A Local Study of Carnot-Carathéodory Metrics

A Dissertation presented

by

John Mitchell

to

The Graduate School

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in

Mathematics

State University of New York

at

Stony Brook

May, 1982

STATE UNIVERSITY OF NEW YORK

AT STONY BROOK

THE GRADUATE SCHOOL

John Mitchell

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of the dissertation.

Detlef Gromoll — Committee Chairman

Michael Taylor — ~~Txxxxx Axxxxxx~~

John Thorpe

Peter VanNieuwenhuizen, Physics (outside member)

The dissertation is accepted by the Graduate School

Dean of the Graduate School

Mr. Mitchell's dissertation adviser is Mikhail Gromov. Professor Gromov is currently in Paris and was not present at the examination.

Abstract of the Dissertation

A Local Study of
Carnot-Carathéodory Metrics

by

John Mitchell

Doctor of Philosophy

in

Mathematics

State University of New York at Stony Brook

1982

A smooth distribution W of k-planes on a smooth Riemannian manifold M is said to be infinitesimally transitive if for some local basis of vector fields $X_1,\ldots X_k$ for the distribution, the X_i's, along with all of their commutators, span the tangent space of M at each point (locally). The inverse function theorem then implies local transitivity: any two points of M may be joined by a piecewise smooth curve γ which is horizontal, i.e., a.e. tangent to W. Define a metric d on M by $d(m_1,m_2) = \inf(\text{length } \gamma)$ where inf is over all horizontal paths γ joining m_1 to m_2. This is called the Carnot-

Carathéodory metric associated to W and it arises, for example, on certain nilpotent Lie groups as a Hausdorff limit of the family of Riemannian metrics obtained by applying a sequence of contractions to a left invariant Riemannian metric. I study the Hausdorff dimension of the non-smooth metric d as well as its tangent cone, in certain cases, by using techniques of approximation by nilpotent Lie groups which were developed by Rothschild and Stein in their study of hypoelliptic operators. In the cases for which these approximation methods work, an effective formula for the Hausdorff dimension is found in terms of the commutator sequence of the distribution.

To my lovely wife, Marilyn,
who has revealed to me beauty
where I had not seen it before,
and brought me joy that I would
otherwise not have known.

Notation

M = smooth manifold, TM = its tangent bundle, $\exp: TM \to M$ the exponential mapping.

If X_i, $i = 1,\ldots,n$ are smooth vector fields on M then, for each multi-index $I = (i_1,i_2,\ldots,i_k)$ $1 \le i_j \le n$ for $j = 1,\ldots,k$, we denote by X_I the commutator of order k: $[X_{i_1},[X_{i_2},\ldots[X_{i_{k-1}},X_{i_k}]\ldots]$.

The order of a multi-index I as above is $|I| \equiv k$.

If $J = (j_1,j_2,\ldots,j_k)$ then $(i,J) \equiv (i_1,j_1,j_2,\ldots j_k)$ etc.

"exp" will also denote the mapping: $\mathbb{R} \times \Gamma(TM) \times M \to M$ defined by:

$$\exp(t,X,m) = \exp_m(t\cdot X).$$

ACKNOWLEDGEMENTS

I am pleased to have the opportunity to acknowledge my gratitude toward the Mathematics Department here at Stony Brook for their generous support, and to Professor Detlef Gromoll for having taught me geometry during my first years here.

Professor Michael Gromov, my advisor, has always been extremely generous in offering his time, his ideas and questions, and his support. I feel privileged to have been able to witness the artistry of his toughts, to what limited extent I could understand them. I will always be deeply thankful to him for all his kindness.

Finally, I want to thank Estella Shivers for her kindness and patience in translating my sloppy manuscript into a neat typed version.

§1. Introduction to Carnot-Carathéodory metrics

We wish to investigate Carnot-Carathéodory metrics which may be defined on a smooth manifold M which is equipped with both a Riemannian metric and a smooth distribution of n-planes satisfying Hörmander's condition. A smooth distribution assigns to each point $p \in M$ an n-dimensional subspace of the tangent space T_pM, and this subspace depends smoothly on p. A set of vector fields $X_1, X_2, \ldots, X_n$ satisfies Hörmander's condition of order k at p if the X_i's and their commutators up to and including those of order k span the tangent space T_pM. A commutator of order k is one of the form $[X_{i_1}, \ldots [X_{i_{k-1}}, X_{i_k}] \ldots]$.

A distribution satisfies Hörmander's condition at p by definition, if a local basis of vector fields for the distribution satisfies this condition (at p). This does not depend on the choice of local basis since the subbundle of TM generated by the X_i's and their commutators of order $\leq m$ depends only on the distribution: if a new set of X_i's is chosen which generates the same distribution, the new commutators of order k are expressed as linear combinations with coefficients in $C^\infty(M)$ of the old commutators of order $\leq k$ (including the X_i's, of course).

Riemannian manifolds equipped with such distributions arise in a variety of situations. For example:

1) Several authors (Hörmander, Folland, Rothschild, Stein, Goodman...) have investigated regularity properties of hypoelliptic operators of the form $X_1^2 + X_2^2 + \cdots + X_n^2 + X_{n+1}$, where the vector fields $X_1, \ldots X_{n+1}$ satisfy Hormander's condition. It is interesting to try to cast this analytical work into a more geometrical form, and the Carno metric, defined later, seems to be a natural and important geometric invariant of this set-up.

2) Carnot metrics also arise as Hausdorff limits of nilpotent Lie groups by replacing a left-invariant Riemannian metric, g, by $B_i g$ where the B_i's are real numbers which tend to zero as $i \to \infty$. Under certain conditions, the sequence of metric spaces $(G, B_i g)$ converges, in the sense of Hausdorff (see Definition (1) below) to the group G equipped with a Carnot metric defined by a left-invariant distribution (see Pansu). Thus, it appears that a local study of Carnot metrics might, in these cases, be related to the global geometry of the group.

3) A third example may be taken from non-linear control theory or non-holonomic mechanics. Here situations arise where one wishes to move a point from one position in its phase space to another, while obeying certain restraints, in an optimal time. A ball rolling on a table without slipping is a simple example ($M = SO(3) \times \mathbb{R}^2$ and the

distribution is defined by the "no slipping" constraint).

We will soon see how to define a singular (i.e., non-smooth) metric in such a situation. Firstly, though, we need to state a result which says that infinitesimal transitivity (Hormander's condition) implies local transitivity.

Chow's Theorem: If a smooth distribution W satisfies Hörmander's condition at $p \in M$ then any point $q \in M$ which is sufficiently close to p can be joined to p by a piecewise smooth curve which is almost everywhere (a.e.) tangent to the distribution.

Remark: Thus, any two points of M may be so joined if M is connected and if the distribution satisfies Hormander's condition at each point of M. A curve which is a.e. tangent to W will be called horizontal.

Since we will need to refer to the proof of Chow's Theorem later, we provide a short version here.

Proof: Firstly, choose a linearly independent set from among the X_i's (a local basis for W near p) and their commutators which spans $T_p(M)$ for each $p \in M$ (locally), and suppose that the associated multi-indices are $I_1, \ldots I_N$. To each multi-index I we associate a flow on M as follows:

Set $\phi_i(t) = \exp(tX_i)$ for $i = 1, \ldots n$ and for $I = (i,J)$, set

$$\phi_I(t) = \phi_J(-\sqrt{t}) \circ \phi_i(-\sqrt{t}) \circ \phi_J(\sqrt{t}) \circ \phi_i(\sqrt{t}).$$

We now define a map $\phi : \mathbb{R}^N \to M$. For each $\vec{t} = (t_1, \dots t_N) \in \mathbb{R}^N$, define $\phi(\vec{t}) = \phi_{I_N}(t_N) \circ \phi_{I_{N-1}}(t_{N-1}) \circ \dots \circ \phi_{I_1}(t_1)$.

Since if ϕ_1 and ϕ_2 are any two smooth flows then $[\phi_1, \phi_2](t)$ has zero first order part $t = 0$, $[\phi_1, \phi_2](t)$ is $O(t^2)$ and so $[\phi_1, \phi_2](\sqrt{t})$ is C^1-smooth.

Therefore, all of the ϕ_I above and ϕ are C^1-smooth. One proves easily by induction that $\dot{\phi}_I(t) = X_I = [X_{i_1}, [X_{i_2}, \dots, X_{i_N}]] \dots]$, so the differential of ϕ maps $\frac{\partial}{\partial t_j}$ to X_{I_j} $j = 1, \dots n$. Thus, ϕ is non-singular at the origin and so it is a local C^1-diffeomorphism there. Now the definition of ϕ shows that any point q near p can be reached by a piecewise-smooth curve where "pieces" are integral curves of the X_i's (because $q = \phi(\vec{t})$ for some $\vec{t}$). Obviously such a curve is horizontal, so the theorem is proved.

Q.E.D.

Notice that, for small t, the point $\phi_I(t)(p)$ lies roughly in the direction X_I and at the distance t from p, and that the piecewise smooth curve joining these two points, which appears implicitely in the definition of ϕ_I, has length of the order $t^{1/2^{k-1}}$ $(k=|I|)$, whereas one would expect from the Campbell-Hausdorff formula for Lie groups that these points could be joined by a horizontal curve of length of the order $t^{1/k}$, at least in the bese cases. We see, then, that the mapping ϕ will not be of use in making sharp estimates for the Carnot-Carathéodory metric.

Definition: The Carnot-Carathéodory distance between the points p and q in M, denoted by $d_1(p,q)$, is defined as the infimum of the lengths of all piecewise smooth horizontal curves joining p to q.

This (locally defined) metric on M is the object of our study.

Remark: Since all Riemannian metrics are locally Lipschitz equivalent, the Carnot metric is independent, up to local equivalence, of the choice of Riemannian metric.

Historical Note: The names of Carnot and Carathéodory are attached to this metric through their work on thermodynamics. In that theory, the infinitesimal amount of heat "δQ" added to a system is expressed as a linear combination of the changes in thermodynamic parameters. These latter define coordinates for the phase space of the system, on which is therefore defined a linear form (1-form) δQ. For adiabatic processes no heat transfer occurs, so the point in phase space describing the state of the system moves along a path for which $\delta Q = 0$, that is, the path is horizontal with respect to the distribution defined by Kernal (δQ).

As an example, consider the Heisenberg group. This is a three dimensional, simply connected, nilpotent Lie group with the one dimensional center.

In $\mathbb{R}^3$ with coordinates (x,y,z) we obtain a model for the Heisenberg group by defining the vector fields

$$X = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial z} \quad Y = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial z}; \ [X,Y] = \frac{\partial}{\partial z},$$

and using the Campbell-Baker-Hausdorff formula to define the group multiplication (see Adams; Bourbaki....).* Notice that X and Y generate the kernal of the 1-form $\omega \equiv dz + \frac{1}{2}(ydx-xdy)$. Thus, a curve C, tangent to the distribution generated by X and Y and which joins (x',y',z') to (x'',y'',z'') may be projected to the x-y plane: $(x,y,z) \to (x,y)$ to obtain a curve joining (x',y') whose "area" $(=\int_C (xdy-ydx))$ is equal to $\int_C dz = z'' - z'$. Evidentaly, any such curve in the (x,y) plane may also be lifted to a horizontal curve joining $(x',y'z')$ to (x'',y'',z''). Thus, any two points in $\mathbb{R}^3$ may be joined by a smooth horizontal curve, and from the isopermetric inequality it follows that there is always a smooth minimal curve, namely, one which projects to an arc of a circle. The minimal curve is unique unless $(x'',y'') = (x',y')$, in which case any circle of the appropriate area will do.

Remark: The projection of the curve C to the x-y plane may be described intrinsically as follows: by the map induced by left multiplication, carry all tangent vectors to C back to (x',y',z'), thus obtaining a curve in $T(M)(x',y',z')$. Now integrate this curve to get a curve which lies in the

---

* Adams, J.F. "Lectures on Lie Groups" New York, W.A. Benjamin 1969
N. Bourbaki "Eléments de Mathématique XXXIV, Groups et Algèbres de Lie", Hermann, Paris

plane which is the kernal of ω at (x',y',z'). This latter curve is isometric to the projection of C. This follows from the fact that left translation carries a tangential vector at (x',y',z'), say $aX + bY$, to the vector $a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}$ at the origin (i.e. identity), and this is just the projection onto the x-y plane.

Notice that, since X and Y are orthonormal with respect to the obvious choice of left-invariant Riemannian metric, the above remark also implies that the length of C (w.r.t. this metric) = length of projection of C w.r.t. Euclidean metric. So, for example, $d_1((0,0,0),(0,0,z))$ = length of circle of area $z \sim \sqrt{z}$. This shows that the Carnot metric is, in general, singular. In general, we may demonstrate the existence of minimal horizontal curves which are absolutely continuous by a standard and elementary method (cf. Filippov): one chooses a minimizing sequence of curves, then shows that an appropriate limit curve exists. We may, for this purpose, restrict our considerations to curves of speed ≤ 1, whereupon the problem becomes that of minimizing the time required to travel from a point p to a point q. We also shall take our manifold M to be isometrically embedded in some Euclidean space with norm $\|\cdot\|$.

Choose a sequence, $x_i(t)$, of horizontal paths with $x_i(0) = p$ and $x_i(t_i) = q$ such that $t_i \to t_\infty \equiv \inf\{t_i\}$. These curves are uniformly bounded and equicontinuous since $\|\dot{x}_i(t)\| \leq 1$, so we may choose a uniformly convergent subsequence with limit $x(t)$. The equicontinuity implies that $x(t_\infty) = q$ and uniform convergence implies that $x(t)$ is Lipschitz with constant 1, since the $x_i(t)$ satisfy this property. Thus, $x(t)$ is absolutely continuous.

To see that $x(t)$ is horizontal almost everywhere, consider a point t_0 of $(0,t_\infty)$ where $\dot{x}(t_0)$ exists, and choose a small, positive number ϵ. For all t sufficiently close to t_0, the secant $\frac{x(t) - x(t_0)}{t - t_0}$ approximates within ϵ the derivative $\dot{x}(t_0)$. The former is, in turn, arbitrarily well approximated (for fixed t) by the secant: $\frac{x_i(t) - x_i(t_0)}{t - t_0}$ if i is large. This last term is equal to $\frac{1}{t - t_0}\int_{t_0}^{t} \dot{x}_i(\tau)d\tau$.

For t close to t_0 and $t_0 \leq \tau \leq t$, $\dot{x}_i(\tau)$ lies in the ϵ-neighborhood of $W(x(t_0))$. W denotes the distribution spanned by the X_i's. Since this neighborhood is convex, the integral also lies in the neighborhood. Therefore, $\dot{x}(t_0)$ is arbitrarily close to $W(x(t_0))$ and so is, in fact, in $D(x(t_0))$.

Q.E.D.

§2. Hausdorff dimension and tangent cone of Carnot-Carathéodory metrics

For the Heisenberg group, there were smooth extremal curves. In general, since d_1 is singular, one may ask what the tangent cone of the Carnot metric looks like, and what its Hausdorff dimension is. In certain cases, we can answer these questions. Note that the answers do not depend upon the original choice of Riemannian metric.

Here are the definitions of these terms:

Definition (1): The Hausdorff distance between two subspaces of a metric space z is defined as $H_Z(X,Y) = \inf\{\varepsilon \mid Y \subset N_\varepsilon(X),\ X \subset N_\varepsilon(Y)\}$, where N_ε denotes the ε-neighborhood.

For two metric spaces X,Y, one takes the infimum of $H_Z(X,Y)$ over all isometric imbeddings $X,Y \to z$, and denotes this by H(X,Y). A sequence X_i of metric spaces converges in the sense of Hausdorff to a metric space Y if $\lim_{i\to\infty} H(X_i,Y) = 0$.

The tangent cone of the metric d_1 is the Hausdorff limit of the sequence of metric spaces (M,rd_1) where rd_1 denotes the metric d_1 multiplied by $r > 0$, and $r \to \infty$. (This gives the usual tangent space when applied to a Riemannian metric.)

(Here one must use limits of metric spaces with base points.)

The Hausdorff δ-measure ($\delta \geq 0$) of a subset X ($H^\delta(X)$) of a metric space is defined (up to a constant factor) by covering X by sets of diameter $\leq \epsilon$, taking the infimum of $(\sum_{\text{covering}} (\text{diam})^\delta)$ over all such coverings, and then taking the limit of this (increasing) sequence of numbers as $\epsilon \to 0$. The Hausdorff dimension of X is equal to

$\inf\{\delta \mid H^\delta(X) = 0\} = \sup\{\delta \mid H^\delta(X) = \infty\}$.

In [Gr], Gromov provides a condition, called uniform compactness, which guarantees the existence of a limit metric space for a sequence (X_i, x_i) of proper metric spaces with distinguished points. (Proper means each closed ball is compact.)

The Heisenberg group may again be used as an illustration. The Lie algebra, g, of G admits a family of dilations, that is, a one parameter group of Lie algebra automorphisms, denoted by δ_t, $t \geq 0$. These are defined on X and Y by $\delta_t(X) = tX$, $\delta_t(Y) = tY$, and extended uniquely to automorphisms of g. The automorphisms may be made to act on G via the exponential map.

The dilation δ_t multiplies volumes by t^4 (rather than t^3), since in exponential coordinates, it takes (x,y,z) to (tx,ty,t^2z). It is clear that the Carnot metric is homogeneous of degree one with respect to δ_t, i.e. δ_t multiplies lengths of horizontal curves by t. Thus, if B_t denotes the ball of radius t w.r.t. d_1,

$$\mathrm{vol}(B_t) = \mathrm{vol}(\delta_t(B_1)) = t^4 \mathrm{vol}(B_1),$$

which easily implies that the Hausdorff dimension of the Carnot metric on the Heisenberg group is four.

The homogeneity of d_1 implies that the tangent cone (at $\vec{0}$) of (G,d_1) is equal to (G,d_1) itself, since δ_t is an isometry between (G,d_1) and (G,td_1).

If the reader would like to become better acquainted with Hausdorff limits, he or she may demonstrate the following (or any of the other examples to be found in [Gr]).

Let X be the free Abelian group of rank two with two fixed generators. Let d be the word metric on X, and let e = identity. Then the sequence $(X,x_i d)$ $x_i \to 0$ converges to the plane $\mathbb{R}^2$ with the Minkowski metric

$$\mathrm{dist}((a,b),(a',b')) = |a-a'|+|b-b'|.$$

We saw above that the homogeneity (i.e., existence of a family of dilations) of the Heisenberg group allowed us to compute its Hausdorff dimension and its tangent cone with ease. Based on this observation, we seek, for the more general case of a distribution on a manifold M, a family of pseudo-dilations on M. These appear most naturally if the picture resembles that of a Lie group with dilations. Fortunately, such an approximation result already exists in the work of Rothschild-Stein, under certain conditions on the distribution.

Definition: A family of dilations on a (finite divisional real) Lie algebra g is a 1-parameter group $\{\gamma_r\}$ $r > 0$ of Lie algebra automorphisms of the form $\gamma_r = \exp(A \log(r)$ where A is a diagonalizable linear transformation of g with positive eigenvalues.

Clearly, a Lie algebra must be nilpotent to admit a family of dilations. However, not all nilpotent Lie groups do admit such families. Dyer constructed a nilpotent Lie algebra all of whose automorphisms are nilpotent. Dilations are not nilpotent, so such groups do not admit dilations.

To state the theorem of Rothschild and Stein, we assume that, besides satisfying Hörmander's condition of order m, the vector fields X_i are free up to step m i.e., the X_i's and their commutators of order $\leq m$ satisfy no linear relations except anticommutivity and the Jacobi identity and their consequences. This is the same as saying that the dimension of the space spanned by the X_i's and their commutators of order $\leq m$ is the same as that of $G_{n,m}$, the simply-connected Lie group whose Lie algebra is the free nilpotent Lie algebra of step m on n generators, denoted by $g_{n,m}$. Denote the generators of $g_{n,m}$ by $Y_1, \ldots Y_n$ and let γ_r be the dilation defined on the Y_i's by $\gamma_r(Y_i) = rY_i$ and extended uniquely to all of $g_{n,m}$ (and $G_{n,m}$).

Set $Y_{j1} = Y_j$ for $j = 1,\ldots n$ and for $k = 2,\ldots n$, let $Y_{1k}, Y_{2k},\ldots$ be a maximal linearly independent set of the commutators of order k.

The Y_{jk}'s form a basis for $g_{n,m}$. Taking coordinates with respect to this basis and composing with the exponential map, one obtains coordinates on $G_{n,m}$ in which the coordinate functions are homogeneous of degrees $1,\ldots m$ (with respect to γ_r).

A vector field X on $G_{n,m}$ may be written $X = \sum_{j,k} a_{jk} Y_{jk}, a_{jk} \in C^{\infty}(M)$. If we expand the a_{jk}'s in their Taylor series about 0 in the coordinates defined above, X will be exhibited as a formal sum of homogeneous differential operator (if f is a homogenous function of degree i (w.r.t. γ_r), then fY_{jk} is a homogeneous operator of degree k_i). We say that X is of <u>local degree $\leq \lambda$</u> if each term in this formal sum is homogeneous of degree $\leq \lambda$. We then have the approximation theorem: (Rothschild-Stein).

<u>Theorem</u>: Given vector fields $X_1,\ldots X_n$ on M, free up to step m and satisfying Hörmander's condition of order m in an open neighborhood U of M, there is, for each $p \in U$, a diffeomorphism $\theta_p : U \to V$ (a neighborhood of the origin in $G_{n,m}$) whose differential, θ_*, maps X_i to $Y_i + R_i$, where R_i is of local degree ≤ 0.

The dilations now act on M locally via the diffeomorphism θ, and they may be used to estimate the (Riemannian) volumes of small balls in the Carnot metric, as well as to determine the structure of the tangent cone of this metric.

We may now work in V, a neighborhood of 0 = origin of $G_{n,m}$.

Definitions (2):

W = the distribution spanned by $X_1, \ldots X_n$ ($= Y_1+R_1 \ldots Y_n+R_n$)

W_r = distribution spanned by $\gamma_{r_*}(X_1), \ldots \gamma_{r_*}(X_n)$

where γ_{r_*} = differential of γ_r

d_r = Carnot metric associated to W_r (and the fixed left invariant Riemannian metric on $G_{n,m}$)

W_∞ = distribution spanned by $Y_1, \ldots Y_n$

d_∞ = Carnot metric associated to W_∞.

$B_r(k)$ (resp. $S_r(k)$) = ball (resp. sphere) of radius k with respect to d_r $1 \le r \le \infty$.

With the above assumptions on the distribution W, we have the following results:

Theorem 1: The tangent cone of (M,d_1) is isometric to $(G_{n,m}, d_\infty)$.

Theorem 2: The Hausdorff dimension of (M,d_1) is equal to

Q, the homogeneous dimension of $G_{n,m}$ (see Folland).

Let $g = g_1 \supset g_2 \supset \ldots \supset g_m, g_{m+1} = \{0\}$ be the descending central sequence of g. Let $r_k = \dim(g_k/g_{k+1})$. Then $q = \sum_{k=1}^{m} kr_k$.

Theorem 1 is proved by showing that:

1.1) The quasi-isometric distance between $r \cdot d_1$ and d_r tends to zero as r tends to ∞. (By quasi-isometric distance we mean log (infimum of metric distortion* over all homeomorphisms $f : M \to M$)).

1.2) d_r converges, in the sense of Hausdorff, to d_∞ as $r \to \infty$.

One may then apply the simple estimate

<u>Lemma 0:</u> † $\frac{H(X,Y)}{\operatorname{diam}(X) + \operatorname{diam}(Y)} \leq C\cdot(X,Y)$

where diam = diameter and (X,Y) = quasi-isometric distance between X and Y, to obtain Theorem 1.

Theorem 2 may be obtained from an estimate of $\operatorname{vol}(B_1(\epsilon))$ (vol. = Riemannian volume):

** $C^{-1}\epsilon^Q \leq \operatorname{vol}(B_1(\epsilon) \leq C\,\epsilon^Q$ for some $C > 1$ and all small ϵ.

This, in turn, follows from the estimate:

---

* The metric distortion of a map $f: (M_1,d_1)(M_2,d_2)$ between metric spaces is defined as $\sup_{x,y \in M_1} \max\left(\frac{d_2(f(x),f(y))}{d_1(x,y)}, \frac{d_1(x,y)}{d_2(f(x),f(y))}\right)$

<u>Lemma 1</u>:

$$B_1(\frac{1}{Cr}) \subset \gamma_{1/r}(B_1(1)) \subset B_1(\frac{C}{r}) \qquad \text{for some } C > 1 \text{ and all large } r,$$

along with the fact that γ_r multiplies volumes by r^Q (one lets $\epsilon = \frac{1}{Cr}$ on $\frac{C}{r}$, in turn).

<u>Proof of Lemma 0</u>: Let us first prove the simple estimate (+); suppose that X has a metric d_1, and Y has a metric d_2. $(X,Y) < \infty \Rightarrow$ there exists a homeomorphism $f : X \to Y$ whose distortion is arbitrarily close to (X,Y). Identify Y with X via f, so we now have a single space X with two metrics d_1 and d_2. We may imbed each of these metric spaces isometrically into a single metric space, namely, $C^0(X)$ F (continuous functions on X with metric d induced from sup norm) by sending $x \in X$ to the distance function $d_i(x,\cdot)$. Call these two imbeddings F_1 and $F_2 : X \to C^0(X)$. For any $X_1, X_2 \in X$, $|d_1(X_1,X_2) - d_2(X_1,X_2)| \leq \text{diam}(X) + \text{diam}(Y)$ which implies $H(X,Y)| \leq \text{diam}(X) + \text{diam}(Y)$, thus we may suppose that that $(X,Y) \leq 1$. Now the two inequalities

$$|\log(\frac{d_1(X_1,X_2)}{d_2(X_1,X_2)})| \leq (X,Y) \text{ and } \max\{d_1(X_1,X_2),\ d_2(X_1,X_2)\} \leq \text{diam}(X) + \text{diam}(Y)$$

imply that $|d_1(x_1,x_2) - d_2(x_1,x_2)| \leq (\text{diam}(X) + \text{diam}(Y)) \cdot (1-e^{-(X,Y)}) < C(X,Y)$ if $(X,Y) \leq 1$ for some constant C.

Thus $d(F_1(X_1),F_2(X_1)) < C(X,Y)$ for all $X_1 \in X$ Q.E.D.

Proof of Lemma 1: Consider now the estimate in Lemma 1: This may be paraphrased as follows: Up to bounded distortion, γ_r, applied to curves or vectors in $\gamma_{1/r}(B_1(1))$ which are tangent to W, multiplies lengths by r. For the proof, let $x_0 \in S_1(1)$. To estimate the Carnot distance of $\gamma_{1/r}(x_0)$ from $0 \in G_{n,m}$, we need to estimate how γ_r acts on vectors in W whose base point lies in $\gamma_{1/r}(B(1))$. Let $y \in B(1)$ and let $V \in W(\gamma_{1/r}(y))$. Then $V = \sum_i v_i Y_i|_{\gamma_{1/r}(y)} + \sum_i v_i R_i|_{\gamma_{1/r}(y)}$ $(v_i \in \mathbb{R})$.

Thus $\gamma_{r*}(V) = r\cdot\sum_i v_i y_i + \sum_i v_i \gamma_{r*}(R_i(\gamma_{1/r}(Y)))$, since $\gamma_{r*}(Y_i) = r\cdot Y_i$. Now, the definition of "R_i has local degree ≤ 0" implies that the length of $\gamma_{r*}(R_i(\gamma_{1/r}(y))$ remains bounded as $r \to \infty$ (Proof: the homogeneous terms in the formal expansion of R_i as a sum of homogeneous operators look like $a_{jk,m}Y_{jk}$ if a_{jk} has the formal expansion $a_{jk} = \sum_{m=0}^{\infty} a_{jk,m}$ where $a_{jk,m}$ is a function homogeneous of degree m. Since $a_{jk,m}(\gamma_{1/r}(y)) = r^{-m}a_{jk,m}(y)$ and $\gamma_{r*}(Y_{jk}(\gamma_{1/r}(y)) = r^k Y_{jk}(y)$, we have: $\gamma_{r*}(a_{jk,m}Y_{jk}(\gamma_{1/r}(y))) = r^{k-m}a_{jk,m}Y_{jk}(y)$. "$R_i$ of local degree ≤ 0" means $k - m \leq 0$, so such a term remains bounded (in fact it decreases) as $r \to \infty$. This implies the result).

Also, $|R_i(\gamma_{1/r}(y))| \to 0$ as $r \to \infty$ ("$|\ \ |$" means Riemannian length), since $R_i(0) = 0$. Therefore,

$$\# \quad \frac{1}{r}\frac{|\gamma_{r*}(V)|}{|V|} = \frac{1}{r}\frac{\left|r\cdot\sum_i v_i Y_i\right|_y + \sum_i v_i \gamma_{r*}(R_i(\gamma_{1/r}(y)))\Big|}{\left|\sum_i v_i Y_i\right|_{\gamma_{1/r}(y)} + \sum_i v_i R_i(\gamma_{1/r}(y))\Big|} \to 1 \text{ as } r \to \infty$$

and so this expression is bounded below and above by $\frac{1}{C}$ and C respectively for some $C > 1$, for all sufficiently large r.

From this estimate on vectors we get the estimate on curves. If $p : [0,1] \to G_{n,m}$ is a path in $G_{n,m}$ joining 0 to $\gamma_{1/r}(x_0)$ which is tangent to the distribution W a.e., and which lies in $\gamma_{1/r}(B(1))$, then $\gamma_r(p)$ is a path joining 0 to x_0, and its length is therefore bounded below by a positive constant (= Riemannian distance from 0 to x_0). From the inequality on vectors given above, we see that

$$\text{const.} \le \text{length}(\gamma_r(p)) \le c\cdot r \text{ length}(p),$$

which gives the left side of the inequality in Lemma 1 (with a new C). We will see below (Proposition 1) that $d_r \to d_\infty$ in the sense of Hausdorff. This implies that $B_\infty(k) \subset B_r(k+\delta)$ for all sufficiently large r, and some δ. Also, it is clear that $B_1(1) \subset B_\infty(k)$ for some k, so $B(1) \subset B_r(k+\delta)$ for all large r. This shows that we may choose a piecewise smooth path, $\tilde{p}$, tangent to W_r and joining 0 to x_0, of length $\le k+\delta$ = const. Then $p = \gamma_{1/r}(\tilde{p})$

is tangent to W, joins 0 to $\gamma_{1/r}(x_0)$ and satisfies

$$\text{length}(p) \le \text{const}/r \qquad \text{for some constant.}$$

This gives the right side of the inequality in Lemma 1, once statement (1.2) is proved.

Proof of (1.1): Note that we have proven that $\lim_{r\to\infty}\left(\frac{\text{length}(\gamma_r(p))}{r\ \text{length}(p)}\right) = 1$ (this follows from #) which is precisely the meaning of statement (1.1).

We now wish to prove (1.2).

Proposition 1: d_r converges, in the sense of Hausdorff, to d_∞ as $r \to \infty$. This will be done by using two lemmas:

Lemma 2: $\lim_{r\to\infty} W_r = W_\infty$, the limit referring to the C^N topology on the space of distributions (for any $N \subset \infty$).

Proof: W_r is spanned by the vector fields $\{\frac{1}{r}\circ\gamma_{r_*}(X_i)\} = \{Y_i + \frac{1}{r}\gamma_{r_*}(R_i\circ\gamma_{1/r})\}$. Again, by definition of "R_i has local degree ≤ 0" one sees that the C^N norm of $\gamma_{r_*}(R_i\circ\gamma_{1/r})$ is bounded on a compact set by the C^N norm of R_i on this set.* The latter if finite since R_i is smooth. Thus $\frac{1}{r}\gamma_{r_*}(R_i\cdot\gamma_{1/r}) \to 0$ in C^N as $r \to \infty$. Q.E.D.

Lemma 3: If W_∞ is a smooth distribution satisfying Hormander's condition and if W_r is a family of smooth distributions converging, in the C^N sense for any N, to W_∞ as $r \to \infty$, then d_r

---

* $\gamma_{r_*}(a_{jk,m}Y_{jk}\circ\gamma_{1/r})(x) = r^{k-m}(a_{jk,m}Y_{jk})(x)$ and $k - m \le 0$.

converges in the sense of Hausdorff to d_∞.

Proof: First let us note that if W_∞ satisfies Hörmander's condition of order m, then taking $N \geq m$, we see that W_r does also, for large r. (Derivatives converge ⇒ commutators converge). We will need the following uniform estimate on the size of the d_r's, whose proof will come later.

Lemma 4: There is a function $F(\rho) > 0$ defined for $\rho > 0$, such that $\lim_{\rho\to 0} F(\rho) = 0$ and $\mathbb{B}(\rho) \subset B_r(F(\rho))$ for all sufficiently large $r : R \leq r \leq \infty$ (this R may depend on ρ).

Here $\mathbb{B}(\rho)$ denotes the Riemannian ball of radius ρ centered at $0 \in G_{n,m}$. This says simply that if two points are Riemannian close, they are d_r-close for all large r.

Now, as in the proof of Lemma 0, in order to estimate $H((M,d_r), (Md_\infty))$, we need to estimate, for every pair of points $p,q \in U \subset M$, the difference $|d_\infty(p,q) - d_r(p,q)|$. This estimate must be uniform for all pairs of points and all $r \geq R$, and it must approach zero as $R \to \infty$. This is done by producing, for each piecewise smooth curve joining p to q which is tangent to W_{r_1} a curve tangent to W_{r_2} joining p to $\tilde{q}$, where $\tilde{q}$ is uniformly close to q (i.e. $d(q,\tilde{q}) \leq \epsilon(R)$ where $\epsilon(R)$ depends only on R, $R \leq r_1 \leq r_2 \leq \infty$, and $\epsilon(R) \to 0$ as $R \to \infty$.) Moreover, the lengths of the two curves will be equal. Then, since by Lemma 4, $d_{r_2}(q,\tilde{q}) \leq F(\epsilon(R))$, we have $|d_{r_2}(p,q) - d_{r_1}(p,q)| \leq F(\epsilon(R))$, so $H(d_{r_1},d_{r_2}) \leq F(\epsilon(R)) \to 0$ as $R \to \infty$ and Lemma 3 will be proved. Notice that by Lemma 4, we may assume that all curves which we consider have length $\le$ const., and unit speed; thus the time of travel is $\le$ const.

A curve c_1, tangent to W_{r_1} and joining p to q satisfies:

$c_1(0) = p$, $\dot{c}_1(t) = \sum_{i=1}^{n} a_i(t) X_i^{r_1}(c_1(t))$ a.e., where the X_i^r's generate W_r (and $\lim_{r\to\infty} X_i^r = X_i^\infty$ in C^N). Define a new curve, tangent to W_{r_2}, by: $c_2(0) = p$, $\dot{c}_2(t) = \sum_{i=1}^{n} a_i(t) X_i^{r_2}(c_2(t))$.

We may assume that $\{X_i^r\}$ $i = 1,\ldots n$, is an orthonormal set for all r, so that $|\dot{c}_1(t)| = |\dot{c}_2(t)|$ a.e. and therefore, length (c_1) = length (c_2). If T is the final time, i.e. $c_1(T) = q$, $c_2(T) = \tilde{q}$, we must estimate $d(c_2(T), c_1(T))$ (d = Riemannian distance)

$$(c_2-c_1)(t) = \int_0^t \sum_{i=1}^{n} a_i(t)(X_i^{r_1}(c_2(t))-X_i^{r_1}(c_1(t)))dt$$

$$+ \int_0^t \sum_{i=1}^{n} a_i(t)(X_i^{r_2}(c_2(t))-X_i^{r_1}(c_2(t)))dt.$$

Now, if $r_1, r_2 \ge R$, then $|X_1^{r_2}(c_2(t)) - X_1^{r_1}(c_2(t))| \le \epsilon = \epsilon(R)$ where $\epsilon(R) \to 0$ as $R \to 0$. This is because the X_i^r's converge. Since $|\dot{c}_1(t)| = 1$, $\sum_{i=1}^{n} a_i^2(t) = 1$ so $\sum_{i=1}^{n} |a_i(t)| \le 1$.

Furthermore, since the X_i^r's are smooth and we're working in some bounded set, these are numbers R and L such that for all $r \ge R$,

$$\|X_i^r(c_2(t)) - X_i^r(c_1(t)\| \leq L\,\|c_2(t)-c_1(t)\| \quad \text{for all } i$$

where we use the norm $\|\vec{v}\| = \|(v_1,v_2,\ldots v_n)\| = \sum_{i=1}^{n}|v_i|$.

(The Lipschitz constant L is uniform in r since the X_i^r's converge in C^1). Thus:

$$\|c_2(t)-c_1(t)\| \leq L\cdot\int_0^t\|c_2(t)-c_1(t)\|dt$$

$$+ \epsilon(R)t \leq L\int_0^t\|c_2(t)-c_1(t)\|dt + \epsilon T$$

for $0 \leq t \leq T$. We now apply the following estimate (see p.11 of ref. 8).

If y(t) is a continuous function satisfying:
$|y(t)| < M(1+k\int_0^t|y(t)|dt)$ then $|y(t)| < M\,e^{kM\cdot t}\,(t>0)$.

Proof: Let $v(t) = \int_0^t|y(t)|dt$. Then $\dot{v}(t) < M(1+kv(t) \Rightarrow \frac{\dot{v}}{1+kv} < M$
$\rightarrow \log(1+kv) < kM \Rightarrow 1 + kv < e^{kM}$, but $|y| < M(1+kv)$ Q.E.D.

In our case, $y(t) = \|c_2(t) - c_1(t)\|$, $M = \epsilon T$ and $kM = L$, so $\|c_2(t) - c_1(t)\| \leq \epsilon\cdot T\cdot e^{Lt}$. Setting t = T and rewriting t for T, we have

$$*** \qquad \|c_2(t) - c_1(t)\| \leq \epsilon(R)te^{Lt}.$$

By the remark at the top of page (21), T < const. So $\|\tilde{q}-q\| \leq \epsilon(R)\cdot\text{const.}$ Thus $d(q,\tilde{q}) \leq \epsilon(R)\cdot\text{const.}$ and Lemma 3 is proved.

Q.E.D.

<u>Proof of Lemma 4</u>: Recall that in the proof of Carnot's Theorem we produced a c'-smooth map $\theta : \mathbb{R}^n \to M$ which is a c'-diffeomorphism near the origin. We may associate such a θ_r to each distribution W_r for large r. If ρ is sufficiently small, then there is a $\rho' > 0$ s.t. $B_\infty(\rho') \supset \mathbb{B}(\rho)$ and θ_∞ is a diffeomorphism on $B_\infty(\rho')$; also, we may choose $\rho' \to 0$ as $\rho \to 0$. By the estimate *** above, we see that for s on the sphere of radius ρ' in $\mathbb{R}^n$, $\mathrm{dist}(\theta_r(s), \theta_\infty(s)) \leq \epsilon(R)T\, e^{LT}$ if $r > R$. If we take R large enough to insure that $\epsilon(R)T\, e^{LT} < \frac{\rho}{2}$, e.g., then θ_r is homotopic to $\theta_\infty|_{\mathrm{sphere}(\rho')}$ <u>in the complement</u> of $\mathbb{B}(\rho/2) \subset M$. On the otherhand, if there is a point $y \in \mathbb{B}(\rho/2)$ such that $y \notin \theta_r(\text{ball of rad.}\rho')$ then $\theta_r|_{\mathrm{sphere}(\rho')}$ is homotopic in $M - \{y\}$ to the constant map $x_0 = \theta_r(0)$. But $\theta_\infty|_{\mathrm{sphere}(\rho')}$ is not homotopic in M-y to x_0. Thus $\mathbb{B}(\rho/2) \subset \theta_r(\text{ball } \rho')$. Q.E.D.

Lemma 1, (1.1) and (1.2) together now give Theorem 1. Theorem 2 follows easily from the volume estimate ** as follows: Choose a maximal set of disjoint balls (in the Carnot-Carathéodory metric) of radius ϵ. The number N_ϵ, of such balls does not exceed $\dfrac{\mathrm{vol}(B_1(1)}{c^{-1}\epsilon^Q}$. The set of concentric balls of radius 2ϵ cover $B_1(1)$. Each has diameter $\leq 4\epsilon$ and, so, the Hausdorff δ-measure of $B_1(1)$ is at most $\lim\limits_{\epsilon\to 0}(\dfrac{\mathrm{vol}(B_1(1)}{c^{-1}\epsilon^Q} \cdot \epsilon^\delta) = 0$ if $\delta > Q$. Thus $\dim \leq Q$.

Conversely, given any covering $B_1(1)$ by sets of diameter ϵ, there is an associated covering by balls of radius ϵ, so the number, N_ϵ, of sets in the covering satisfies:

$N_\epsilon \cdot c \cdot \epsilon^Q \geq \sum_{i=1}^{N_\epsilon} \mathrm{vol}(i^{th}\ \mathrm{ball}) \geq \mathrm{vol}(B_1(1))$. Thus

$\sum_{\text{covering}} \epsilon^\delta \geq \frac{\mathrm{vol}(B_1(1))}{c\epsilon^Q} \epsilon^\delta$. Taking in f over all coverings by sets of diameter $\leq \epsilon$, then taking $\lim_{\epsilon \to 0}$ gives Hausdorff δ-measure of $B_1(1) = \infty$ if $\delta < Q$. Thus dim $\geq Q$. This proves Theorem 2.

Q.E.D.

Remarks:

a) The estimates above imply more: Hausdorff Q-measure is commensurate with Lebesque measure (on $B_1(1)$):

$$\left(\frac{v_Q}{c2^Q}\right)\mu \leq \mu^Q \leq (c \cdot v_Q)\mu$$

where v_Q = volume of unit ball in $\mathbb{R}^Q$ and μ = Lebesque measure, μ^Q = Q-dimensional Hausdorff measure.

b) Lemma 3 says that the mapping:

infinitesimally transitive distribution $\to$ associated Carnot-Carathéodory metric is continuous in the topologies used there. An invariant of an infinitesimally transitive distribution is the order of this map w.r.t. c'-distance on the left side and Hausdorff distance on the right. The proof of Lemma 3 implicitely contains the estimate: $H(d_r, d_\infty) \leq [c'\text{-dist.}(W_r, W_\infty)]^{1/m}$ if W_∞ satisfies Hormander's condition of order m.

The method described above for approximating (M,W) by a nilpotent Lie group with a left-invariant distribution applies in more general situations. We may extend Theorems 1 and 2 from the free case to the homogeneous case.

Definition: Let $\mathfrak{L} = \mathfrak{L}^1$ be the free Lie algebra of rank r.* For $i = 1,\ldots r$ let $\mathfrak{L}^{i+1} = [\mathfrak{L}^i, \mathfrak{L}]$. An ideal I of $\mathfrak{L}$ is homogeneous if the vector space I is isomorphic to the direct sum of $I\wedge\mathfrak{L}^i / I\wedge\mathfrak{L}^{i+1}$ $i = 1,2\ldots r$. We shall say that the Lie algebra L is homogeneous if it is isomorphic to the quotient of a free Lie algebra by a homogeneous ideal.

We then have the following simple theorem:

Theorem: Any homogeneous Lie algebra admits expanding automorphisms (dilations) (see Dyer).

Proof: $\mathfrak{L}$ admits the obvious canonical family of dilations defined by $\delta_t(Y_i) = tY_i$ and linearity. Homogeneity implies that $\delta_t(I) \subset I$ and so the dilations δ_t act on the quotient $\mathfrak{L}/I = L$.

Q.E.D.

Example: Let $M = \{z,w) : \mathrm{Im}(w) = \|z\|^2,\ z \in \mathbb{C}^n,\ w \in \mathbb{C}\}$. This is a hypersurface (codimension one) in $\mathbb{C}^{n+1}$. The holomorphic tangent space to M has complex dimension n and is spanned by $L_j = \frac{\partial}{\partial z_j} + 2i\ \bar{z}_j \frac{\partial}{\partial w}$ $1 \le j \le n$. (To

---

* That is, there exist r elements $Y_1,\ldots Y_r$ which generate $\mathfrak{L}$ as a Lie algebra, and $\mathfrak{L}$ enjoys the universal mapping property.

see that the L_j are tangent to M, note that $M = \{f = 0\}$ where $f = w - \bar{w} - 2i\|z\|^2$ and $L_j f = 0$.) Let $\bar{L}_j$ denote the conjugate vector fields. Then one may check that $0 = (L_i, L_j) = (\bar{L}_i, \bar{L}_j)$, $(L_i, \bar{L}_j) = \delta_{ij}(-2_i)(\frac{\partial}{\partial w} + \frac{\partial}{\partial \bar{w}})$. If we define $X_j = \mathrm{Re}(L_j)$, $Y_j = \mathrm{Im}(L_j)$, $z = (X_j, Y_j)$, it is easy to check that X_j, Y_j and z together span the real tangent space M at each point and satisfy the Heisenberg commutation relations:

$$(X_j, Y_k) = \delta_{jk} z,$$

all other commutators commute. This Lie algebra g is the real Heisenberg Lie algebra of dimension 2n + 1. It may be written as $\mathcal{L}/I$ where $\mathcal{L}$ = free two-step nilpotent Lie algebra $I = \langle [X_i, X_j], [X_i, Y_j], [X_i, Y_i] - [X_j, Y_j] \rangle$ $1 \le i < j \le n$. Thus we are dealing with a homogeneous Lie algebra.

We will say that a Lie algebra L is <u>homogeneous up to step m</u> if L/L^{m+1} is a homogeneous Lie algebra. Suppose now that the vector fields $X_1, \ldots X_n$ on M satisfy Hormander's condition of order m and are homogeneous up to step m. Let L be the Lie algebra spanned by the X_i's, $g = L/L^{m+1}$, G = simply connected nilpotent Lie group associated to g. With notations as in Theorems 1 and 2, we have the following results:

<u>Theorem 1'</u>: The tantent cone of (M, d_1) is isometric to (G, d_∞).

Theorem 2': The Hausdorff dimension of (M,d_1) is equal to $Q(G)$. Here d_∞ is the Carnot-Carathéodory metric associated to the distribution W_∞ generated by the images of the X_i's in L/L^{m+1}, denoted Y_i.

Proof: The proof is the same as for the free case, once the theorem of Rothschild-Stein has been replaced the following generalization (which is equivalent to the combination of their previously quoted theorem plus their "lifting theorem").

Theorem (Rothschild-Stein; see Goodman): Suppose λ is a partial homomorphism (i.e., a homomorphism "up to step m") from a graded Lie algebra g into the Lie algebra of vector fields on a manifold M which is surjective at a point $p \in M$. Define a map: $g \to M$ by $v \to \exp(\lambda(0))$, and the associated map $\theta : G \to M$. Then the vector fields $\lambda(Y_i)$ may be lifted to vector fields $Y_i + R_i$ on G (locally) where R_i is of local degree ≤ 0.

When the distribution generated by the X_i's is homogeneous up to step m and satisfies Hormander's condition of order m, one defines λ by $\lambda(Y_i) = X_i$ and extends to a partial homomorphism. Since the differential λ_* is bijective at 0 in this case, θ is a local diffeomorphism. The proofs of Theorems 1 and 2 now apply. Q.E.D.

§3. Further questions:

a) The space of horizontal curves is itself of interest. The path space of a smooth manifold M is a Serre fibration over M x M. Is this so for the space of horizontal paths? What can be said about existence and regularity of minimizing horizontal paths?

b) Are there isoperimetric inequalities for the Carnot-Carathéodory metric? Note that Folland and Stein have proven generalized Sobolev inequalities for the situation discussed here, Pierre Pansu has obtained an isoperimetric inequality for the three-dimensional Heisenberg group.

c) Can one reasonably define curvatures and a Laplace operator for the Carnot metrics? One approach might be to define a Laplacian on 1-form through a variational problem: minimize the L^2-norm of the restriction to W of the 1-form w, where w varies within a cohomology class. The associated Euler-Lagrange operator is one candidate for a Laplacian. A "crude" Laplacian may be defined, and the difference of the two, if a zero order operator (Bocher's formula) may be taken to be a mean curvature for the Carnot metric.

Bibliography

[1] Chow, W.L., "Systeme von linearen partiellen differential gleichungen erster ordnung", Math. Ann. 117 (1939), pp. 98-105.

[2] Dyer, J., "A nilpotent Lie algebra with nilpotent automorphism group", Bull. Amer. Math. Soc. 76 (1970), pp. 52-56.

[3] Filippov, A.F., "On certain questions in the theory of optimal control", SIAM Journal on Control, 1 (1962), pp. 76-84.

[4] Folland, G.B., "Applications of analysis on nilpotent groups to partial differential equations", Bull. Amer. Math. Soc., 83 No.5 (1977), pp.912-930.

[5] Gaveau, B., "Principe de moindre action, propagation de la chaleur et estimées sous-elliptiques sur certains groupes nilpotents", Acta Math. 139 (1977), pp. 95-153.

[6] Goodman, Roe, "Nilpotent Lie groups: structure and applications to analysis", Lecture Notes in Mathematics No.562, Springer-Verlag (1970).

[7] Gromov, M., "Groups of polynomial growth and expanding maps", Publications de l'Institut des Hautes Etudes Scientifiques No.53.

[8] Nemytskii and Stepanov, "Qualitative Theory of Ordinary Differential Equations", Princeton Univ. Press (1960).

[9] Pansu, Pierre - to appear.

[10] Rothschild, L.P. and Stein, E.M., "Hypoelliptic differential operators and nilpotent groups", Acta Math. 137 (1976), 247-320.