I \ne 0\}$. If we consider $I = (i_1, \ldots, i_k)$, $1 \le i_j \le k$, then $d(I) = |I|$.

It is shown in [FS] that any polynomial on $\mathbb{G}$ satisfies (P1) and (P2) in Definition 2.1 when ν is Lebesgue measure. Moreover, if $\Omega \subset \mathbb{G}$ and

$f \in C^m(\Omega)$, then for any ball $B \subset \Omega$ the Poincaré inequality for Lebesgue measure, i.e.,

$$\|f - P_m(B,f)\|_{L^1(B,dx)} \le Cr(B)^m \|X^m f\|_{L^1(B,dx)}, \tag{2.4}$$

holds for several choices of polynomials $P_m(B,f)$ with homogeneous degree less than m by the results in [L3], [L4]. Here the constant C is independent of f as well as B. One choice is to take $P_m(B,f)$ such that

$$\int_B X^\alpha (f - P_m(B,f))\,dx = 0 \tag{2.5}$$

for all $|\alpha| < m$, where $X^\alpha = X_{i_1}^{\alpha_1} \dots X_{i_k}^{\alpha_k}$, $1 \le i_j \le k$, α_j's are nonnegative integers and $|\alpha| = \alpha_1 + \dots + \alpha_k$. We can also take $P_m(B,f)$ to be a polynomial $\pi_m(B,f)$ which is defined by a projection operator. More precisely, let $\mathcal{P}_m$ be the collection of polynomials on $\mathbb{G}$ of degree less than m. Then there is a linear operator $\pi_m(B,\cdot) : C^m(\Omega) \to \mathcal{P}_m$ for $\Omega \subset \mathbb{G}$ satisfying $\pi_m(B,P) = P$ for all $P \in \mathcal{P}_m$, and

$$\sup_{x\in B} |\pi_m(B,f)(x)| \le \frac{C}{|B|} \int_B |f(y)|\,dy \tag{2.6}$$

and

$$\|X^\alpha \pi_m(B,f)\|_{L^q(B,dx)} \le C(Q,q)\|X^l f\|_{L^q(B,dx)} \tag{2.7}$$

for all $1 \le q \le \infty$, $0 \le |\alpha| = l < \infty$, where $|X^l f| = (\sum_{|\alpha|=l} |X^\alpha f|^2)^{1/2}$. The existence of such polynomials is proved in [L3], [L4]. When $\mathbb{G} = \mathbb{H}^n$, the Heisenberg group, polynomials $P_m(B,f)$ satisfying (2.5) were found earlier in [N].

In [FS], Taylor polynomials are introduced and used to prove a stratified Taylor estimate. More precisely, there exist positive constants C and b such that

$$|f(xy) - P_x(y)| \le C|y|^m \sup_{|z|\le b^{k+1}|y|,\ |\alpha|=m} |X^\alpha f(xz)|,$$

where $P_x(y)$ is the Taylor polynomial of f of degree $m-1$ based at x. Since the right-hand side involves the supremum of mth order derivatives, this formula does not appear to imply the Poincaré inequality we have assumed in (H1).

It is clear that polynomials for which (2.5) holds satisfy (2.4) by using repeatedly the known case $m = 1$ of (2.4). To see that the projection polynomials also satisfy (2.4), we use the following fact with $p = 1$.

LEMMA 2.8. *If P is any polynomial of degree less than m and $1 \le p \le \infty$, then*

$$\Big(\int_B |f(x) - \pi_m(B,f)(x)|^p\,dx\Big)^{1/p} \le C\Big(\int_B |f(x) - P(x)|^p\,dx\Big)^{1/p}$$

where C only depends on Q, m.

The proof is a simple corollary of (2.6) and the fact that $\pi_m(B,f) - P = \pi_m(B, f-P)$ if the degree of P is less than m.

We mention in passing that the existence of polynomials satisfying either (2.5) or (2.6) and (2.7) for general Carnot–Carathéodory vector fields is an open question. We are now ready to state the representation theorems on a stratified group.

COROLLARY E. *Let B_0 be a ball in $\mathbb{G}$, $f \in C^m(B_0)$, and m be a positive integer. Then there is a polynomial $P_m(B_0,f)$ of order less than m such that for $x \in B_0$,*

$$\begin{aligned}|f(x) &- P_m(B_0,f)(x)| \\ &\le C\int_{B_0} |X^m f(y)| \frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy + C\frac{r(B_0)^m}{|B_0|}\int_{B_0} |X^m f|\,dy\end{aligned}$$

with C independent of f, x and B_0. Moreover, if $m \le Q$ then

$$|f(x) - P_m(B_0,f)(x)| \le C\int_{B_0} |X^m f(y)| \frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy.$$

For a weak Boman chain domain, we have

COROLLARY F. *Let Ω be a weak Boman chain domain in $\mathbb{G}$ with a central ball B_0, and let $f \in C^\infty(\Omega)$. Then for any $m \le Q$ there is a polynomial $P_m(B_0,f)$ of order less than m such that for $x \in \Omega$,*

$$|f(x) - P_m(B_0,f)(x)| \le C\int_\Omega |X^m f(y)| \frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy.$$

The methods used in this paper are extensions of several techniques from [FW] and [LW2], together with the use of properties of polynomials. The presence of polynomials $P_m(B,f)$ of positive order causes some technical difficulties which do not occur in the case of order zero, i.e., when $P_m(B,f) = f_B$. If we instead prove Theorem B by adapting the technique used in [FLW2], we will have to assume the reverse doubling condition of order $m + \varepsilon$, i.e.,

$$\frac{\mu(B)}{\mu(\widetilde{B})} \ge c\Big(\frac{r(B)}{r(\widetilde{B})}\Big)^{m+\varepsilon} \quad \text{if } \widetilde{B} \subset B \subset cB_0$$

for some $\varepsilon > 0$, but then the segment property (H3) is not needed.

In passing, we note that the Poincaré assumption in Definition 2.1 can be weakened by allowing an enlarged ball on the right, namely by assuming that there is a constant $a_1 > 1$ such that for all balls B with $a_1 B \subset \Omega$, there is a polynomial $P_m(B,f)$ such that

$$\frac{1}{\nu(B)}\int_B |f-P_m(B,f)|\,d\nu \le C\frac{r(B)^m}{\mu(B)}\int_{a_1B}|g|\,d\mu,$$

and by also assuming that the measures ν,μ are related by the following balance condition: if B and $\widetilde{B}$ are balls with $a_1B\subset\widetilde{B}\subset\Omega$, then

$$\left(\frac{r(B)}{r(\widetilde{B})}\right)^m\frac{\nu(B)}{\nu(\widetilde{B})}\le C\frac{\mu(B)}{\mu(\widetilde{B})}.$$

This balance condition clearly holds automatically if $\nu=\mu$; moreover, for any two doubling measures, it is necessary for the L^1 Poincaré estimate if $m=1$, $g=|Xf|$ and X is a differential operator (see [CW], p. 1194). For $m>1$ and $g=|X^mf|$, this balance condition is also necessary. To see why we can allow an enlarged ball a_1B for $a_1>1$ on the right of the Poincaré inequality, we need the following lemmas.

LEMMA 2.9. *Let $\Omega\in\mathcal{F}(\sigma,M)$, μ and ν be measures and ν be doubling. Suppose that f is a function so that for each ball B with $\sigma B\subset\Omega$,*

$$\int_B |f-P_m(B,f)|\,d\nu\le A\int_{\sigma B}|g|\,d\mu$$

with A independent of B. Then

$$\int_\Omega |f-P_m(B_0,f)|\,d\nu\le cA\int_\Omega|g|\,d\mu,$$

where B_0 is a central ball for Ω and c depends only on σ,M and ν.

This sort of lemma is well known for $m=1$, even for quasimetrics; see, e.g., the comments following Theorem 5.2 of [FGW]. For $m>1$, it can be found in [B] in the Euclidean case and in [L4] for stratified groups. The main ingredient needed in the proof is those properties of polynomials we listed in (P1) and (P2).

LEMMA 2.10. *Let $(\mathcal{S},\varrho)$ be a metric space equipped with a doubling measure μ. If the segment hypothesis* (H3) *holds for a domain Ω, then every ball $B\subset\Omega$ is a Boman chain domain of type σ,M for any given σ with M depending only on σ and μ.*

This is Theorem 5.4 of [FGW].

If (H3) holds, we can use Lemmas (2.9) and (2.10) to show that if the Poincaré hypothesis (H1) holds for some $a_1>1$, then it also holds for $a_1=1$. In fact, fix any ball $\widetilde{B}\subset\Omega$. If B is a ball with $a_1B\subset\widetilde{B}$, condition (H1) and then the balance condition give

$$\int_B|f-P_m(B,f)|\,d\nu\le C\nu(B)\frac{r(B)^m}{\mu(B)}\int_{a_1B}|g|\,d\mu\le C\nu(\widetilde{B})\frac{r(\widetilde{B})^m}{\mu(\widetilde{B})}\int_{a_1B}|g|\,d\mu.$$

By Lemma 2.10, $\widetilde{B}$ is a Boman chain domain of type σ,M with $\sigma=a_1$ for some M. Thus, by Lemma 2.9 with Ω, σ and A there chosen to be $\widetilde{B}$, a_1 and $C\nu(\widetilde{B})r(\widetilde{B})^m/\mu(\widetilde{B})$ respectively, we obtain

$$\int_{\widetilde{B}}|f-P_m(\widetilde{B}_0,f)|\,d\nu\le C\nu(\widetilde{B})\frac{r(\widetilde{B})^m}{\mu(\widetilde{B})}\int_{\widetilde{B}}|g|\,d\mu$$

for some polynomial $P_m(\widetilde{B}_0,f)$, where $\widetilde{B}_0$ is the central ball in the Boman chain condition of $\widetilde{B}$. In this case we can take $P_m(\widetilde{B}_0,f)$ as $P_m(\widetilde{B},f)$. Thus we can assume without loss of generality that $a_1=1$ in the Poincaré part of hypothesis (H1), provided that (H3) and the balance condition hold.

3. Proof of Theorem D. We consider two cases: $x\in\varepsilon B_0$ and $x\in B_0\backslash\varepsilon B_0$ for $0<\varepsilon<1/5$. Parts of the argument are similar to the construction in [FW], but there is a difference now since the chain of balls will lie inside B_0.

CASE 1: $x\in\varepsilon B_0$. Let $B_0=B_0(x_0,r_0)$ and fix $x\in\varepsilon B_0$. We first show that there is a point z_0, depending on x, with $\varrho(x_0,z_0)=((1-3\varepsilon)/2)r_0$ such that if z is any point on the geodesic segment containing z_0 and x, then $B(z,\varrho(x,z))\subset B_0$. In fact, pick any point z_0 with $\varrho(x_0,z_0)=((1-3\varepsilon)/2)r_0$ and let z be a point on the geodesic γ connecting z_0 and x. If $\xi\in B(z,\varrho(x,z))$, first note that

$$\varrho(\xi,x_0)\le\varrho(\xi,z)+\varrho(z,x)+\varrho(x,x_0)<\varrho(x,z)+\varrho(x,z)+\varepsilon r_0=2\varrho(x,z)+\varepsilon r_0.$$

Also, since $z\in\gamma$,

$$\varrho(x,z)=\varrho(x,z_0)-\varrho(z_0,z)\le\varrho(x,z_0)\le\varrho(x,x_0)+\varrho(x_0,z_0)<\frac{1-\varepsilon}{2}r_0.$$

Combining estimates shows that $\xi\in B_0$ as desired.

Now fix z_0 as above and let $\varrho(x,z_0)=r_1$. The fact that $\varepsilon<1/5$ implies that $x\ne z_0$, i.e., $r_1>0$. Let $\gamma=\gamma(t)$, $0\le t\le r_1$, be the geodesic connecting z_0 and x with $\gamma(0)=z_0$, $\gamma(r_1)=x$ and $\varrho(\gamma(t),\gamma(s))=|t-s|$. Define $t_1=0$ and $t_{k+1}=(r_1+t_k)/2$ for $k\ge1$. Then $r_1-t_k=r_1/2^{k-1}$ for $k\ge1$, and $t_k\nearrow r_1$. For $0<\theta<1$ to be chosen and $k\ge1$, let

$$x_k=\gamma(t_k)\quad\text{and}\quad B_k=B(x_k,\theta(r_1-t_k)).$$

Note that $\varrho(x,x_k)=\varrho(\gamma(r_1),\gamma(t_k))=r_1-t_k$, and then

$$B_k\subset B(x_k,r_1-t_k)=B(x_k,\varrho(x,x_k))\subset B_0$$

by what we showed earlier, since $x_k\in\gamma$. Thus property (1) follows. Clearly, $r(B_k)\to0$ and

$$r(B_1)=\theta(r_1-t_1)=\theta r_1=\theta\varrho(x,z_0)\ge\theta[\varrho(x_0,z_0)-\varrho(x,x_0)]>\theta\frac{1-5\varepsilon}{2}r_0.$$

Thus, $r(B_1) \approx r(B_0)$ and property (2) holds. To show (3), fix $k \geq 1$ and $y \in B_k$. Then

$$\varrho(x,y) \leq \varrho(x,x_k) + \varrho(x_k,y) < (r_1 - t_k) + \theta(r_1 - t_k) = \frac{1+\theta}{\theta} r(B_k),$$

and similarly

$$\varrho(x,y) \geq \varrho(x,x_k) - \varrho(x_k,y) > \frac{1-\theta}{\theta} r(B_k).$$

To prove (4), note first that for each $k \geq 1$, $r(B_k) = 2r(B_{k+1})$ and

$$\varrho(x_k,x_{k+1}) = t_{k+1} - t_k = \frac{1}{2\theta} r(B_k).$$

Pick z_k on γ halfway between x_k and x_{k+1}:

$$\varrho(x_k,z_k) = \frac{t_{k+1}-t_k}{2} = \frac{\varrho(x_k,x_{k+1})}{2}.$$

Then also $\varrho(x_{k+1},z_k) = \varrho(x_k,z_k)$. Let

$$S_k = B(z_k, r(B_k)/8).$$

To show that $S_k \subset B_{k+1}$, note that if $\xi \in S_k$ then

$$\begin{aligned}\varrho(\xi,x_{k+1}) &\leq \varrho(\xi,z_k) + \varrho(z_k,x_{k+1}) < \frac{r(B_k)}{8} + \frac{t_{k+1}-t_k}{2} \\ &= \left(\frac{1}{8} + \frac{1}{4\theta}\right) r(B_k) = \left(\frac{1}{4} + \frac{1}{2\theta}\right) r(B_{k+1}) \\ &< r(B_{k+1}) \quad \text{if } \theta > 2/3.\end{aligned}$$

Thus $S_k \subset B_{k+1}$ if $\theta > 2/3$. Similarly, $\varrho(\xi,x_k) < r(B_k)$ if $\xi \in S_k$ and $\theta > 2/3$, so that $S_k \subset B_k \cap B_{k+1}$ if $\theta > 2/3$.

To prove (5), fix $j < k$ and let $\xi \in B_k$. Then

$$\varrho(\xi,x_j) \leq \varrho(\xi,x_k) + \varrho(x_k,x_j) < r(B_k) + |t_k - t_j|.$$

Since $k > j$, $t_k > t_j$ and $r(B_k) < r(B_j)$. Also,

$$|t_k - t_j| = r_1\left(\frac{1}{2^{j-1}} - \frac{1}{2^{k-1}}\right) < \frac{r_1}{2^{j-1}} = \frac{1}{\theta} r(B_j).$$

Thus, $\varrho(\xi,x_j) < (1 + 1/\theta) r(B_j)$, so that $B_k \subset (1+1/\theta)B_j$ as desired.

It remains to show (6). If $B_j \cap B_k \neq \emptyset$ then by the triangle inequality,

$$\varrho(x_j,x_k) < r(B_j) + r(B_k), \quad \text{i.e.,} \quad |t_j - t_k| < r(B_j) + r(B_k).$$

Since $t_j = r_1(1 - 1/2^{j-1})$ and $r(B_j) = \theta r_1/2^{j-1}$, we easily obtain

$$\left|\frac{1}{2^{j-1}} - \frac{1}{2^{k-1}}\right| < \theta \frac{1}{2^{j-1}} + \theta \frac{1}{2^{k-1}}.$$

Therefore,

$$|k - j| < \log_2 \frac{1+\theta}{1-\theta}.$$

Property (6) follows immediately, and the proof of Theorem D is complete in Case 1.

CASE 2: $x \in B_0 \setminus \varepsilon B_0$. The argument is similar to the one for Case 1. Recall that $B_0 = B(x_0, r_0)$. Fix $x \in B_0 \setminus \varepsilon B_0$ and let $\varrho(x,x_0) = r_1$. Then $\varepsilon r_0 \leq r_1 < r_0$. Let $\gamma = \gamma(t)$, $0 \leq t \leq r_1$, be a geodesic connecting x_0 and x with $\gamma(0) = x_0$, $\gamma(r_1) = x$ and $\varrho(\gamma(t),\gamma(s)) = |t - s|$. We first show that if $z \in \gamma$ then $B(z,\varrho(x,z)) \subset B_0$. In fact, if $z = \gamma(s)$ and $\xi \in B(z,\varrho(x,z))$, then $\xi \in B_0$ since

$$\begin{aligned}\varrho(\xi,x_0) &\leq \varrho(\xi,z) + \varrho(z,x_0) \\ &< \varrho(x,z) + \varrho(z,x_0) = (r_1 - s) + s = r_1 < r_0.\end{aligned}$$

Define a sequence $\{B_k\}_{k\geq 1}$ of balls as follows. Let $t_1 = \varepsilon r_0/2$, $t_{k+1} = (r_1 + t_k)/2$ for $k \geq 1$, $x_k = \gamma(t_k)$ and $B_k = B(x_k, \theta(r_1 - t_k))$ for $k \geq 1$, with $0 < \theta < 1$ to be chosen. Then $r_1 - t_k = (r_1 - t_1)/2^{k-1}$ for $k \geq 1$, and $t_k \nearrow r_1$. Also, since $0 < \theta < 1$ and $r_1 - t_k = \varrho(x,x_k)$, we get $B_k \subset B_0$ as before since $x_k \in \gamma$. Thus (1) is satisfied by $\{B_k\}_{k\geq 1}$. Property (2) holds since $r(B_k) \to 0$ and $r(B_1) = \theta(r_1 - \varepsilon r_0/2) \geq \theta\varepsilon r_0/2$. The arguments for the remaining properties are like those for Case 1. The only difference occurs for (5) and (6). For (5), if $j < k$, we now have

$$t_k - t_j = \left(r_1 - \frac{r_1 - t_1}{2^{k-1}}\right) - \left(r_1 - \frac{r_1 - t_1}{2^{j-1}}\right) < \frac{r_1 - t_1}{2^{j-1}} = \frac{1}{\theta} r(B_j),$$

and the rest of the proof of (5) is then as before. For (6), the estimate $|t_j - t_k| < r(B_j) + r(B_k)$ now yields

$$(r_1 - t_1)\left|\frac{1}{2^{j-1}} - \frac{1}{2^{k-1}}\right| < \theta(r_1 - t_1)\left\{\frac{1}{2^{j-1}} + \frac{1}{2^{k-1}}\right\},$$

so that we again obtain

$$\left|\frac{1}{2^j} - \frac{1}{2^k}\right| < \theta\left(\frac{1}{2^j} + \frac{1}{2^k}\right),$$

and the rest is the same. This completes the proof of Theorem D.

4. Proofs of Theorems A, B and C

Proof of Theorem A. We will use Theorem D. Let B_0 be a ball in $\mathcal{S}$ and suppose that (H1) and the segment property (H3) hold for B_0. Given $x \in B_0$, let $\{B_k\}_{k\geq 1}$ be a sequence of balls with the properties in Theorem D. Then

$$\begin{aligned}(4.1) \quad &|f(x) - P_m(B_0,f)(x)| \\ &\leq |f(x) - P_m(B_1,f)(x)| + |P_m(B_1,f)(x) - P_m(B_0,f)(x)|.\end{aligned}$$

For the second term on the right in (4.1), we get, for ν-a.e. $x \in B_0$,

$$\begin{aligned}|P_m(B_1,f)(x) - P_m(B_0,f)(x)| &\le \|P_m(B_1,f)(x) - P_m(B_0,f)(x)\|_{L^\infty_\nu(B_0)}\\ &\le \|P_m(B_1,f)(x) - P_m(B_0,f)(x)\|_{L^\infty_\nu(B_1)}\end{aligned}$$

by (P2) since $B_1 \subset B_0$ and $\nu(B_1) \approx \nu(B_0)$ by property (1) of Theorem D (ν is doubling). In view of (P1) this is bounded by

$$\begin{aligned}&\frac{C}{\nu(B_1)}\|P_m(B_1,f) - P_m(B_0,f)\|_{L^1_\nu(B_1)}\\
&\quad\le \frac{C}{\nu(B_1)}\int_{B_1}|f - P_m(B_1,f)|\,d\nu + \frac{C}{\nu(B_1)}\int_{B_1}|f - P_m(B_0,f)|\,d\nu\\
&\quad\le \frac{C}{\nu(B_1)}\int_{B_1}|f - P_m(B_1,f)|\,d\nu + \frac{C}{\nu(B_0)}\int_{B_0}|f - P_m(B_0,f)|\,d\nu\\
&\quad\le C\frac{r(B_1)^m}{\mu(B_1)}\int_{B_1}|g|\,d\mu + C\frac{r(B_0)^m}{\mu(B_0)}\int_{B_0}|g|\,d\mu\\
&\qquad\qquad\text{by the Poincaré inequality (H1)}\\
&\quad\le C\frac{r(B_0)^m}{\mu(B_0)}\int_{B_0}|g|\,d\mu\end{aligned}$$

since $B_1 \subset B_0$, $r(B_1) \approx r(B_0)$ and $\mu(B_1) \approx \mu(B_0)$.

Assuming as we may that x is a Lebesgue point for both $|f - P_m(B_1,f)|$ and $|g|$ with respect to ν and using properties (1)–(3) from Theorem D, we deduce for the first term on the right in (4.1) that

$$\begin{aligned}&|f(x) - P_m(B_1,f)(x)|\\
&\quad= \lim_{k\to\infty}\fint_{B_k}|f(y) - P_m(B_1,f)(y)|\,d\nu(y)\\
&\quad\le \limsup_{k\to\infty}\fint_{B_k}|f(y) - P_m(B_k,f)(y)|\,d\nu(y)\\
&\qquad+ \limsup_{k\to\infty}\fint_{B_k}|P_m(B_k,f)(y) - P_m(B_1,f)(y)|\,d\nu(y) = I_1 + I_2,\end{aligned}$$

where I_1 and I_2 are defined by the last equality. To show that $I_1 = 0$ for every Lebesgue point x of $|g|$, we use the Poincaré inequality:

$$I_1 \le C\limsup_{k\to\infty}\frac{r(B_k)^m}{\mu(B_k)}\int_{B_k}|g(y)|\,d\mu(y) = 0\cdot|g(x)| = 0.$$

Thus we only need to estimate I_2. We have

$$\begin{aligned}I_2 &\le \limsup_{k\to\infty}\sum_{j=1}^{k-1}\frac{1}{\nu(B_k)}\int_{B_k}|P_m(B_{j+1},f) - P_m(B_j,f)|\,d\nu\\
&\le \limsup_{k\to\infty}\sum_{j=1}^{k-1}\|P_m(B_{j+1},f) - P_m(B_j,f)\|_{L^\infty_\nu(B_k)}\\
&\le \limsup_{k\to\infty}\sum_{j=1}^{k-1}\|P_m(B_{j+1},f) - P_m(B_j,f)\|_{L^\infty_\nu(cB_j)}\end{aligned}$$

by property (5) of Theorem D. The last expression equals

$$\begin{aligned}&\sum_{j=1}^{\infty}\|P_m(B_{j+1},f) - P_m(B_j,f)\|_{L^\infty_\nu(cB_j)}\\
&\quad\le C\sum_{j=1}^{\infty}\|P_m(B_{j+1},f) - P_m(B_j,f)\|_{L^\infty_\nu(S_j)}\quad\text{by property (4) and (P2)}\\
&\quad\le C\sum_{j=1}^{\infty}\frac{1}{\nu(S_j)}\|P_m(B_{j+1},f) - P_m(B_j,f)\|_{L^1_\nu(S_j)}\quad\text{by (P1).}\end{aligned}$$

Using the triangle inequality, we see this is bounded by

$$\begin{aligned}&C\sum_{j=1}^{\infty}\fint_{S_j}|P_m(B_{j+1},f) - f|\,d\nu + C\sum_{j=1}^{\infty}\fint_{S_j}|P_m(B_j,f) - f|\,d\nu\\
&\quad\le C\sum_{j=1}^{\infty}\fint_{B_{j+1}}|P_m(B_{j+1},f) - f|\,d\nu + C\sum_{j=1}^{\infty}\fint_{B_j}|P_m(B_j,f) - f|\,d\nu\end{aligned}$$

since $S_j \subset B_j \cap B_{j+1}$ and $\nu(S_j) \approx \nu(B_j) \approx \nu(B_{j+1})$ by Theorem D. Combining estimates and applying (H1) to the terms of each of the last two sums, we obtain

$$I_2 \le C\sum_{j=1}^{\infty} r(B_j)^m \fint_{B_j}|g(y)|\,d\mu(y).$$

If $y \in B_j$, then

$$\frac{r(B_j)^m}{\mu(B_j)} \approx \frac{\varrho(x,y)^m}{\mu(B(y,\varrho(x,y)))} \approx \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}$$

by part (3) of Theorem D and the fact that μ is a doubling measure. Thus,

$$I_2 \le C\sum_{j=1}^{\infty}\int_{B_j}|g(y)|\frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\,d\mu(y)$$

$$\le C \int_{B_0} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\, d\mu(y)$$

by (6) and (1) of Theorem D. Theorem A now follows by combining estimates.

Proof of Theorem B. Theorem B is a simple corollary of Theorem A and the reverse doubling hypothesis (H2). In fact, if $x, y \in B_0$, then $\varrho(x,y) \le 2r(B_0)$ and consequently by (H2), we have

$$\frac{r(B_0)^m}{\mu(B_0)} \le C \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))} \quad \text{if } x, y \in B_0.$$

Thus, the second term on the right in the conclusion of Theorem A is majorized by the first term. This completes the proof.

Proof of Theorem C. Let $x \in \Omega$. By the definition of weak Boman chain domain, we may select B^* with $x \in B^*$ and a chain $\{B_j\}_{j=0}^k$ connecting $B^* = B_k$ to the central ball B_0. Then

$$\text{(4.2)} \quad \begin{aligned} &|f(x) - P_m(B_0, f)(x)| \\ &\quad \le |f(x) - P_m(B^*, f)(x)| + |P_m(B^*, f)(x) - P_m(B_0, f)(x)|. \end{aligned}$$

By Theorem B, the first term on the right side of (4.2) is at most

$$C \int_{B^*} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\, d\mu(y)$$

for ν-a.e. point of B^*, and we may assume this holds for our fixed x by initially excluding from Ω the set of measure zero formed by the union of the exceptional sets in each Boman ball. Since $B^* \subset \Omega$, we obtain the desired estimate

$$|f(x) - P_m(B^*, f)(x)| \le C \int_{\Omega} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\, d\mu(y).$$

For the second term on the right in (4.2), by using the chain $\{B_j\}$ connecting B_0 to $B_k = B^*$ and by noticing that $B^* \subset MB_j$ and $x \in B^*$, we have

$$\begin{aligned} &|P_m(B^*, f)(x) - P_m(B_0, f)(x)| \\ &\qquad \le \sum_{j=1}^k |P_m(B_j, f)(x) - P_m(B_{j-1}, f)(x)| \\ &\qquad \le \sum_{j=1}^k \|P_m(B_j, f) - P_m(B_{j-1}, f)\|_{L^\infty_\nu(MB_j)}. \end{aligned}$$

If D_j is a ball with $D_j \subset B_j \cap B_{j-1} \subset MB_j$ and $r(D_j) \approx r(B_j) \approx r(B_{j-1})$, then by (P1) and (P2), the last sum is in turn bounded by a constant times

$$\begin{aligned} &\sum_{j=1}^k \|P_m(B_j, f) - P_m(B_{j-1}, f)\|_{L^\infty_\nu(D_j)} \\ &\qquad \le \sum_{j=1}^k \frac{C}{\nu(D_j)} \|P_m(B_j, f) - P_m(B_{j-1}, f)\|_{L^1_\nu(D_j)} \\ &\qquad \le \sum_{j=1}^k \frac{C}{\nu(D_j)} \int_{D_j} |P_m(B_j, f) - f|\, d\nu \\ &\qquad\quad + \sum_{j=1}^k \frac{C}{\nu(D_j)} \int_{D_j} |P_m(B_{j-1}, f) - f|\, d\nu \\ &\qquad \le \sum_{j=1}^k \frac{C}{\nu(B_j)} \int_{B_j} |P_m(B_j, f) - f|\, d\nu \\ &\qquad\quad + \sum_{j=1}^k \frac{C}{\nu(B_{j-1})} \int_{B_{j-1}} |P_m(B_{j-1}, f) - f|\, d\nu \\ &\qquad \le \sum_{j=0}^k \frac{C}{\nu(B_j)} \int_{B_j} |P_m(B_j, f) - f|\, d\nu = I, \end{aligned}$$

where we have used the doubling property of ν. By Poincaré's inequality,

$$\begin{aligned} I &\le C \sum_{j=0}^k \frac{r(B_j)^m}{\mu(B_j)} \int_{B_j} |g(y)|\, d\mu(y) \\ &= C \int_{\Omega} \Big\{ \sum_{j=0}^k \frac{r(B_j)^m}{\mu(B_j)} \chi_{B_j}(y) \Big\} |g(y)|\, d\mu(y). \end{aligned}$$

The proof will be complete if we show that the sum above in curly brackets is bounded by a fixed multiple of $\varrho(x,y)^m/\mu(B(x,\varrho(x,y)))$ for each $y \in \Omega$. Fix $y \in \Omega$. By the definition of weak Boman chain domain, y belongs to at most M balls B_j in the chain, and for each such B_j, the fact that $B^* \subset MB_j$ implies that $x, y \in MB_j$, and consequently that $B(x,\varrho(x,y)) \subset 3MB_j$. Thus, for such j, by (H2),

$$\frac{r(3MB_j)^m}{\mu(3MB_j)} \le C \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}$$

and so by doubling,

$$\frac{r(B_j)^m}{\mu(B_j)} \le C\frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}.$$

Therefore, for any $y \in \Omega$, the sum above is at most M times the last expression, and therefore the proof is complete.

5. Applications to embedding theorems involving high order derivatives. Embedding theorems for first order vector fields have been studied extensively. Our purpose in this section is to show that many analogous results for higher order derivatives can be derived by using the representation formulas given earlier. For simplicity, we sometimes state the results only for smooth functions, but they remain valid for functions in appropriate Sobolev spaces of order m without assuming any extra smoothness. This can be seen by showing that the associated polynomials exist for general Sobolev functions (see [LLW]). We also note that Theorem 5.11 in §5.3 proves some particular smoothness for certain Sobolev functions.

5.1. *Weighted Poincaré inequalities.* We begin by using the representation formulas to derive weighted L^q to L^p Poincaré inequalities for high order vector field gradients on a stratified group $\mathbb{G}$. We consider only the range $1<p<q<\infty$, but the cases $1=p<q<\infty$ and $1<p\le q<\infty$ can also be treated by adapting the results in [FGW] and [PW].

If $w \in L^1_{\rm loc}(\mathbb{G})$ and $w(x)\ge 0$, we say that w is a *weight* and use the notation $w(E)=\int_E w(x)\,dx$ for any measurable set E. If w is a weight, we say that $w\in A_p$, $1<p<\infty$, if there is a constant C such that

$$\left(\frac{1}{|B|}\int_B w(x)\,dx\right)^{1/p}\left(\frac{1}{|B|}\int_B w(x)^{-p'/p}\,dx\right)^{1/p'} \le C$$

for all metric balls B.

THEOREM 5.1. *Let $(\mathbb{G},\varrho)$ be a stratified Lie group with metric ϱ, and let m be any positive integer less than the homogeneous dimension Q. Let $1<p<q<\infty$, B_0 be a metric ball, and w_1, w_2 be two weights satisfying the balance condition*

$$\left(\frac{r(B)}{r(B_0)}\right)^m\left(\frac{w_2(B)}{w_2(B_0)}\right)^{1/q} \le C\left(\frac{w_1(B)}{w_1(B_0)}\right)^{1/p} \tag{5.2}$$

for all metric balls B with $B\subset cB_0$. Suppose also that $w_1\in A_p$ and w_2 is doubling with respect to Lebesgue measure. If $f\in C^m(B_0)$, there is a polynomial $P_m(B_0,f)$ of degree less than m such that

$$\left(\frac{1}{w_2(B_0)}\int_{B_0}|f-P_m(B_0,f)|^q w_2\,dx\right)^{1/q} \le Cr(B_0)^m\left(\frac{1}{w_1(B_0)}\int_{B_0}|X^m f|^p w_1\,dx\right)^{1/p}. \tag{5.3}$$

The nonweighted case of (5.3) was given in [L3], [L4]. In the usual Euclidean case, Poincaré inequalities involving high order derivatives have been studied by Sobolev [So1], [So2] (see also [AH] and [Z]).

Proof of Theorem 5.1. First observe by Corollary E that there is a polynomial $P_m(B_0,f)$ such that

$$|f(x)-P_m(B_0,f)(x)| \le C\int_{B_0}|X^m f(y)|\frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy, \quad x\in B_0. \tag{5.4}$$

Define the integral operator

$$Tg(x)=\int_{\mathbb{G}} g(y)\frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy$$

for $g\ge 0$. Then we may rewrite (5.4) as

$$|f(x)-P_m(B_0,f)(x)|\chi_{B_0}(x) \le CT(|X^m f|\chi_{B_0})(x),$$

and consequently (5.3) will follow by verifying the norm estimate

$$\left(\int_{B_0}|T(g\chi_{B_0})(x)|^q w(x)\,dx\right)^{1/q} \le C\left(\int_{B_0}|g(x)|^p v(x)\,dx\right)^{1/p} \tag{5.5}$$

with weights w, v chosen to be

$$w(x)=\frac{1}{w_2(B_0)}w_2(x) \quad\text{and}\quad v(x)=\frac{r(B_0)^{mp}}{w_1(B_0)}w_1(x).$$

Since both w and $v^{-p'/p}$ are then doubling measures, Theorem 3B of [SW] guarantees that (5.5) holds for $1<p<q<\infty$ if for all balls $B\subset cB_0$, we have

$$\frac{r(B)^m}{|B|}\left(\int_B w(x)\,dx\right)^{1/q}\left(\int_B v(x)^{-p'/p}\,dx\right)^{1/p'} \le C. \tag{5.6}$$

(In fact, by combining Theorem 1.4 of [SWZ] with the sort of localization argument in [SW, p. 833], (5.6) suffices for (5.5) if $w, v^{-p'/p}$ just satisfy a reverse doubling condition of some positive order.) However, a simple computation shows that (5.6) follows from (5.2) since $v\in A_p$. This completes the proof of Theorem 5.1.

5.2. *Exponential type inequalities.* Let $(\mathcal{S},\varrho)$ be a metric space. A measure μ is said to be *doubling of order N* if there is a constant $C>0$ such that for any balls B_1 and B_2 with $B_1 \subset B_2$,

$$\mu(B_2) \le C\left(\frac{r(B_2)}{r(B_1)}\right)^N \mu(B_1). \tag{5.7}$$

We now fix a ball $B \subset \mathcal{S}$ and let

$$Tg(x) = T_B g(x) = \int_B g(y) \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\, d\mu(y).$$

THEOREM 5.8. *Let $(\mathcal{S},\varrho)$ be a metric space and μ be a doubling measure of order N. Let Tg be defined as above for a fixed ball $B \subset \mathcal{S}$. Suppose that $pm = N$ and $p > 1$. Then there exists a constant $C > 0$ independent of B and g such that*

$$\frac{1}{\mu(B)} \int_B \exp\left\{\left(\frac{r(B)^m}{C\mu(B)^{1/p}} \frac{|T_B g(x)|}{\|g\|_{L^p(B,d\mu)}}\right)^{p/(p-1)}\right\} d\mu(x) \le C.$$

Proof. The proof uses ideas of Hedberg [He] and is included for completeness. We may assume without loss of generality that g is supported in B and also that $\|g\|_{L^p(B,d\mu)} = 1$. For $x \in B$ and $0 < \delta < R = 2r(B)$,

$$\begin{aligned}|Tg(x)| &\le \int_{\varrho(x,y)\le\delta} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\, d\mu(y)\\ &\quad + \int_{\delta<\varrho(x,y)\le R} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\, d\mu(y) = I_1 + I_2.\end{aligned}$$

It is easy to see that

$$\begin{aligned}I_1 &= \sum_{k=0}^{\infty} \int_{2^{-k-1}\delta<\varrho(x,y)\le 2^{-k}\delta} |g(y)| \frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\, d\mu(y)\\ &\le C\sum_{k=0}^{\infty} \frac{(2^{-k}\delta)^m}{\mu(B(x,2^{-k}\delta))} \int_{\varrho(x,y)\le 2^{-k}\delta} |g(y)|\, d\mu(y)\\ &\le C\sum_{k=0}^{\infty} (2^{-k}\delta)^m M(g)(x) \le C\delta^m M(g)(x)\end{aligned}$$

where $M(g)$ is the Hardy–Littlewood maximal function of g with respect to μ. For I_2, we use Hölder's inequality to get

$$\begin{aligned}I_2 &\le \Big(\int_{\delta\le\varrho(x,y)\le R} |g|^p d\mu(y)\Big)^{1/p}\\ &\quad\times \Big(\int_{\delta\le\varrho(x,y)\le R} \Big(\frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}\Big)^{p'} d\mu(y)\Big)^{1/p'}.\end{aligned}$$

By dividing the integral domain into annuli $\varrho(x,y) \approx 2^{-k}R$ and using (5.7), we find that the second factor on the right above is at most

$$\begin{aligned}&\Big\{C \sum_{k=0}^{C\log(R/\delta)} \frac{(2^{-k}R)^{mp'}}{((2^{-k}R/R)^N \mu(B))^{p'-1}}\Big\}^{1/p'}\\ &\qquad \le \frac{CR^m}{\mu(B)^{1/p}} \Big\{\sum_{k=0}^{C\log(R/\delta)} 2^{k(N(p'-1)-mp')}\Big\}^{1/p'}\\ &\qquad \le \frac{CR^m}{\mu(B)^{1/p}} (\log(R/\delta))^{1/p'}\end{aligned}$$

since $N(p'-1) - mp' = 0$. Thus, using the fact that $\|g\|_{L^p(d\mu)} = 1$, we obtain

$$I_2 \le C(\log(R/\delta))^{1/p'} \frac{R^m}{\mu(B)^{1/p}}.$$

Combining estimates gives

$$|Tg(x)| \le C\delta^m M(g)(x) + C\frac{R^m}{\mu(B)^{1/p}} (\log(R/\delta))^{1/p'}.$$

With C as above, let $A = CR^{-m}\mu(B)^{1/p}$ and choose δ such that

$$\delta^m = \min\{2^{-1}A^{-1}M(g)(x)^{-1}, R^m\}.$$

Hence,

$$|Tg(x)| \le \frac{R^m}{2\mu(B)^{1/p}} + C\frac{R^m}{\mu(B)^{1/p}} [\log^+(R2^{1/m}A^{1/m}M(g)(x)^{1/m})]^{1/p'},$$

and then for those x such that $|Tg(x)| > R^m/\mu(B)^{1/p}$, we deduce by subtracting and rewriting that

$$\Big(\frac{\mu(B)^{1/p}}{2CR^m} |Tg(x)|\Big)^{p'} \le \log^+(R2^{1/m}A^{1/m}M(g)(x)^{1/m}).$$

Exponentiating and integrating over the set

$$E = \{x \in B : |Tg(x)| > R^m/\mu(B)^{1/p}\},$$

we obtain

$$\frac{1}{\mu(B)}\int_E \exp\left\{\left(\frac{\mu(B)^{1/p}}{2CR^m}|T(g)(x)|\right)^{p'}\right\}d\mu(x)$$
$$\le \frac{c}{\mu(B)}\int_B R2^{1/m}A^{1/m}M(g)(x)^{1/m}\,d\mu(x)$$
$$\le \frac{c}{\mu(B)}\mu(B)^{1/pm}\Big(\int_B M(g)(x)^p\,d\mu(x)\Big)^{1/pm}\mu(B)^{1/(pm)'}$$

by Hölder's inequality with exponents pm and $(pm)'$

$$\le c\Big(\int_B g^p(x)\,d\mu(x)\Big)^{1/(pm)} = c$$

by the boundedness of the Hardy–Littlewood maximal function and the fact that g is supported in B. Clearly,

$$\frac{1}{\mu(B)}\int_{B\setminus E}\exp\left\{\left(\frac{\mu(B)^{1/p}}{2CR^m}|T(g)(x)|\right)^{p'}\right\}d\mu(x)$$
$$\le \frac{1}{\mu(B)}\int_B \exp\left\{\left(\frac{1}{2C}\right)^{p'}\right\}d\mu(x) \le c.$$

The theorem now follows by combining the estimates above.

COROLLARY 5.9. *Let B be a metric ball in a stratified group $\mathbb{G}$ of homogeneous dimension Q, and let $p>1$ and m be a positive integer with $pm=Q$. If $f\in C^m(B)$, then there is a polynomial $P_m(B,f)$ of order less than m such that*

$$\frac{1}{|B|}\int_B \exp\left\{\left(\frac{|f(x)-P_m(B,f)(x)|}{C\|X^m f\|_{L^p(B,dx)}}\right)^{p/(p-1)}\right\}dx \le C,$$

with C independent of f and B. Moreover, for the same p and m, a similar result holds for any weak Boman domain Ω in $\mathbb{G}$: if B_0 is a central ball for Ω, then

$$\frac{1}{|\Omega|}\int_\Omega \exp\left\{\left(\frac{|f(x)-P_m(B_0,f)(x)|}{C\|X^m f\|_{L^p(\Omega,dx)}}\right)^{p/(p-1)}\right\}dx \le C.$$

Proof. The first statement follows immediately from the representation Corollary E and Theorem 5.8. Note that on a stratified group $\mathbb{G}$, $|B| = C_Q r(B)^Q$. The second statement can be obtained from Corollary F by modifying the proof used for Theorem 5.8, now applied to the integral operator

$$Tg(x) = \int_\Omega g(y)\frac{\varrho(x,y)^m}{|B(x,\varrho(x,y))|}\,dy.$$

The main change needed is a different way to choose the number R in the proof of Theorem 5.8. Given x, pick R so that $|B(x,R)|=|\Omega|$. The choice of R is independent of x. For δ to be chosen with $0<\delta\le R/2$,

$$|Tg(x)| \le I_1 + I_2 = \int_{\varrho(x,y)\le\delta}\ldots + \int_{\varrho(x,y)>\delta;\,\Omega}\ldots$$

We estimate I_1 as before. Again by Hölder's inequality, as $(m-Q)p' = -Q$,

$$I_2 \le c\|g\|_{L^p(\Omega,dy)}\Big(\int_{\varrho(x,y)>\delta;\,\Omega}\varrho(x,y)^{-Q}\,dy\Big)^{1/p'}$$
$$\le c\|g\|_{L^p(\Omega,dy)}\Big(\int_{\delta<\varrho(x,y)<R}\varrho(x,y)^{-Q}\,dy\Big)^{1/p'}$$

since

$$\int_{\varrho(x,y)\ge R;\,\Omega}\varrho(x,y)^{-Q}\,dy \le R^{-Q}|\Omega| = R^{-Q}|B(x,R)|$$
$$\le c\int_{\delta<\varrho(x,y)<R}\varrho(x,y)^{-Q}\,dy$$

by the choice of R and since $\delta\le R/2$. Dividing into annuli gives $I_2 \le c\|g\|_{L^p(\Omega,dy)}[\log(R/\delta)]^{1/p'}$. We then proceed as before but with μ replaced by Lebesgue measure and B replaced by Ω. To ensure $\delta\le R/2$, define δ by

$$\delta = \min\{2^{-1}A^{-1}M(g)(x)^{-1}, (R/2)^m\}$$

with A as before. Now if we choose $E=\{x\in\Omega : |Tg(x)|>c\}$ for a suitably large geometric constant c and note that $r^m|B(x,r)|^{-1/p}$ is a constant independent of x and r, the remainder of the estimation is largely as before with B replaced by Ω. This completes the proof of Corollary 5.9.

We now list exponential estimates for f itself instead of for $|f-P_m(B,f)|$.

COROLLARY 5.10. *Let B be a metric ball in a stratified group $\mathbb{G}$ of homogeneous dimension Q, and let $p>1$ and m be a positive integer with $pm=Q$. If $f\in C^m(B)$ and $\|X^m f\|_{L^p(B,dx)}\ne 0$, then*

$$\frac{1}{|B|}\int_B \exp\left\{\left(\frac{|f(x)|}{C\|X^m f\|_{L^p(B,dx)}}\right)^{p/(p-1)}\right\}dx$$
$$\le C\exp\left\{\left(\frac{C\fint_B |f|\,dx}{\|X^m f\|_{L^p(B,dx)}}\right)^{p/(p-1)}\right\},$$

with C independent of f and B. Moreover, for the same p and m, a similar result holds for any weak Boman domain Ω in $\mathbb{G}$ and any f with $\|X^m f\|_{L^p(\Omega,dx)} \neq 0$, we have

$$\frac{1}{|\Omega|}\int_\Omega \exp\left\{\left(\frac{|f(x)|}{C\|X^m f\|_{L^p(\Omega,dx)}}\right)^{p/(p-1)}\right\}dx \le C\exp\left\{\left(\frac{C(\Omega)\fint_{B_0}|f|\,dx}{\|X^m f\|_{L^p(\Omega,dx)}}\right)^{p/(p-1)}\right\},$$

where B_0 is a central ball for Ω and $C(\Omega)$ is independent of f.

Proof. Let C be the constant in Corollary 5.9 and pick $C' \ge 2C$. Fix a ball B and select a polynomial $P_m(B,f)$ with (see (2.6))

$$\|P_m(B,f)\|_{L^\infty(B,dx)} \le \frac{C''}{|B|}\int_B |f|\,dx.$$

Then the result for B follows from Corollary 5.9:

$$\begin{aligned}\int_B \exp&\left\{\left(\frac{|f(x)|}{C'\|X^m f\|_{L^p(B,dx)}}\right)^{p/(p-1)}\right\}dx\\ &\le \exp\left\{\left(\frac{2\|P_m(B,f)\|_{L^\infty(B,dx)}}{C'\|X^m f\|_{L^p(B,dx)}}\right)^{p/(p-1)}\right\}\\ &\quad\times\int_B \exp\left\{\left(\frac{2|f(x)-P_m(B,f)(x)|}{C'\|X^m f\|_{L^p(B,dx)}}\right)^{p/(p-1)}\right\}dx\\ &\le \exp\left\{\left(\frac{2C''\fint_B|f|\,dx}{C'\|X^m f\|_{L^p(B,dx)}}\right)^{p/(p-1)}\right\}C|B|.\end{aligned}$$

The second part of the corollary is proved similarly using the fact that

$$\|P_m(B_0,f)\|_{L^\infty(\Omega,dx)} \le C(\Omega)\|P_m(B_0,f)\|_{L^\infty(B_0,dx)} \le C(\Omega)\fint_{B_0}|f|\,dx.$$

5.3. *L^∞ estimates and Hölder continuity.* We now prove some estimates on stratified groups in case either $p=1$ and $m \ge Q$ or $p>1$ and $mp>Q$; these complement the results in §5.2 where $p>1$ and $pm=Q$.

THEOREM 5.11. *Let B be a metric ball in a stratified group $\mathbb{G}$ of homogeneous dimension Q. Let $p \ge 1$, $mp > Q$ if $p>1$ and $m \ge Q$ if $p=1$. Then if $f \in C^m(B)$, there exists a polynomial $P_m(B,f)$ of order less than m such that*

$$\|f-P_m(B,f)\|_{L^\infty(B,dx)} \le Cr(B)^{m-Q/p}\|X^m f\|_{L^p(B,dx)} \tag{5.12}$$

with C independent of f and B. In particular,

$$\|f\|_{L^\infty(B,dx)} \le \frac{C}{|B|}\int_B |f(y)|\,dy + Cr(B)^{m-Q/p}\|X^m f\|_{L^p(B,dx)}. \tag{5.13}$$

Moreover, if $p>1$, $m \le Q$ and $mp>Q$, a similar result holds for any weak Boman domain Ω in $\mathbb{G}$: if B_0 is a central ball, then

$$\|f-P_m(B_0,f)\|_{L^\infty(\Omega,dx)} \le C|\Omega|^{m/Q-1/p}\|X^m f\|_{L^p(\Omega,dx)}. \tag{5.14}$$

In particular, for the same ball B_0,

$$\|f\|_{L^\infty(\Omega,dx)} \le \frac{C(\Omega)}{|B_0|}\int_{B_0}|f(y)|\,dy + C|\Omega|^{m/Q-1/p}\|X^m f\|_{L^p(\Omega,dx)}. \tag{5.15}$$

Proof. By Theorem A there is a polynomial $P_m(B,f)$ of degree less than m so that

$$\begin{aligned}|f(x)&-P_m(B,f)(x)|\\ &\le C\int_B \varrho(x,y)^{m-Q}|X^m f(y)|\,dy + Cr(B)^{m-Q}\int_B |X^m f(y)|\,dy\end{aligned}$$

for $x \in B$. If $p=1$ and $m \ge Q$, then the right side is at most $Cr(B)^{m-Q}\times\int_B |X^m f(y)|\,dy$ and (5.12) follows. If $p>1$ and $mp>Q$, then by Hölder's inequality both terms on the right are easily seen to be bounded by $Cr(B)^{m-(Q/p)}\|X^m f\|_{L^p(B,dx)}$, which proves (5.12). If we choose $P_m(B,f)$ to satisfy

$$\|P_m(B,f)\|_{L^\infty(B,dx)} \le \frac{C}{|B|}\int_B |f(y)|\,dy, \tag{5.16}$$

then (5.13) follows from (5.12) by the triangle inequality.

If $m \le Q$ and B_0 is a central ball for Ω, by Theorem C and then Hölder's inequality, there is a polynomial $P_m(B_0,f)$ such that for all $x \in \Omega$,

$$\begin{aligned}|f(x)&-P_m(B_0,f)(x)|\\ &\le C\Big(\int_\Omega \varrho(x,y)^{(m-Q)p'}\,dy\Big)^{1/p'}\|X^m f\|_{L^p(\Omega,dx)}.\end{aligned} \tag{5.17}$$

By selecting R with $|B(x,R)| = |\Omega|$ and using $-Q < (m-Q)p' \le 0$, we see that the first factor on the right in (5.17) is bounded by

$$\Big(\int_{B(x,R)} \varrho(x,y)^{(m-Q)p'}\,dy + \int_{\varrho(x,y)>R;\,\Omega} R^{(m-Q)p'}\,dy\Big)^{1/p'} \le C|\Omega|^{m/Q-1/p}.$$

This completes the proof of (5.14). To prove (5.15), note that if $P_m(B_0,f)$ satisfies (5.16) for B_0 then

$$\begin{aligned}\|f\|_{L^\infty(\Omega,dx)} &\le \|P_m(B_0,f)\|_{L^\infty(\Omega,dx)} + C|\Omega|^{m/Q-1/p}\|X^m f\|_{L^p(\Omega,dx)}\\ &\le C(\Omega)\|P_m(B_0,f)\|_{L^\infty(B_0,dx)} + C|\Omega|^{m/Q-1/p}\|X^m f\|_{L^p(\Omega,dx)}\\ &\le \frac{C(\Omega)}{|B_0|}\int_{B_0}|f(y)|\,dy + C|\Omega|^{m/Q-1/p}\|X^m f\|_{L^p(\Omega,dx)}.\end{aligned}$$

We emphasize that $C(\Omega)$ is independent of f. This ends the proof of (5.15).

5.4. *Embedding theorems in Campanato–Morrey spaces.* Embedding theorems on Campanato–Morrey spaces are useful for studying regularity of solutions of partial differential equations. In the classical case, the spaces and some of their applications have been studied for example in [C1], [C2], [C3] and in [S1], [S2]. For degenerate vector fields of first order, embedding theorems on non-isotropic Campanato–Morrey spaces were studied in [L5], [L6]. In this section, we prove an embedding theorem for analogous spaces involving high order vector field gradients on a stratified group $\mathbb{G}$. This sort of embedding theorem allows the larger gap $1/p-1/q$ than the gap in the L^p to L^q Poincaré inequalities.

Let $\Omega\subset\mathbb{G}$, $f\in C^m(\Omega)$ and $\pi_m(B,f)$ be the projection polynomial of degree less than m associated with f and a ball $B\subset\Omega$ (see §2). Let $\mathcal{P}_m$ be the collection of polynomials in $\mathbb{G}$ of degree less than m. We define the following two types of Campanato–Morrey norms: first, for $\lambda\ge 0$ and $1\le p<\infty$, let $\mathcal{L}_m^{p,\lambda}(\Omega)$ be the space of all $f\in L^p_{\rm loc}(\Omega)$ with

$$\|f\|_{\mathcal{L}_m^{p,\lambda}(\Omega)} = \sup_{B\subset\Omega}\Big(r(B)^\lambda\mu(B)^{-1}\inf_{P\in\mathcal{P}_m}\int_B|f-P|^p\,d\mu\Big)^{1/p}<\infty.$$

Hereafter, by using Lemma 2.8, we redefine

$$\|f\|_{\mathcal{L}_m^{p,\lambda}(\Omega)} = \sup_{B}\Big(r(B)^\lambda\mu(B)^{-1}\int_B|f-\pi_m(B,f)|^p\,d\mu\Big)^{1/p}.$$

We also define the space $M^{p,\lambda}(\Omega)$ of all $f\in L^p_{\rm loc}(\Omega)$ with

$$\|f\|_{M^{p,\lambda}(\Omega)} = \sup_{B\subset\Omega}\Big(r(B)^\lambda\mu(B)^{-1}\int_B|f|^p\,d\mu\Big)^{1/p}<\infty.$$

Note that if $q\ge p$, $\mu/q\le\lambda/p$ and Ω is bounded, then $\mathcal{L}_m^{q,\mu}(\Omega)\subset\mathcal{L}_m^{p,\lambda}(\Omega)$; in fact,

$$\begin{aligned}\Big(r(B)^\lambda\fint_B|f-\pi_m(B,f)|^p\,d\mu\Big)^{1/p} &= r(B)^{\lambda/p}\Big(\fint_B|f-\pi_m(B,f)|^p\,d\mu\Big)^{1/p}\\ &\le r(B)^{\lambda/p}\Big(\fint_B|f-\pi_m(B,f)|^q\,d\mu\Big)^{1/q}\le C\Big(r(B)^\mu\fint_B|f-\pi_m(B,f)|^q\,d\mu\Big)^{1/q}\end{aligned}$$

since $r(B)^{\lambda/p-\mu/q}\le C$ (all balls are in Ω and thus $r(B)\le C$). Taking the sup over B gives $\|f\|_{\mathcal{L}_m^{p,\lambda}(\Omega)}\le C\|f\|_{\mathcal{L}_m^{q,\mu}(\Omega)}$. The same remark applies to the spaces $M_m^{p,\lambda}$.

In the definitions of $M_m^{p,\lambda}$ and $\mathcal{L}_m^{p,\lambda}$, the role played by balls could instead be played by any family $\{E\}$ of sets which are comparable to balls in the sense that there is a constant c so that for each set E, there are balls B' and B'' with $B'\subset E\subset B''$ and $\mu(B'')\le c\mu(B')$.

We now state the main theorem of this section. In case $m=1$, it was proved in [L5], [L6] for any vector fields for which a representation formula holds.

THEOREM 5.18. *Let $\mathbb{G}$ be a stratified Lie group with homogeneous dimension Q, and define the Campanato–Morrey spaces as above with μ taken to be Lebesgue measure. If $1<p<\lambda/m\le Q$ and $p^*=\lambda p/(\lambda-pm)$, then*

$$\|f\|_{\mathcal{L}_m^{p^*,\lambda}(\Omega)}\le C\|X^m f\|_{M^{p,\lambda}(\Omega)}$$

with C independent of f.

Given a real-valued function f on Ω, a doubling measure μ, and $\gamma>0$, define the fractional maximal function of f with respect to μ by

$$M_\gamma f(x)=\sup_B\frac{r(B)^\gamma}{\mu(B)}\int_B|f(y)|\,d\mu(y),$$

where the sup is taken over all metric balls $B\subset\Omega$ with center x. The analogous function when $\gamma=0$ is the Hardy–Littlewood maximal function

$$Mf(x)=\sup_B\frac{1}{\mu(B)}\int_B|f(y)|\,d\mu(y),$$

with the supremum taken as above. Also, for any ball $B\subset\Omega$, let

$$Tg(x)=T_Bg(x)=\int_B\frac{\varrho(x,y)^m}{\mu(B(x,\varrho(x,y)))}g(y)\,d\mu(y).$$

To prove Theorem 5.18, we need a lemma relating these operators.

LEMMA 5.19. *Let $\mathbb{G}$ be a stratified Lie group with homogeneous dimension Q, let μ be a doubling measure and $1<p<\lambda/m\le Q$. There is a constant C such that for any metric ball $B\subset\Omega$, any $x\in\mathbb{G}$ and any function g,*

$$|T_Bg(x)|\le CM_{\lambda/p}(|g|\chi_B)(x)^{pm/\lambda}\cdot M(|g|\chi_B)(x)^{1-pm/\lambda}.$$

In particular, if $f\in C^\infty(\Omega)$ and M_γ and M are the maximal functions above with $d\mu=dx$, there is a constant C such that for any metric ball $B\subset\Omega$ and any $x\in B$,

$$\begin{aligned}|f(x)-\pi_m(B,f)(x)|&\le CM_{\lambda/p}(|X^mf|\chi_B)(x)^{pm/\lambda}\cdot M(|X^mf|\chi_B)(x)^{1-pm/\lambda}\\ &\quad+Cr(B)^{m-Q}\int_B|X^mf(y)|\,dy,\end{aligned}$$

where $\pi_m(B,f)$ is any polynomial satisfying the L^1 to L^1 Poincaré inequality of order m for f, B.

For the first statement, the proof when $m = 1$ and $d\mu = dx$ is given in [L5], [L6] by adapting ideas from [He], and the general case is similar. We omit the details. The second statement follows from the first one and Theorem A. We are now ready to prove the main theorem.

Proof of Theorem 5.18. Fix any ball B_0. By the second part of 5.19,

$$\begin{aligned}|f(x)-\pi_m(B_0,f)(x)|\chi_{B_0}(x) &\le C[M_{\lambda/p}(|X^m f|\chi_{B_0})(x)]^{pm/\lambda}[M(|X^m f|\chi_{B_0})(x)]^{1-pm/\lambda}\\ &\quad + Cr(B_0)^m\frac{1}{|B_0|}\int_{B_0}|X^m f(y)|\,dy = I_1+I_2.\end{aligned}$$

Clearly,

$$\begin{aligned}\int_{B_0} I_1^{p^*}\,dx &\le C\|M_{\lambda/p}(|X^m f|\chi_{B_0})\|_{L^\infty(B_0)}^{p^*pm/\lambda}\\ &\quad\times\int_{B_0}\{M(|X^m f|\chi_{B_0})(x)\}^{(1-pm/\lambda)p^*}\,dx\\ &\le C\|M_{\lambda/p}(|X^m f|\chi_{B_0})\|_{L^\infty(B_0)}^{p^*pm/\lambda}\int_{B_0}|X^m f|^p\,dx,\end{aligned}$$

where for the second factor we have used the fact that $p^*(1-pm/\lambda) = p$ and the L^p boundedness of the Hardy–Littlewood operator for $p > 1$. By the definition of $M_{\lambda/p}$,

$$\|M_{\lambda/p}(|X^m f|\chi_{B_0})\|_{L^\infty(B_0)} \le C\sup_B r(B)^{\lambda/p}|B|^{-1}\int_B|X^m f|\,dy,$$

where the sup is taken over all balls B. By Hölder's inequality, this is bounded by

$$C\sup_B\Big(r(B)^\lambda|B|^{-1}\int_B|X^m f|^p\,dy\Big)^{1/p} = C\|X^m f\|_{M^{p,\lambda}(\Omega)},$$

and consequently,

$$\int_{B_0} I_1^{p^*}\,dx \le C\|X^m f\|_{M^{p,\lambda}(\Omega)}^{pp^*m/\lambda}\int_{B_0}|X^m f|^p\,dx.$$

Also,

$$\int_{B_0} I_2^{p^*}\,dx \le C|B_0|\frac{r(B_0)^{mp^*}}{|B_0|^{p^*/p}}\Big(\int_B|X^m f|^p\,dx\Big)^{p^*/p}.$$

Hence,

$$\begin{aligned}\Big(\frac{r(B_0)^\lambda}{|B_0|}\int_{B_0}|f-\pi_m(B_0,f)|^{p^*}\,dx\Big)^{1/p^*} &\le C\|X^m f\|_{M^{p,\lambda}(\Omega)}^{pm/\lambda}\Big(\frac{r(B_0)^\lambda}{|B_0|}\int_{B_0}|X^m f|^p\,dx\Big)^{1/p^*}\\ &\quad + Cr(B_0)^{\lambda/p^*+m-\lambda/p}\|X^m f\|_{M^{p,\lambda}(\Omega)},\end{aligned}$$

and since $pm/\lambda + p/p^* = 1$, the right side of the inequality is bounded by

$$C\|X^m f\|_{M^{p,\lambda}(\Omega)}^{pm/\lambda}\|X^m f\|_{M^{p,\lambda}(\Omega)}^{p/p^*} + C\|X^m f\|_{M^{p,\lambda}(\Omega)} = C\|X^m f\|_{M^{p,\lambda}(\Omega)}.$$

Since B_0 is any ball in Ω, we get the desired conclusion.

REMARK. By extending the techniques of this paper, the authors of [LP] have weakened the assumption in our main theorems that an L^1 to L^1 Poincaré inequality holds, showing that it is enough to assume an L^p to L^1 Poincaré inequality for some p with $0 < p < 1$.

References

[AH] D. Adams and L. Hedberg, *Function Spaces and Potential Theory*, Grundlehren Math. Wiss. 314, Springer, New York, 1996.

[B] B. Bojarski, *Remarks on Sobolev imbedding theorems*, in: Lecture Notes in Math. 1351, Springer, New York, 1988, 52–68.

[BH] B. Bojarski and P. Hajłasz, *Pointwise inequalities for Sobolev functions and some applications*, Studia Math. 106 (1993), 77–92.

[C1] S. Campanato, *Proprietà di inclusione per spazi di Morrey*, Ricerche Mat. 12 (1963), 67–86.

[C2] —, *Proprietà di hölderianità di alcune classi di funzioni*, Ann. Scuola Norm. Sup. Pisa 17 (1963), 175–188.

[C3] —, *Proprietà di una famiglia di spazi funzionali*, ibid. 18 (1964), 137–160.

[CDG] L. Capogna, D. Danielli and N. Garofalo, *Subelliptic mollifiers and a basic pointwise estimate of Poincaré type*, Math. Z. 226 (1997), 147–154.

[CW] S. Chanillo and R. L. Wheeden, *Weighted Poincaré and Sobolev inequalities and estimates for weighted Peano maximal functions*, Amer. J. Math. 107 (1985), 1191–1226.

[FP] C. Fefferman and D. H. Phong, *Subelliptic eigenvalue problems*, in: Proc. Conf. on Harmonic Analysis (Chicago 1980), W. Beckner *et al.* (eds.), Wadsworth, 1981, Vol. 2, 590–606.

[FS] G. B. Folland and E. M. Stein, *Hardy Spaces on Homogeneous Groups*, Princeton Univ. Press, Princeton, 1980.

[F] B. Franchi, *Inégalités de Sobolev pour des champs de vecteurs lipschitziens*, C. R. Acad. Sci. Paris Sér. I Math. 311 (1990), 329–332.

[FGW] B. Franchi, C. Gutiérrez and R. L. Wheeden, *Weighted Sobolev–Poincaré inequalities for Grushin type operators*, Comm. Partial Differential Equations 19 (1994), 523–604.

[FHK] B. Franchi, P. Hajłasz and P. Koskela, *Definitions of Sobolev classes on metric spaces*, preprint, 1998.

[FL] B. Franchi and E. Lanconelli, *Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coefficients*, Ann. Scuola Norm. Sup. Pisa 10 (1983), 523–541.

[FLW1] B. Franchi, G. Z. Lu and R. L. Wheeden, *Representation formulas and weighted Poincaré inequalities for Hörmander vector fields*, Ann. Inst. Fourier (Grenoble) 45 (1995), 577–604.

[FLW2] —, —, —, *A relationship between Poincaré-type inequalities and representation formulas in spaces of homogeneous type*, Internat. Math. Res. Notices 1996, no. 1, 1–14.

[FW] B. Franchi and R. L. Wheeden, *Some remarks about Poincaré type inequalities and representation formulas in metric spaces of homogeneous type*, J. Inequal. Appl. 3 (1999), 65–89.

[GT] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Springer, Berlin, 1983.

[H] P. Hajłasz, *Sobolev spaces on an arbitrary metric space*, Potential Anal. 5 (1996), 403–415.

[HK] P. Hajłasz and P. Koskela, *Sobolev met Poincaré*, Mem. Amer. Math. Soc., to appear.

[He] L. Hedberg, *On certain convolution inequalities*, Proc. Amer. Math. Soc. 36 (1972), 505–510.

[LLW] Y. Liu, G. Z. Lu and R. L. Wheeden, *Several equivalent definitions of high order Sobolev spaces on stratified groups and generalizations to metric spaces*, preprint, 1998.

[L1] G. Z. Lu, *Weighted Poincaré and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications*, Rev. Mat. Iberoamericana 8 (1992), 367–439.

[L2] —, *The sharp Poincaré inequality for free vector fields: an endpoint result*, ibid. 10 (1994), 453–466.

[L3] —, *Local and global interpolation inequalities for the Folland–Stein Sobolev spaces and polynomials on the stratified groups*, Math. Res. Lett. 4 (1997), 777–790.

[L4] —, *Polynomials, higher order Sobolev extension theorems and interpolation inequalities on weighted Folland–Stein spaces on stratified groups*, Acta Math. Sinica Ser. B, to appear.

[L5] —, *Embedding theorems on Campanato–Morrey spaces for vector fields and applications*, C. R. Acad. Sci. Paris Sér. I Math. 320 (1995), 429–434.

[L6] —, *Embedding theorems on Campanato–Morrey spaces for vector fields of Hörmander type*, Approx. Theory Appl. 14 (1998), 69–80.

[LP] G. Z. Lu and C. Pérez, *L^p to L^1 Poincaré inequalities for $0 < p < 1$ imply representation formulas*, preprint, 1999.

[LW1] G. Z. Lu and R. L. Wheeden, *Poincaré inequalities, isoperimetric estimates and representation formulas on product spaces*, Indiana Univ. Math. J. 47 (1998), 123–151.

[LW2] —, —, *An optimal representation formula for Carnot–Carathéodory vector fields*, Bull. London Math. Soc. 30 (1998), 578–584.

[N] D. Nhieu, *Extension of Sobolev spaces on the Heisenberg group*, C. R. Acad. Sci. Paris Sér. I Math. 321 (1995), 1559–1564.

[PW] C. Pérez and R. L. Wheeden, *Uncertainty principle estimates for vector fields*, to appear.

[SW] E. Sawyer and R. L. Wheeden, *Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces*, Amer. J. Math. 114 (1992), 813–874.

[SWZ] E. Sawyer, R. L. Wheeden and S. V. Zhao, *Weighted norm inequalities for operators of potential type and fractional maximal functions*, Potential Anal. 5 (1996), 523–580.

[So1] S. L. Sobolev, *On a boundary value problem for polyharmonic equations*, Amer. Math. Soc. Transl. 33 (1963), 1–40.

[So2] —, *Some Applications of Functional Analysis in Mathematical Physics*, Transl. Math. Monogr. 90, Amer. Math. Soc., Providence, RI, 1991.

[S1] G. Stampacchia, *The spaces $L^{p,\lambda}$, $N^{p,\lambda}$ and interpolation*, Ann. Scuola Norm. Sup. Pisa 19 (1965), 443–462.

[S2] —, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble) 15 (1965), no. 1, 189–258.

[Z] W. Ziemer, *Weakly Differentiable Functions*, Grad. Texts in Math. 120, Springer, New York, 1989.

Department of Mathematics
Wayne State University
Detroit, MI 48202, U.S.A.
E-mail: gzlu@math.wayne.edu

Department of Mathematics
Rutgers University
Hill Center for Mathematical Studies
New Brunswick, NJ 08903, U.S.A.
E-mail: wheeden@math.rutgers.edu

Received November 24, 1998
Revised version April 25, 2000
(4213)