Microlocal Quasianalyticity for Distributions and Ultradistributions

By

Jan BOMAN*

§ 1. Introduction

Let S be a real analytic submanifold of $\boldsymbol{R}^d$ and let f be a distribution defined in an open set X containing S. Assume that the analytic wave front set of f, $WF_A(f)$, is disjoint from the conormal bundle of S, i.e.

$$WF_A(f) \cap N^*(S) = \emptyset. \tag{1.1}$$

Then the restrictions to S of f and all its derivatives $\partial^\alpha f$ are well-defined distributions on S, so it makes sense to assume

$$\partial^\alpha f \,|_S = 0 \quad \text{for all} \quad \alpha = (\alpha_1, \dots, \alpha_d) \in \boldsymbol{N}^d. \tag{1.2}$$

It was proved in [Bo1] that (1.1) and (1.2) imply that f must vanish in some neighborhood of S. The purpose of this note is to strengthen that result by replacing $WF_A(f)$ with $WF_M(f)$, the wave front set of f with respect to an arbitrary quasianalytic Denjoy-Carleman class C^M (Theorem 1), and by allowing f to be an ultradistribution in the dual of a non-quasianalytic class (Theorem 2).

By a counterexample of M. Sato ([Ka], Note 3.3) (1.1) and (1.2) do not imply that $f = 0$ in a neighborhood of S, if f is only assumed to be a hyperfunction. In fact we have recently proved [Bo2] that Sato's example can be strengthened as follows: for an arbitrary quasianalytic class C^M there exists a hyperfunction in the dual of C^M such that (1.1) and (1.2) hold for some S but the support of f meets S.

In the analytic case there is a well known closely related theorem proved by Hörmander for distributions ([H1] ; [H2], Theorem 8.5.6) and independently by Kawai and Kashiwara for hyperfunctions ([Ka], Theorem 4.4.1), which reads as follows. If the distribution f vanishes on one side of the C^1 hypersurface S and

---

* Department of Mathematics, Stockholm University, S-10691 Stockholm, Sweden.

$(x, \xi) \notin WF_A(f)$, where ξ is one of the two conormals to S at x, then f must vanish in some neighborhood of x. The corresponding statement where $WF_A(f)$ is replaced by $WF_M(f)$, C^M quasianalytic, has also been proved by Hörmander [H4].

The special case of Theorem 2 where f satisfies (1.1) has been considered by Tanabe and Takiguchi [TT].

Let $M = \{M_n\}_{n=1}^{\infty}$ be an increasing sequence of positive numbers and X an open connected subset of $\mathbf{R}^d$. The class $C^M = C^M(X)$ is defined as the set of all $\varphi \in C^{\infty}(X)$ such that for each compact set $K \subset X$ there exists a constant C such that

$$\sup_K | \partial^{\alpha} \varphi | \leq C^{|\alpha|+1} M_{|\alpha|} \quad \text{for all} \quad \alpha \in \mathbf{N}^d.$$

The class C^M is called quasianalytic, if no non-trivial function in C^M vanishes together with all its derivatives at any point. According to the Denjoy-Carleman Theorem the class C^M is quasianalytic if and only if $\Sigma \bar{M}_n/\bar{M}_{n+1} = \infty$, or equivalently, $\Sigma \bar{M}_n^{-1/n} = \infty$, where $\bar{M}_n$ is the largest logarithmically convex minorant of M_n. A class $C^M(\mathbf{R}^d)$ can always be defined by a logarithmically convex sequence M_n. Let M_n be a logarithmically convex sequence such that

$$M_0 = 1, \quad M_n \geq n!, \quad \text{and} \quad M_{n+1} \leq C^{n+1} M_n \tag{1.3}$$

for some C. Those conditions ensure that C^M is closed under multiplication and differentiation and contains the analytic class. The set of Schwartz distributions in X is denoted $\mathscr{D}'(X)$ and the set of distributions with compact support in X is denoted $\mathscr{E}'(X)$. The wave front set with respect to C^M for f, denoted $WF_M(f)$, introduced by Hörmander in [H1], is defined as follows. The Fourier transform of f is denoted $\hat{f}$.

Definition. $WF_M(f)$ is defined as the complement with respect to $X \times (\mathbf{R}^d \backslash 0)$ of the set of (x^0, ξ^0) such that there exists a neighborhood $V \subset X$ of x^0, a conic neighborhood Γ of ξ^0, and a bounded sequence $f_n \in \mathscr{E}'(X)$, where $f_n = f$ in V and

$$| \hat{f}_n(\xi) | \leq C^{n+1} M_n / | \xi |^n \quad \text{for} \quad n = 0, 1, \ldots \text{ and all } \quad \xi \in \Gamma.$$

The requirement that the sequence f_n be bounded is equivalent to the existence of constants C and q such that

$$| \hat{f}_n(\xi) | \leq C(1 + | \xi |)^q$$

for all n (cf. [H2], Prop. 8.4.2).

Theorem 1. *Let M_n be a logarithmically convex sequence satisfying* (1.3), *and*

assume that the class C^M is quasianalytic. Let f be a distribution defined in some neighborhood of the real analytic submanifold $S \subset \mathbf{R}^d$. Assume that the wave front set with respect to C^M for f, $WF_M(f)$, is disjoint from the conormal bundle of S, that is,

$$WF_M(f) \cap N^*(S) = \emptyset, \tag{1.4}$$

and that (1.2) holds. Then $f = 0$ in some neighborhood of S.

The more general statement where f is allowed to be an ultradistribution is proved in Section 7 (Theorem 2).

Note that the condition (1.2) makes sense in Theorem 1, since $WF_M(f)$ is obviously contained in $WF(f)$, the wave front set with respect to C^∞, and $WF(f) \cap N^*(S) = \emptyset$ is sufficient for the restrictions $\partial^\alpha f \mid_S$ to be well-defined.

Here is an outline of the proof of the theorem. To simplify the notation a little we shall first assume that S is a hypersurface. The assumptions of the theorem are preserved under a real analytic change of coordinates, so we may assume that S is contained in the plane $x_d = 0$. We write $x = (x', x_d)$. The problem is local, so it is enough to study a neighborhood of one point, for instance the origin. As in [Bo1] we shall construct a sequence of cut-off functions $\chi_n(x) = \chi_n(x', x_d)$, whose derivatives of order $\leq n$ satisfy good estimates, and study the functions

$$w_n(x_d) = \int \chi_n(x', x_d) f(x', x_d) dx'. \tag{1.5}$$

The sequence χ_n can be chosen to tend to an arbitrary test function, so it will be sufficient to prove that $w_n(x_d)$ tends to zero as $n \to \infty$ for sufficiently small x_d. The assumption (1.2) implies that all derivatives of w_n vanish at the origin, and (1.4) implies that the derivatives $w_n^{(s)}$ satisfy good estimates for $s \leq n$, for s close to n essentially $\mid w_n^{(s)} \mid \leq C^s M_s$. If, as in [Bo1], C^M is the analytic class, we can use Taylor's formula at this point to conclude $\mid w_n(x_d) \mid \leq (C \mid x_d \mid)^n$, which implies $w_n \to 0$ for $\mid x_d \mid < 1/C$. Here we are interested in the case when $M_n^{1/n}/n$ tends to infinity; we must therefore replace the remainder estimate in Taylor's formula by an estimate which takes into account bounds for derivatives $\mid w_n^{(s)} \mid$ of orders less than n also. However, the derivatives $\partial^\alpha \chi_n$ satisfy rather bad estimates for $\mid \alpha \mid /n$ small, so our estimate for $\mid w_n^{(s)} \mid$ will be bad for s/n small. Our substitute for Taylor's formula will therefore be an estimate for $\mid w_n(x_d) \mid$ in terms of bounds for $\mid w_n^{(s)} \mid$ for s in some suitable interval $m \leq s \leq n$. For the same reason special care is needed in the construction of χ_n.

§2. An inequality for Flat Functions

The following lemma and its proof are based on Bang's ideas [Ba] as simplified by Cohen [C]. The lemma can also be deduced from Lemma 1.3.6 in [H2]; see

Remark at the end of this section. The special case $m = 1$ proves of course that the class C^M is quasianalytic if $\sum \bar{M}_k / \bar{M}_{k+1}$ is divergent. However, for a given finite sequence $B_0, \ldots, B_n$ (for example $B_k = n^k$ for $0 \leq k \leq n$) and sufficiently small t, larger values of m often give better estimate for $u(t)$. We shall only need the case when B_k is a logarithmically convex sequence, but we treat the general case, since it requires very little additional effort.

Lemma 1. *Let $u(t)$ be a C^∞ function on $\mathbf{R}$, vanishing for $t \leq 0$, let B_k, $0 \leq k \leq n$, be a sequence of positive numbers, and assume $\sup |u^{(k)}| \leq B_k$ for $0 \leq k \leq n$. Denote by $\bar{B}_k$ the largest logarithmically convex minorant of B_k. Then*

$$| u(t) | \leq 2B_0 \left(\frac{2t}{\sum_m^n \bar{B}_{k-1}/\bar{B}_k} \right)^m , \qquad (2.1)$$

provided the expression within brackets is $\leq 1/2$.

Proof. We first consider the special case when B_k is logarithmically convex; we shall prove that $u(t)$ then satisfies (2.1) with the factor 2 omitted, i.e.

$$| u(t) | \leq B_0 \left(\frac{2t}{\sum_m^n B_{k-1}/B_k} \right)^m , \qquad (2.1')$$

provided the expression within brackets is $\leq 1/2$. Consider first the case $m = 1$. Let $t_1 > 0$ and divide the interval $[0, t_1]$ into n subintervals, $I_1, I_2, \ldots, I_n$, numbered *from right to left*, so that the length of I_k is proportional to B_{k-1}/B_k, that is, $| I_k | = aB_{k-1}/B_k$, where $a = t_1 / \sum_1^n B_{k-1}/B_k$. Writing

$$F(s, k) = \sup_{I_k} | u^{(s)} |$$

we shall prove the estimate

$$F(s, k) \leq (2a)^{k-s} B_s \quad \text{for} \ \ 0 \leq s \leq k \leq n . \qquad (2.2)$$

The proof will be carried out by double induction with ascending k and s. By assumption the estimate (2.2) is true for $k = s$, and if we set $F(s, n+1) = 0$ for all s, (2.2) will be true for $k = n+1$ also. For the induction step we use the trivial estimate

$$F(s, k) \leq F(s, k+1) + | I_k | F(s+1, k) . \qquad (2.3)$$

Let $0 \leq s < k \leq n$ and assume (2.2) is true for $(s, k+1)$ and $(s+1, k)$. Then by (2.3)

$$F(s, k) \leq (2a)^{k+1-s}B_s + \frac{aB_{k-1}}{B_k}(2a)^{k-s-1}B_{s+1}. \tag{2.4}$$

The logarithmic convexity implies that B_{k-1}/B_k is decreasing. Hence

$$\frac{B_{k-1}}{B_k} \leq \frac{B_s}{B_{s+1}},$$

so that (2.4) gives

$$F(s, k) \leq (2a)^{k-s}B_s\left(2a + \frac{a}{2a}\right).$$

If now $2a \leq 1/2$, the induction works, and we obtain for $s = 0$ with $t_1 = t$

$$F(0, k) \leq B_0(2a)^k \leq B_0(2a) = B_0\frac{2t}{\sum_1^n B_{k-1}/B_k}, \tag{2.5}$$

which completes the proof of (2.1′) in the special case when $m = 1$.

To prove (2.1′) for arbitrary m we restrict all considerations to the intervals I_k with $k \geq m$. If we choose

$$a = t/\sum_m^n B_{k-1}/B_k,$$

we obtain from (2.5)

$$|u(t)| \leq F(0, m) \leq B_0(2a)^m,$$

which proves (2.1′).

Finally we consider the case of an arbitrary sequence B_k. Setting $F_k = \sup |u^{(k)}|$ we know by Kolmogorov's inequality that

$$F_k \leq 2F_m^{(n-k)/(n-m)}F_n^{(k-m)/(n-m)}, \quad m < k < n.$$

This means that we can apply (2.1′) to the logarithmically convex sequence $2\bar{B}_k$, and observing that $\bar{B}_0 = B_0$ we then obtain (2.1).

Remark. Lemma 1 can be deduced from Lemma 1.3.6 in [H2] as follows. If a_j has the same meaning as in that lemma, we choose $a_j = cB_{j-1}/B_j$ for $j \geq m$ and $a_j = a_m$ for $1 \leq j \leq m$, where c is the constant determined by the condition $\sum a_j = t$. Then the conclusion of the lemma implies (2.1′).

§3. The Cut-off Functions

Given a compact set $K \subset \mathbf{R}^d$ and an open set $U \supset K$ it is well known how to

construct a sequence χ_n of C^∞ functions with support contained in U, equal to 1 on K, and satisfying estimates

$$\sup | \partial^\alpha \chi_n | \leq C(Cn)^{|\alpha|}, \quad | \alpha | \leq n, \tag{3.1}$$

with C independent of n (see [H2], Section 8.4). In this section we shall construct, for an arbitrary logarithmically convex sequence M_k satisfying (1.3), a sequence χ_n satisfying (3.6) below; this estimate is slightly weaker than (3.1) for $| \alpha | = n$, but stronger for smaller $| \alpha |$.

To motivate the construction below we will first explain why the estimate (3.1) is not sufficient for our purpose. As explained in the introduction we are going to estimate the function (1.5) using Lemma 1. It will be necessary then to have bounds $B_{k,n}$ for the derivatives $\partial^k w_n$, $k \leq n$, that are good enough for

$$\sum_{k=m}^{n} B_{k-1,n}/B_{k,n} \geq c > 0 \tag{3.2}$$

to hold for arbitrarily large m and suitable n. Using a sequence χ_n satisfying (3.1) and assuming f satisfies (1.4) we would be able to prove $| \partial^k w_n | \leq C^{k+1} B_{k,n}$ for $k \leq n$, where $B_{k,n} = \max(n^k, M_k)$, but no better, since those are the estimates valid for a product $\chi_n f$ where $f \in C^M$. And it is easy to see that (3.2) does not hold then, if $M_k = (k \log k)^k$. In fact, if k_0 is the largest integer k such that $k \log k < n$, then $k_0 = (n/\log n)(1+o(1))$ as $n \to \infty$, and

$$\sum_{1}^{n} \frac{B_{k-1,n}}{B_{k,n}} \leq \frac{k_0}{n} + \sum_{k_0}^{n} \frac{1}{k \log k} \leq$$
$$\frac{C}{\log n} + \log \frac{\log n}{\log k_0} = \log \frac{\log n}{\log n - \log \log n} + o(1) \to 0 \quad \text{as} \quad n \to \infty,$$

which contradicts (3.2).

Take a non-negative, even function $v \in C^\infty$ with support in $(-1, 1)$ and with integral equal to 1. It is clearly possible to choose v such that $\int | v'(t) | \, dt \leq 2$. For $0 < \varepsilon < 1$ take

$$\phi(x) = (\sqrt{d}/\varepsilon)^d \prod_{\nu=1}^{d} v(x_\nu \sqrt{d}/\varepsilon).$$

Then ϕ is non-negative, even, $\phi \in C^\infty$, the support of ϕ is contained in an ε-neighbourhood of the origin, U_ε, and

$$\int | \partial_\nu \phi(x) | \, dx = \frac{\sqrt{d}}{\varepsilon} \int | v'(t) | \, dt \leq \frac{2\sqrt{d}}{\varepsilon}, \quad 1 \leq \nu \leq d. \tag{3.3}$$

Let θ_n be the convolution

$$\theta_n(x) = \frac{1}{a_1}\,\phi\left(\frac{x}{a_1}\right)* \ldots * \frac{1}{a_n}\,\phi\left(\frac{x}{a_n}\right), \tag{3.4}$$

where $a_1, \ldots, a_n$ are positive numbers with $\Sigma\, a_k = 1$. For the constants a_k we make the choice

$$a_k = a_k(n) = \frac{1}{s_n}\cdot\frac{M_{k-1}}{M_k},$$

where

$$s_n = \sum_{k=1}^{n} M_{k-1}/M_k. \tag{3.5}$$

Then supp $\theta_n \subset U_\varepsilon$, $\int \theta_n dx = 1$, and for $|\beta| \leq n$ we can compute $\partial^\beta \theta_n$ by applying a first order partial derivative to each of the first $|\beta|$ factors in (3.4), hence

$$\int |\,\partial^\beta\theta_n\,|\,dx \leq Bs_n\frac{M_1}{M_0}\cdot \ldots \cdot Bs_n\frac{M_{|\beta|}}{M_{|\beta|-1}} = (Bs_n)^{|\beta|}M_{|\beta|}.$$

Here the constant B is the number $(2\sqrt{d}\,)/\varepsilon$ appearing in (3.3). For $\varphi \in C_0^\infty(\boldsymbol{R}^n)$ we choose

$$\chi_n = \varphi * \theta_n.$$

Then

$$|\,\partial^{\alpha+\beta}\chi_n\,| = |\,\partial^\alpha\varphi * \partial^\beta\theta_n\,| \leq \sup |\,\partial^\alpha\varphi\,|\int|\,\partial^\beta\theta_n\,|\,dx \leq C_\alpha(Bs_n)^{|\beta|}M_{|\beta|}, \quad |\beta| \leq n. \tag{3.6}$$

We next prove that $\chi_n \to \varphi$ in C_0^∞. The support of χ_n is contained in an ε-neighborhood of the support of φ, and for an arbitrary derivative we have the estimate

$$|\,\partial^\alpha\chi_n - \partial^\alpha\varphi\,| = |\,\theta_n * \partial^\alpha\varphi - \partial^\alpha\varphi\,| \leq \int|\,(\hat{\theta}_n(\xi)-1)\widehat{\partial^\alpha\varphi}(\xi)\,|\,d\xi.$$

Since $\hat{\theta}_n$ is bounded, it is therefore sufficient to verify that $\hat{\theta}_n(\xi)$ tends to 1 uniformly on compact sets. To do this, note that $\hat{\phi}'(0) = 0$ since ϕ is even, and hence

$$|\log\hat{\theta}_n(\xi)| = \Big|\sum_{k=1}^{n}\log\hat{\phi}(a_k\xi)\Big| \leq C\sum_{k=1}^{n}|\,a_k\xi\,|^2 \leq C_1\sum_{k=1}^{n}a_k^2$$

for ξ in a compact set. Since a_k is decreasing and $\Sigma\, a_k = 1$ we have

$$\sum_1^n a_k^2 \le a_1 \sum_1^n a_k = a_1 = \frac{M_0}{M_1 s_n} \to 0, \text{ as } n \to \infty,$$

which proves the claim.

If $\varphi = 1$ in some neighborhood of the compact set K and ε is sufficiently small, then also $\chi_n = 1$ in some fixed neighborhood of K. This justifies the term cut-off functions for χ_n.

It is natural to ask for which sequences $B_{k,n}$ with $B_{0,n} = 1$ for all n there exists a sequence χ_n as above satisfying

$$\sup |\partial^\alpha \chi_n| \le C^{k+1} B_{k,n} \text{ for } |\alpha| = k \le n, \quad n = 0, 1, \dots . \tag{3.7}$$

It is easy to see that the construction above works provided

$$\sup_n \sum_{k=1}^n \bar{B}_{k-1,n}/\bar{B}_{k,n} < \infty; \tag{3.8}$$

here $k \mapsto \bar{B}_{k,n}$ is the largest logarithmically convex minorant of $k \mapsto B_{k,n}$. On the other hand, if (3.8) does not hold and $\chi_n \in C^\infty(\boldsymbol{R})$ is supported in a fixed bounded interval and satisfies (3.7), then Lemma 1 with $m = 1$ shows that some subsequence of χ_n must tend uniformly to zero. Taking $B_{k,n} = k!$ for $k \le n$ and all n we see in particular that cut-off functions satisfying the stronger estimates

$$\sup |\partial^\alpha \chi_n| \le C^{k+1} k!, \quad |\alpha| = k \le n,$$

do not exist.

Repeated convolutions (3.4) with dilation factors a_k proportional to M_{k-1}/M_k were used by Cohen [C] in his proof of the Denjoy-Carleman Theorem; in fact, if $\Sigma M_{k-1}/M_k < \infty$, then θ_n as given by (3.4) converges as $n \to \infty$ to a non-trivial function in C^M with compact support.

§ 4. Estimates for $\widehat{\chi_n f}$

Assume f satisfies the assumptions of the theorem and S is an open subset of the hyperplane $x_d = 0$. Let $x^0 \in S$. The definition of the wave front set with respect to C^M then means that there exists a conic open neighborhood Γ of $\pm dx_d = (0, \dots, 0, \pm 1)$ and distributions f_n, $n = 1, 2, \dots$, with compact support, equal to f in a fixed open neighborhood V of x^0, and constants C and A_0 such that

$$|\hat{f}_n(\xi)| \le C(A_0/|\xi|)^n M_n \text{ for all } \xi \in \Gamma \text{ and all } n, \tag{4.1}$$

and, if f has order q,

$$|\hat{f}_n(\xi)| \le C(1 + |\xi|)^q \text{ for all } n. \tag{4.2}$$

Choose a number b, $0 < b < 1/2$, and a conic neighborhood Γ_1 of $\pm dx_d$ such that

$$\xi \in \Gamma_1 \text{ and } |\eta| < b|\xi| \text{ implies } \xi + \eta \in \Gamma.$$

Let $\varphi \in C_0^\infty(V)$, choose ε, $0 < \varepsilon < 1$, so small that a 2ε-neighborhood of supp φ is contained in V, let $\chi_n = \theta_n * \varphi$ be the sequence of functions constructed above, and let the constant B be chosen as in Section 3. We must now estimate $\widehat{\chi_n f}$.

Lemma 2. *Assume (4.1) and (4.2) hold, and let ε, B, b, Γ_1, and χ_n be chosen as described above. Then*

$$|\widehat{\chi_n f}(\xi)| \le C \frac{A^k(1 + s_n)^k}{(1 + |\xi|)^k} M_k \text{ for all } \xi \in \Gamma_1 \text{ and all } k \le n \tag{4.3}$$

with $A = 2 \max\{2A_0, B/b\}$; here s_n is defined by (3.5). The constant C is independent of k and n, but may depend on φ.

Proof. As in [Bo1] we write

$$|\widehat{\chi_n f}(\xi)| = |\widehat{\chi_n f_k}(\xi)| = |\int \hat{\chi}_n(\eta)\hat{f}_k(\xi - \eta)d\eta| \le \sup_{|\eta| < b|\xi|} |\hat{f}_k(\xi - \eta)| \int |\hat{\chi}_n(\eta)| d\eta$$

$$+ \sup_{|\eta| > b|\xi|}(1 + |\eta|)^{-q}|\hat{f}_k(\xi - \eta)| \int_{|\eta| > b|\xi|} |\hat{\chi}_n(\eta)|(1 + |\eta|)^q d\eta = I + II. \tag{4.4}$$

We first estimate the term I. From (3.6) with $|\beta| = 0$ and $|\alpha| = s$ we get

$$|\hat{\chi}_n(\xi)| \le C_s(1 + |\xi|)^{-s}, \tag{4.5}$$

and with $s = d + 1$ here we see that $\int |\hat{\chi}_n(\eta)| d\eta \le C$. Here and below the letter C will denote constants that are independent of n but may depend on f and on the function φ which was used in the construction of the sequence χ_n. Noting that $|\xi - \eta| \ge |\xi|/2$ when $|\eta| < b|\xi|$ we see that (4.1) implies

$$I \le C \left(\frac{2A_0}{|\xi|}\right)^k M_k \le C \left(\frac{4A_0}{1 + |\xi|}\right)^k M_k \text{ for all } \xi \in \Gamma_1, \ |\xi| > 1, \ k \ge 0.$$

For $|\xi| < 1$ we can use (4.2) and get $I \le C \le C2^k(1 + |\xi|)^{-k}$; hence I is bounded by the right hand side of (4.3) if $A = \max\{4A_0, 2\}$ (note that $B/b > 1$).

We now estimate the term II. With $s = q + d + 1$ in (4.5) and using (4.2) we see that

$$II \le C \text{ for all } \xi. \tag{4.6}$$

Taking $|\alpha| = q+d+1$ and $|\beta| = k$ we can also deduce from (3.6) that

$$|\hat{\chi}_n(\xi)| \le C(1+|\xi|)^{-q} |\xi|^{-d-1-k} (Bs_n)^k M_k \text{ for } k \le n;$$

this leads to

$$II \le C(Bs_n)^k M_k \int_{|\eta|>b|\xi|} \frac{d\eta}{|\eta|^{d+1+k}} = C\left(\frac{Bs_n}{b|\xi|}\right)^k M_k \int_{|\eta|>1} \frac{d\eta}{|\eta|^{d+1+k}}$$

$$\le C_1 \left(\frac{Bs_n}{b|\xi|}\right)^k M_k \text{ for all } \xi \text{ and all } k \le n.$$

Using this estimate for $|\xi| > 1$ and (4.6) for $|\xi| < 1$ we now easily see that II is bounded by the right hand side of (4.3) for all ξ and $k \le n$. This completes the proof of the lemma.

§5. Estimates for w_n

Let π be the projection $(x', x_d) \mapsto x_d$ from $\boldsymbol{R}^d$ to $\boldsymbol{R}$, and let g be a distribution on $\boldsymbol{R}^d$ such that the restriction of π to the support of g is proper. Then the push-forward $\pi_* g$ is a distribution on $\boldsymbol{R}$ defined by $\langle \pi_* g, \varphi \rangle = \langle g, \varphi \circ \pi \rangle$ for $\varphi \in C_0^\infty(\boldsymbol{R})$. Here $\langle \cdot, \cdot \rangle$ is the pairing between distributions and test functions. We now define w_n as $\pi_*(\chi_n f)$; this is the precise meaning of (1.5). Since f satisfies (1.4), w_n must be C^∞, and since f is flat along $S = \{x_d = 0\}$, w_n must be flat at the origin. Restricting to $\xi' = 0$ in (4.3) we obtain

$$|\hat{w}_n(\xi_d)| = |\widehat{\chi_n f}(0, \xi_d)| \le C\frac{A^k(1+s_n)^k}{(1+|\xi_d|)^k} M_k, \quad k \le n. \tag{5.1}$$

Hence

$$\sup |w_n^{(k)}| \le \int |\tau^k \hat{w}_n(\tau)| \, d\tau \le C_1 A^{k+2}(1+s_n)^{k+2} M_{k+2}, \quad k \le n-2.$$

With $B_k = B_{k,n} = C_1 A^{k+2}(1+s_n)^{k+2} M_{k+2}$ we now compute a sum of the kind occurring in Lemma 1:

$$\sum_{m-2}^{n-2} \frac{B_{k-1}}{B_k} = \frac{1}{A(1+s_n)} \sum_m^n \frac{M_{k-1}}{M_k} = \frac{1}{A} \cdot \frac{\sum_m^n M_{k-1}/M_k}{1 + \sum_1^n M_{k-1}/M_k}.$$

Since $\sum M_k/M_{k+1}$ is divergent, we can choose a sequence of numbers $m = m_n$ tending to infinity as $n \to \infty$, such that the quotient is $> 1/2$. Applying Lemma 1 we can now conclude that $|w_n(x_d)| \to 0$ as $n \to \infty$, provided $|x_d| < 1/4A$. But we have also seen that $\chi_n \to \varphi$ in C^∞, hence $w_n(x_d) = \int \chi_n f dx' \to \int f \varphi dx'$ as $n \to \infty$. But $\varphi \in C_0^\infty$ was arbitrary, hence f must vanish in $V \cap \{x; |x_d| < 1/4A\}$. This completes the proof of Theorem 1 in the special case when S is a hypersurface.

Assume now that S is a submanifold of arbitrary dimension r, $0 \le r \le d-1$.

Let S be defined by $x'' = 0$, where $x = (x', x'')$, $x' \in \mathbf{R}^r$, $x'' \in \mathbf{R}^{d-r}$, and define w_n by (1.5) with x_d replaced by x''. Then $w_n \in C^\infty(\mathbf{R}^{d-r})$, and writing $\xi = (\xi', \xi'')$ we have (5.1) with ξ_d replaced by ξ''. Hence

$$| \partial^\alpha w_n | \leq C_1 A^k (1 + s_n)^k M_k \text{ for } |\alpha| + d - r + 1 \leq k \leq n.$$

The definition of the class C^M is obviously invariant under a rotation of the coordinate system; hence directional derivatives of w_n of order $|\alpha|$ satisfy similar estimates. Applying the reasoning above to the restrictions of w_n to lines through the origin in $\mathbf{R}^{d-r}$ with all directions therefore finishes the proof as before.

§ 6. Preliminaries on Ultradistributions

We will now briefly review the definition and basic properties of non-quasianalytic ultradistributions, that is, elements in the dual of a non-quasianalytic class C^L. Such distributions were introduced by Beurling [B] and Roumieu [Ro1], [Ro2]; Beurling's theory was worked out in detail by Björck [Bj]. Extensive studies of non-quasianalytic ultradistributions have later been undertaken by Komatsu [Ko1], [Ko2], [Ko3]. One obtains still larger spaces of distributions by considering the dual of quasianalytic classes; then one speaks about quasianalytic ultradistributions or infrahyperfunctions. Basic properties of spaces of ultradistributions, quasianalytic or not, have been established by Hörmander [H3] and de Reuver [dR]. The largest class of quasianalytic ultradistributions is formed by the hyperfunctions of Sato, which can locally be identified with linear forms on the space of real analytic functions.

Let L_k be a logarithmically convex sequence satisfying (1.3) and assume C^L is non-quasianalytic, that is, $\sum L_k^{-1/k} < \infty$. Let X be an open connected subset of $\mathbf{R}^d$ and denote by $\mathscr{D}_L(X)$ the set of functions in $C^L(\mathbf{R}^d)$ that have compact support in X. Observe that a function $\varphi \in C_0^\infty(\mathbf{R}^d)$ belongs to $C^L(\mathbf{R}^d)$ if and only if the norm

$$\| \varphi \|_{L, r} = \sup_\alpha r^{|\alpha|} L_{|\alpha|}^{-1} \sup | \partial^\alpha \varphi |$$

is finite for some $r > 0$. The set $\mathscr{D}'_L(X)$ is defined as the set of linear forms f on $\mathscr{D}_L(X)$ such that for every compact set $K \subset X$ and every $r > 0$ there exists a constant $C_{r, K}$ such that

$$| f(\varphi) | \leq C_{r, K} \| \varphi \|_{L, r} \tag{6.1}$$

for all $\varphi \in \mathscr{D}_L(X)$ with support in K. Since there exist partitions of unity in $\mathscr{D}_L(X)$, one can define the restriction of an ultradistribution f to an open set $\Omega \subset X$ by restricting the linear form f to the set of φ with support in Ω. The support of f can then be defined as the complement of the largest open set on

which f is equal to zero. The set of elements in $\mathscr{D}'_L(X)$ with compact support is denoted $\mathscr{E}'_L(X)$, and this space can of course be identified with a subspace of $\mathscr{E}'_L(\boldsymbol{R}^d) = \mathscr{E}'_L$. For $f \in \mathscr{E}'_L$ it is clear that $f(\varphi)$ can be defined in a unique way for arbitrary $\varphi \in C^L(\boldsymbol{R}^d)$.

If $B_k \leq L_k$, C^B and C^L both non-quasianalytic, then $\mathscr{D}_B(X)$ is dense in $\mathscr{D}_L(X)$, that is, if $\varphi \in \mathscr{D}_L(X)$, there exists $\varphi_n \in \mathscr{D}_B(X)$, such that $\| \varphi_n - \varphi \|_{L,r} \to 0$ for some r; hence $\mathscr{D}'_L(X)$ is naturally identified with a subspace of $\mathscr{D}'_B(X)$.

Since $C^L(X)$ is a ring, we can define multiplication of $\phi \in C^L(X)$ with $f \in \mathscr{D}'_L(X)$ by $(\phi f)(\varphi) = f(\phi\varphi)$ for $\varphi \in \mathscr{D}_L(X)$.

The classes C^L are stable under real analytic maps as shown in [H2], Proposition 8.4.1. Let Y be another open set and let ρ be a proper real analytic map $Y \to X$. The argument in the cited proposition proves also that the composition map $\varphi \mapsto \varphi \circ \rho$ from $\mathscr{D}_L(X)$ to $\mathscr{D}_L(Y)$ is continuous in the sense that

$$\| \varphi \circ \rho \|_{L,r} \leq C \| \varphi \|_{L,Ar},$$

where C and A depend only on the map ρ and on the support of φ. This proves that the space $\mathscr{D}'_L(X)$ is invariant under real analytic coordinate transformations and that it makes sense to speak about the spaces $C^L(Y)$, $\mathscr{D}'_L(Y)$, etc., where Y is a real analytic submanifold of $\boldsymbol{R}^d$.

The Fourier transform of $\varphi \in \mathscr{D}_L(\boldsymbol{R}^d)$ satisfies an estimate

$$| \hat{\varphi}(\xi) | \leq C_r / \tilde{L}(r |\xi|), \tag{6.2}$$

for some $r > 0$, where $\tilde{L}(\rho)$ is defined by

$$\tilde{L}(\rho) = \sup_k \rho^k / L_k, \quad \rho > 0.$$

For $f \in \mathscr{E}'_L$ the Fourier transform $\hat{f}$ can be defined as $\hat{f}(\xi) = f(e_{-\xi})$, where $e_\xi(x) = \exp(ix \cdot \xi)$. It follows from (6.1) that $\hat{f}$ is a real analytic function (in fact entire) and satisfies the estimate

$$| \hat{f}(\xi) | \leq C_r \tilde{L}(r |\xi|), \tag{6.3}$$

for every $r > 0$.

For further basic properties of ultradistributions we refer to the papers cited above.

We now need to define the standard (C^∞) wave front set for an ultradistribution $f \in \mathscr{D}'_L(X)$ as well as the wave front set of f with respect to some other class C^M, quasianalytic or not.

Definition. Let $f \in \mathscr{D}'_L(X)$. Then $WF(f)$ is defined as the complement with

respect to $X \times (\mathbf{R}^d \backslash 0)$ of the set of (x^0, ξ^0) such that there exists $g \in \mathscr{E}'_L(X)$ with $g = f$ in some neighborhood of x^0, a conic neighborhood Γ of ξ^0, and constants C_n, such that

$$| \hat{g}(\xi) | \leq C_n / | \xi |^n, \text{ for } n = 1, 2, \ldots \text{ and all } \xi \in \Gamma. \tag{6.4}$$

If M_k is another logarithmically convex sequence satisfying (1.3), we define $WF_M(f)$, the wave front set of f with respect to C^M, as the complement with respect to $X \times (\mathbf{R}^d \backslash 0)$ of the set of (x^0, ξ^0) such that there exists a neighborhood $V \subset X$ of x^0, a conic neighborhood Γ of ξ^0, and a sequence $f_n \in \mathscr{E}'_L(X)$, where $f_n = f$ in V, satisfying

$$| \hat{f}_n(\xi) | \leq C^{n+1} M_n / | \xi |^n, \text{ for } n = 0, 1, \ldots \text{ and all } \xi \in \Gamma,$$

and

$$| \hat{f}_n(\xi) | \leq C \tilde{L}(r | \xi |) \tag{6.5}$$

for some C and r independent of n.

If $f \in \mathscr{D}'_L(X)$ and $\phi \in C^L(X)$ we have $WF(\phi f) \subset WF(f)$. Similarly, if $\phi \in C^L(X) \cap C^M(X)$, then $WF_M(\phi f) \subset WF_M(f)$. Using (6.2) and (6.3) one proves these facts in the same way as the corresponding statements for Schwartz distributions (cf. Lemma 8.1.1 and Theorem 8.4.7 in [H2]). Note that $WF(f)$ and $WF_M(f)$ do not depend on the sequence $\{L_k\}$.

For a Schwartz distribution f satisfying (1.4) the restriction $f|_S$ to a smooth submanifold S was defined by Theorem 8.2.4 in [H2]. For hyperfunctions satisfying (1.4) the restriction to a submanifold has also been defined (see e.g. [Ka]). Ultradistributions can be imbedded in the space of hyperfunctions, so we could refer to the theory of hyperfunctions for the definition of $f|_S$ for the ultradistributions considered here. However, we prefer to make a minor extension of Hörmander's construction in order to avoid relying on the theory of hyperfunctions at this point.

If S is an open part of a hyperplane in $\mathbf{R}^d$, say the hyperplane $x_d = 0$, and $f \in C_0^\infty(\mathbf{R}^d)$, then the restriction $f|_S$ satisfies

$$\langle f|_S, \varphi \rangle = \int f(x', x_d) \varphi(x') dx' = (2\pi)^{-d} \int \hat{f}(\xi', \xi_d) \hat{\varphi}(-\xi') d\xi' d\xi_d \tag{6.6}$$

for $\varphi \in C_0^\infty(S)$. Let $f \in \mathscr{D}'_L(\mathbf{R}^d)$ and assume f satisfies

$$WF(f) \cap N^*(S) = \emptyset. \tag{6.7}$$

Using a partition of unity we can then write $f = f_1 + \ldots + f_N + h$, where h vanishes

in a neighborhood of S and each f_ν satisfies (6.4), which in this case means that there is an $\varepsilon > 0$ and constants C_n such that

$$|\hat{f}_\nu(\xi', \xi_d)| \le C_n/|\xi_d|^n, \quad n = 1, 2, \ldots,$$

provided $|\xi'| < \varepsilon|\xi_d|$. Then the integral on the right hand side of (6.6) makes sense for f_ν, so we can take this as our definition of $\langle f_\nu|_S, \varphi \rangle$ and define $f|_S = \Sigma f_\nu|_S$. We need to check that this definition has the right invariance properties and extend it to arbitrary analytic surfaces S. Following Hörmander ([H2], Section 8.2), for a closed subset $E \subset T^*(X)$, conic in the cotangent variable, we denote by $\mathscr{D}'_{L, E}(X)$ the set of $f \in \mathscr{D}'_L(X)$ such that $WF(f) \subset E$, and say that $f_j \to f$ in $\mathscr{D}'_{L, E}(X)$ if $f_j \to f$ in $\mathscr{D}'_L(X)$ (weakly) and in addition

$$\sup_F |\xi|^j| \widehat{\phi f_j}(\xi) - \widehat{\phi f}(\xi)| \to 0, \quad \text{as } j \to \infty,$$

for every closed cone F in $\mathbf{R}^d$ such that $\operatorname{supp}\phi \times F \cap E = \emptyset$. Now, if S is contained in the plane $x_d = 0$, and $f|_S$ is defined as above, it is easy to see that the map $\rho_S: f \mapsto f|_S$ is continuous from $\mathscr{D}'_{L, E}(X)$ into $\mathscr{D}'_L(S)$, and that this is the unique extension of ρ_S from $C_0^\infty(\mathbf{R}^d)$ to $\mathscr{D}'_{L, E}(X)$ with the continuity property just described. This shows that ρ_S, originally defined on $C^\infty(X)$, has a unique continuous extension to $\mathscr{D}'_{L, E}(X)$ also for an arbitrary analytic hypersurface S, and this defines $f|_S$ in the general case. The same reasoning can of course be applied if S is an analytic surface with codimension greater than 1.

Let $g \in \mathscr{E}'_L(\mathbf{R}^d)$, let S be the plane $x_d = 0$ and assume g satisfies (6.7). Let π be the projection $(x', x_d) \mapsto x_d$ as in Section 5. For the proof of Theorem 2 we need to verify that $w(x_d) = (\pi_* g)(x_d)$ must be flat at the origin if g is flat at S. (A distribution satisfying (1.2) is said to be *flat* at S.) By the definition of w we have $\hat{w}(\tau) = \hat{g}(0, \tau)$, hence $w \in C^\infty$ by virtue of (6.7) and $w^{(k)}(0) = (1/2\pi)\int \tau^k \hat{g}(0, \tau)d\tau$. On the other hand, the definition of restriction shows that the Fourier transform of $(\partial_d^k g)|_S$ is

$$\frac{1}{2\pi}\int (i\xi_d)^k \hat{g}(\xi', \xi_d)d\xi_d.$$

Now, the fact that g is flat at S implies that this expression vanishes for all ξ', in particular for $\xi' = 0$, hence $w^{(k)}(0) = 0$.

The verification that χf is flat along S if $\chi \in \mathscr{D}_L(\mathbf{R}^d)$ and $f \in \mathscr{E}'_L(\mathbf{R}^d)$ is flat along S is even easier (see [Bo1], p. 1234).

§7. A Vanishing Theorem for Ultradistributions

We can now state our extension of Theorem 1 to ultradistributions associated to a non-quasianalytic class.

Theorem 2. *Let L_n and M_n be logarithmically convex sequences satisfying* (1.3) *such that $M_n \leq L_n$. Assume that the class C^L is non-quasianalytic and that C^M is quasianalytic. Let $f \in \mathscr{D}'_L(X)$ be an ultradistribution defined in some neighborhood X of the real analytic submanifold $S \subset \mathbf{R}^d$. Assume that the wave front set with respect to C^M for f, $WF_M(f)$, is disjoint from the conormal bundle of S, that is, that* (1.4) *holds. Assume furthermore that f is flat along S in the sense that* (1.2) *holds. Then $f = 0$ in some neighborhood of S.*

The proof of Theorem 2 is essentially parallel to that of Theorem 1; to avoid repetitions we will therefore describe only those parts of the proof that require modifications of the proof given above.

For the cut-off functions χ_n we take $\phi \in \mathscr{D}_L(\mathbf{R}^d)$, construct θ_n with (3.6) as before, and finally set $\chi_n = \varphi * \theta_n$, where $\varphi \in \mathscr{D}_L(\mathbf{R}^d)$. For the proof of an improvement of the estimate (3.6) to be discussed below we shall in fact need to take φ such that $\| \phi \|_{L, r} < \infty$ for *every* $r > 0$. That such functions φ exist is easily seen from the fact that for every logarithmically convex sequence L_k such that $\Sigma L_k^{-1/k} < \infty$ there exists another logarithmically convex sequence B_k such that $\lim L_k^{-1/k}/B_k^{-1/k} = 0$ and $\Sigma B_k^{-1/k} < \infty$, and we can take $\varphi \in \mathscr{D}_B(\mathbf{R}^d)$. We need to prove first of all that $\| \chi_n - \varphi \|_{L, r} \to 0$ for some $r > 0$ (in fact true for any $r > 0$), that is

$$\sup_\alpha r^{|\alpha|} L_{|\alpha|}^{-1} \sup |\partial^\alpha(\chi_n - \varphi)| \to 0 \quad \text{as} \quad n \to \infty. \tag{7.1}$$

Since $\varphi \in \mathscr{D}_L(\mathbf{R}^d)$, we have $\sup |\partial^\alpha \varphi| \leq C^{|\alpha|+1} L_{|\alpha|}$ for some C, hence

$$|\partial^\alpha(\chi_n - \varphi)| \leq |\theta_n * \partial^\alpha \varphi| + |\partial^\alpha \varphi| \leq 2C^{|\alpha|+1} L_{|\alpha|}.$$

If we take $r \leq 1/2C$, this means that

$$r^{|\alpha|} L_{|\alpha|}^{-1} \sup |\partial^\alpha(\chi_n - \varphi)| \leq 2C \cdot 2^{-|\alpha|}$$

for all α. It remains therefore only to recall that $\partial^\alpha(\chi_n - \varphi)$ tends to zero as $n \to \infty$ for each fixed α, which was proved in Section 3.

In estimating $\widehat{\chi_n f}(\xi)$ we shall need an improved version of (3.6). By the choices of φ and ϕ we obtain as in (3.6)

$$|\xi^{\alpha+\beta} \hat{\chi}_n(\xi)| \leq C_r L_{|\alpha|} r^{-|\alpha|} (Bs_n)^{|\beta|} M_{|\beta|}, \quad |\beta| \leq n,$$

or, with a new C_r,

$$(r|\xi|)^k L_k^{-1} |\hat{\chi}_n(\xi)| \leq C_r (Bs_n/|\xi|)^j M_j, \quad j \leq n,$$

for arbitrarily large r. Replacing k by $k+d+1$ and using the fact that L_k is increasing we obtain with still a new C_r

$$|\xi|^{d+1}(r|\xi|)^k L_k^{-1}|\hat{\chi}_n(\xi)| \le C_r(Bs_n/|\xi|)^j M_j, \quad j \le n.$$

Taking supremum over k we finally obtain

$$|\hat{\chi}_n(\xi)| \le C_r|\xi|^{-d-1}\tilde{L}(r|\xi|)^{-1}(Bs_n/|\xi|)^j M_j, \quad j \le n. \qquad (7.2)$$

Lemma 2 is valid for $f \in \mathscr{D}_L'(\mathbf{R}^d)$ with the only change that the assumption (4.2) has to be replaced by (6.5), which means that f_n is a bounded sequence in $\mathscr{E}_L'$. In the proof of Lemma 2 the term I in (4.4) can be treated as before. In estimating the term II we need to replace the factor $(1+|\eta|)^q$ by $\tilde{L}(r|\eta|)$, since we do not have (4.2). Then the term II becomes

$$\sup_{|\eta|>b|\xi|}\tilde{L}(r|\eta|)^{-1}|\hat{f}_k(\xi-\eta)|\int_{|\eta|>b|\xi|}|\hat{\chi}_n(\eta)|\tilde{L}(r|\eta|)d\eta,$$

and using (6.5) we can estimate this expression by a constant times

$$\int_{|\eta|>b|\xi|}|\hat{\chi}_n(\eta)|\tilde{L}(r|\eta|)d\eta.$$

To estimate this integral we use (7.2) and obtain

$$II \le C\int_{|\eta|>b|\xi|}|\eta|^{-d-1}(Bs_n/|\eta|)^j M_j d\eta = C(Bs_n/b|\xi|)^j M_j \int_{|\eta|>1}\frac{d\eta}{|\eta|^{d+j+1}}$$

$$\le C_1(Bs_n/b|\xi|)^j M_j, \quad j \le n,$$

which completes the proof of Lemma 2 for ultradistributions.

To complete the proof of Theorem 2 we argue exactly as in Section 5 with the only difference that we must use the fact that $\chi_n \to \varphi$ in $\mathscr{D}_L(V)$ to be able to conclude $w_n(x_d) \to \int f\varphi dx'$.

References

[B] Beurling, A., Quasianalyticity and general distributions, *Lectures 4 and 5, A.M.S. Summer institute*, Standford, 1961.

[Ba] Bang, Th., Om quasi-analytiske funktioner, *thesis*, Copenhagen, 1946.

[Bj] Björck, G., Linear partial differential operators and generalized distributions, *Arkiv för Mat.*, **6** (1966), 351-407.

[Bo1] Boman, J., A local vanishing theorem for distributions, *C.R. Acad. Sci. Paris* Sér. I, **315** (1992), 1231-1234.

[Bo2] ———, Uniqueness and non-uniqueness for microanalytic continuation of hyperfunctions, *Preprint*, 1996.

[C] Cohen, P., A simple proof of the Denjoy-Carleman theorem, *Amer. Math. Monthly*, **75**

(1968), 26-31.

[H1] Hörmander, L., Uniqueness theorems and wave front sets for solutions of linear differential equations with analytic coefficients, *Comm. Pure Appl. Math.*, **24** (1971), 671-704.

[H2] ———, *The analysis of linear partial differential operators* 1, Springer-Verlag, Berlin, Heidelberg, and New York, 1983.

[H3] ———, Between distributions and hyperfunctions, in *Colloque en l'honneur de Laurent Schwartz, Astérisque*, **131** (1985), 89-106.

[H4] Hörmander, L., Remarks on Holmgren's uniqueness theorem, *Ann. Inst. Fourier (Grenoble)*, **43** (1993), 1223-1251.

[Ka] Kaneko, A., *Introduction to hyperfunctions*, Kluwer Academic Publishers, Dordrecht, Boston, and London, 1988.

[Ko1] Komatsu, H., Ultradistributions, I. Structure theorems and a characterization, *J. Fac. Sci. Univ. Tokyo Section IA*, **20** (1973), 25-105.

[Ko2] ———, Ultradistributions, II. The kernel theorem and ultradistributions with support in a submanifold, *J. Fac. Sci. Univ. Tokyo Section IA*, **24** (1977), 607-628.

[Ko3] ———, Ultradistributions, III. Vector valued ultradistributions and the theory of kernels, *J. Fac. Sci. Univ. Tokyo Section IA*, **29** (1982), 653-717.

[dR] de Roever, J. W., Hyperfunctional singular support of ultradistributions, *J. Fac. Sci. Univ. Tokyo Section IA*, **31** (1984), 585-631.

[Ro1] Roumieu, C., Sur quelques extensions de la notion de distribution, *Ann. Sci. École Norm. Sup.*, **77** Série 3 (1960), 41-121.

[Ro2] ———, Ultradistributions définies sur R^n et sur certaines classes de variétés différentiables, *J. d'Anal. Math.*, **10** (1962-63), 153-192.

[TT] Tanabe, S., and Takiguchi, T., A local vanishing theorem for ultradistributions with analytic parameters, *J. Fac. Sci. Univ. Tokyo Section IA*, **40** (1993), 607-621.