A THREE TERM SUBLEVEL SET INEQUALITY

MICHAEL CHRIST

Abstract. An upper bound is established for the Lebesgue measure of the set on which a certain type of variable coefficient linear combination of measurable functions is small.

Contents

Part 1. Introduction
1. An inequality
2. On the hypotheses

Part 2. Proof of the main theorem
3. Notation and preliminary remarks
4. Some properties of smooth functions
5. Preliminaries related to transversality and curvature
6. Differences
7. Microscale Analysis
8. A two term sublevel set inequality
9. Mesoscale Analysis
10. Macroscale Analysis
11. Conclusion of proof for the nondegenerate case
12. The degenerate case

Part 3. Complements and further results
13. Reformulation of the main hypothesis
14. Supplementary remarks
15. Approximation by regular functions
16. A special case with arbitrarily many summands
17. Another form of degeneracy
References

Part 1. Introduction

1. An inequality

This paper establishes an inequality in differentiable combinatorics. The inequality is intended as a tool for the analysis of integrals of products of irregularly oscillating functions. To formulate it, let $B \subset \mathbb{R}^2$ be a closed ball of finite, positive radius. Let $\bar{B} \supset B$ be open and connected. For each $j \in \{1, 2, 3\}$ let $a_j : \bar{B} \to \mathbb{R}$ be a $C^\omega$ scalar-valued function that

Date: February 14, 2022.
Research supported by NSF grant DMS-1901413.
does not vanish identically. Let \( \varphi_j : \tilde{B} \to \mathbb{R}^1 \) be a nonconstant real analytic mapping. Write \( \mathbf{a} = (a_1, a_2, a_3) \), and \( \Phi = (\varphi_1, \varphi_2, \varphi_3) \). To any ordered triple \( \mathbf{f} = (f_1, f_2, f_3) \) of Lebesgue measurable functions \( f_j : \phi_j(B) \to \mathbb{R} \) and any \( \varepsilon > 0 \) associate the sublevel set

\[
S(\mathbf{f}, \varepsilon) = \{ x \in B : \left| \sum_{j=1}^{3} a_j(x)(f_j \circ \varphi_j)(x) \right| < \varepsilon \}.
\]

Our theme is that under natural and relatively minimal hypotheses, if \( \varepsilon \) is small then the set \( S(\mathbf{f}, \varepsilon) \) has small Lebesgue measure, unless the functions \( f_j \) are themselves small. This theme is expressed by a sublevel set inequality, by which we mean an upper bound of the form

\[
|S(\mathbf{f}, \varepsilon) \cap \{ x : |f \circ \Phi(x)| \geq 1 \}| \leq C\varepsilon^\tau
\]

for some finite, positive constants \( \varepsilon, \tau \) independent of \( \mathbf{f} \). \(|E|\) denotes the Lebesgue measure of a set \( E \).

Upper bounds of this type have arisen in a study \([2]\) of quadrilinear forms \( \int_B \prod_{j=1}^{3} (g_j \circ \varphi_j) \), for which an upper bound in terms of a product of negative order Sobolev norms of the functions \( g_j \) is sought. Scalar-valued sums such as \( \sum_{j=1}^{3} a_j(f_j \circ \varphi_j) \), with three summands, arise as components of the gradients of net phase functions, for fourfold products. The interpretation of our inequality is that if \( g_j \) are rapidly oscillating, then approximate stationarity can occur only on a set of small Lebesgue measure. In this way, an inequality \((1.2)\) provides one of the principal elements of the analysis in the companion paper \([2]\). Simpler sublevel set inequalities have been applied to trilinear oscillatory forms in \([1]\), and subsequently in \([3]\) and \([4]\).

Our main hypothesis is a simple necessary condition for even a very weak form of the inequality \((1.2)\) to hold. We write \( \mathbf{f} \in C^\omega(\Phi(U)) \) to indicate that \( f_j \in C^\omega(\varphi_j(U)) \) for each \( j \in \{1,2,3\} \).

**Main hypothesis.** For any open set \( U \subset \tilde{B} \) and any \( \mathbf{f} \in C^\omega(\Phi(U)) \) that satisfies

\[
\sum_{j=1}^{3} a_j(x)(f_j \circ \varphi_j)(x) = 0 \quad \forall x \in U,
\]

\( f_j \) vanishes identically in \( \varphi_j(U) \) for each \( j \in \{1,2,3\} \).

This hypothesis is equivalently reformulated in a more quantitative way, involving finite order Taylor expansions, in Proposition \((13.1)\).

The main result of this paper is the following theorem. It involves an auxiliary hypothesis formulated below.

**Theorem 1.1.** Let \( \mathbf{a}, \Phi \) be real analytic in a neighborhood \( \tilde{B} \) of a closed ball \( B \subset \mathbb{R}^2 \). Assume that none of the mappings \( \varphi_j \) are constant on \( B \), and that there exists no pair of distinct indices \( i \neq j \in \{1,2,3\} \) for which \( \nabla \varphi_i \) and \( \nabla \varphi_j \) are everywhere linearly dependent. Assume that \( (\mathbf{a}, \Phi) \) satisfies both the main and the auxiliary hypotheses. Then there exist \( C < \infty \) and \( \tau > 0 \), depending only on \( B, \mathbf{a}, \Phi \), such that the sublevel set inequality \((1.2)\) holds for every ordered triple \( \mathbf{f} \) of Lebesgue measurable functions.

Two special cases of Theorem 1.1 were treated in \([1]\): firstly, when the coefficient functions \( a_j \) are all constant, and secondly, when the web in \( \tilde{B} \subset \mathbb{R}^2 \) defined by \( \Phi \) has identically
vanishing curvature. The author has not been able to apply those special analyses to the general case treated here.

Theorem 1.1 treats only a special case of a more general problem. It would be desirable to allow $B \subset \mathbb{R}^D$ for arbitrary $D \geq 2$, to allow $\varphi_j : \mathbb{R}^D \to \mathbb{R}^{d_j}$ for $d_j \geq 1$, and to sum over an arbitrarily large finite set of indices $j$. In the special case in which all mappings $\varphi_j$ are linear, an extension to arbitrarily many summands is developed in Theorem 1.1. Our purpose, besides establishing the specific Theorem 1.1, is to develop an argument that can potentially be extended to these generalizations in future work. Comments about such an extension are ventured in §17.

The auxiliary hypothesis of Theorem 1.1 is a variant of the main hypothesis. For each index $j \in \{1, 2, 3\}$, let $W_j$ be a real analytic vector field in a neighborhood of $B$ that does not vanish identically, and satisfies $W_j(\varphi_j) \equiv 0$.

**Auxiliary hypothesis.** For each index $j \in \{1, 2, 3\}$, for any nonempty open set $U \subset B$, for each $\tau \in \mathbb{R}$, the only $C^\omega$ solution $g = (g_i : i \neq k)$ of

\[
\sum_{i \neq k} a_i \cdot |W_k \varphi_i|^{\tau} (g_i \circ \varphi_i) \equiv 0 \text{ in } U
\]

is the trivial solution $g_i \equiv 0$ for each $i \neq k$.

This hypothesis is independent of the choice of vector fields $W_j$. The special case $\tau = 0$ of the auxiliary hypothesis is equivalent to what we will call the weak auxiliary hypothesis: For any $i \neq j \in \{1, 2, 3\}$, the ratio $a_i / a_j$ cannot be expressed as $(g_i \circ \varphi_i)/(g_j \circ \varphi_j)$ in any nonempty open set, with $g_i, g_j \in C^\omega$ nonvanishing. The weak auxiliary hypothesis is also the special case of the main hypothesis in which one of the three functions $f_j$ is required to vanish identically.

In principle, the proof produces an effective lower bound for the exponent $\tau$ in the conclusion (1.2) in terms of $(\mathbf{a}, \Phi)$, and more specifically, in terms of properties of an auxiliary function defined in terms of $(\mathbf{a}, \Phi)$ in (10.29). But we believe that any lower bound obtained in this way would be far from optimal, and we have neither made an effort to make the proof more efficient in that respect, nor developed a concrete bound for $\tau$.

2. On the hypotheses

The main hypothesis implies that no coefficient $a_j$ vanishes identically on $B$. It does not imply that all $\varphi_j$ are nonconstant, nor that $\nabla \varphi_i(x)$ cannot take the form $b(x) \nabla \varphi_j(x)$, with $b$ scalar-valued, for some indices $i \neq j$.

We say that $(\mathbf{a}, \Phi)$ is nondegenerate in $B$ if all of the following conditions hold. Firstly, each $a_j$ vanishes nowhere. Secondly, for any $i \neq j \in \{1, 2, 3\}$ and any $x \in B$, $\nabla \varphi_i(x)$ and $\nabla \varphi_j(x)$ are linearly independent. Thirdly, the curvature of the web defined by $(\mathbf{a}, \Phi)$ does not vanish identically in $B$. We will prove Theorem 1.1 in the nondegenerate case, then will explain in §12 how the general case follows, essentially as a consequence of the nondegenerate case.

In the real analytic context, under the second of these nondegeneracy conditions, the identical vanishing of the curvature of this web is locally equivalent to linearizability, that is, to the existence of choices of coordinates in which all mappings $\varphi_j$ are simultaneously linear.

The hypothesis of real analyticity of the datum $(\mathbf{a}, \Phi)$ represents a compromise between simplicity and generality; if $\mathbf{a}, \Phi$ are merely $C^\omega$ then a more quantitative variant of the
main hypothesis is required to exclude situations in which there exist smooth \( f \) for which \( \sum_j a_j \cdot (f_j \circ \varphi_j) \) vanishes to infinite order at a point. In [13] an alternative hypothesis appropriate for \( C^\infty \) data is formulated, and shown to be equivalent to the main hypothesis for real analytic data \((a, \Phi)\).

The auxiliary hypothesis can be formulated more directly, without reference to unknowns \( g_j \). Let \((i,j,k)\) be an arbitrary permutation of \((1,2,3)\). Any measurable solution, almost everywhere, of \((1.4)\) in any open set agrees almost everywhere with a real analytic function. Indeed, since the level curves of \( \varphi_i, \varphi_j \) are transverse, analyticity follows directly by restricting the equation

\[
(g_j \circ \varphi_j) = -\left( a_i |W_k \varphi_i|^\tau / a_j |W_k \varphi_j|^\tau \right) (g_i \circ \varphi_i)
\]

to almost any level curve of \( \varphi_i \). Moreover, after replacing \( g_i, g_j \) by smooth functions agreeing with them almost everywhere, \( g_j \) is determined, up to constant multiples, by this equation. The same holds for \( g_i \). If there exists a set of positive measure on which \( g_j \) does not vanish, then it follows that \( g_j \) vanishes only on a set of measure zero; and conversely. Thus \((1.4)\) has a nonzero solution if and only if \( a_i |W_k \varphi_i|^\tau / a_j |W_k \varphi_j|^\tau \) can be expressed as the ratio of a nonvanishing \( C^\omega \) function of \( \varphi_i \) to a nonvanishing \( C^\omega \) function of \( \varphi_j \).

The author believes the auxiliary hypothesis to be superfluous.

**Part 2. Proof of the main theorem**

3. **Notation and preliminary remarks**

\[
\pi : \mathbb{R}^2 \times \mathbb{R}^1 \to \mathbb{R}^2
\]
denotes the projection mapping

\[
\pi(x, t) = x.
\]

(3.1) For any smooth vector field \( V \) in a Euclidean space, \( e^{tV}x = h(t) \) is the solution of the ordinary differential equation \( h'(t) = V(h(t)) \) with initial condition \( h(0) = x \).

\( c, C \) denote constants in \((0, \infty)\) whose values are permitted to change from one occurrence to the next. These depend only on \( a, \Phi \), on choices made in the constructions, and on constants that occur prior to their appearance in the discussion; they do not depend on \( f, \varepsilon \). For quantities \( u, v \in \mathbb{R} \), typically depending on \( f, \varepsilon \), \( u \lesssim v \) means \( u \leq C v \), while \( u \gtrsim v \) means \( u \geq c v \). The three notations \( O(\varepsilon) \), \( \lesssim C \varepsilon \), and \( \lesssim \varepsilon \) all have the same meaning.

In proving Theorem [1.1] it suffices to prove that there exists \( \varepsilon_0 > 0 \) such that the conclusion holds for every \( \varepsilon \in (0, \varepsilon_0) \). We will implicitly assume throughout the discussion that \( \varepsilon \) is sufficiently small.

Since all of our hypotheses are invariant under permutation of the indices \( j \in \{1,2,3\} \), an equivalent formulation is that if \( |f_3(y)| \geq 1 \) for every \( y \in \varphi_3(B) \) then \( |S(f, \varepsilon)| \leq C \varepsilon^\tau \).

It suffices to prove Theorem [1.1] under the additional hypothesis that \( \|f_j\|_{L^\infty} \leq 2 \) for each index \( j \). Indeed, to recover the general case from this base case, assume that \( |f_3(y)| \geq 1 \) for every \( y \in \varphi_3(B) \) and that \( |f_j| < \infty \) almost everywhere for all three indices. For \( n \geq 0 \) and \( j \in \{1,2,3\} \) define

\[
S_{n,j} = \{ x \in S(f, \varepsilon) : |f_i \circ \varphi_i(x)| < 2^{n+1} \text{ for every } i \in \{1,2,3\} \text{ and } |f_j \circ \varphi_j(x)| \geq 2^n \}.
\]

Then \( S(f, \varepsilon) = \bigcup_{n=0}^\infty \bigcup_{j=1}^3 S_{n,j} \). Applying the base case to \( S_{n,j} \subset S(2^{-n}f, 2^{-n} \varepsilon) \) gives \( |S_{n,j}| \leq C(2^{-n} \varepsilon)^\tau \). Summing over \( n, j \) completes the proof.

We assume throughout the proof of Theorem [1.1] that the curvature of the web in \( \tilde{B} \) defined by \( \Phi \) does not vanish identically. This curvature vanishes identically in a connected
open set $U \subset \tilde{B}$, if and only if there exist real analytic coordinates for $U$ and associated real analytic coordinates for $\varphi_j(U)$ in which the three mappings $\varphi_j$ are simultaneously linear.

The analysis is structured in three successive steps, at scales $\varepsilon^1$, $\varepsilon^2$ for an arbitrary choice of $\gamma \in \left(\frac{1}{2}, 1\right)$, and $\varepsilon^0$, respectively, with information established at each of the smaller scales exploited at the next larger scale. These steps are developed in §7 in §9 and in §10 respectively. They are combined in §11 to complete the proof in the nondegenerate case. The general case of Theorem 1.1 is discussed in §12 Preliminary results used in the analysis are developed in earlier sections of the paper.

4. Some properties of smooth functions

Lemma 4.1. Let $K \subset \mathbb{R}^d$ be a connected compact set. Let $\Omega \subset \mathbb{R}^D$ be compact. Let $H : K \times \Omega \rightarrow \mathbb{R}$ be real analytic in a neighborhood of $K \times \Omega$. Write $H_\omega(x) = H(x, \omega)$. Suppose that there exists no $\omega \in \Omega$ for which $H_\omega$ vanishes identically on $K$. Then there exist $\tau > 0$ and $C < \infty$ such that for every $\varepsilon > 0$,

$$\left|\{x \in K : |H_\omega(x)| \leq \varepsilon\}\right| \leq C\varepsilon^\tau.$$  

Lemma 4.2. Let $D, N \in \mathbb{N}$. Let $B$ be a closed ball in $\mathbb{R}^D$ of positive, finite radius. There exist $c, C \in (0, \infty)$ with the following property.

Let $F$ be a real-valued $C^{N+1}$ function defined in a neighborhood of $B$. Suppose that $\|F\|_{C^{N+1}} \leq 1$. Let $\delta > 0$. Suppose that for each $x \in B$,

$$\sum_{0 \leq |\alpha| \leq N} |\partial^\alpha F(x)| \geq \delta.$$  

Then for every $\eta > 0$,

$$\left|\left\{x \in B : |F(x)| \leq \eta\right\}\right| \leq C\eta^\delta C^\tau.$$  

Proofs of Lemma 4.1 and of Lemma 4.2 are sketched in [1].

Lemma 4.3. Let $(x, y) \mapsto f(x, y) \in \mathbb{R}$ be a $C^\omega$ function in a connected open set $\Omega$ in some Euclidean space. Let $K \times K' \subset \Omega$ be compact. There exist $N \in \mathbb{N}$ and $C \in (0, \infty)$ such that for every $y \in K'$,

$$\sup_{x \in K} \sum_{0 \leq |\alpha| \leq N} |\partial^\alpha f(x, y)| \leq C \inf_{x \in K} \sum_{0 \leq |\alpha| \leq N} |\partial^\alpha f(x, y)|.$$  

Proof. Let $K \subset \Omega$ be a connected compact neighborhood of $K \times K'$. Consider the function $g(x, y, t) = f(x + t, y)$ in $K \times B$, where $B$ is a closed ball of sufficiently small positive radius that $(x + t, y) \in \Omega$ whenever $(x, y) \in K$ and $t \in B$.

Expand $g$ in power series with respect to $t$:

$$g(x, y, t) = \sum_{\alpha} t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying

$$f(x + t, y) \equiv \sum_{|\alpha| \leq N} h_\alpha(x, y, t) t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying

$$f(x + t, y) \equiv \sum_{|\alpha| \leq N} h_\alpha(x, y, t) t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying

$$f(x + t, y) \equiv \sum_{|\alpha| \leq N} h_\alpha(x, y, t) t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying

$$f(x + t, y) \equiv \sum_{|\alpha| \leq N} h_\alpha(x, y, t) t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying

$$f(x + t, y) \equiv \sum_{|\alpha| \leq N} h_\alpha(x, y, t) t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying

$$f(x + t, y) \equiv \sum_{|\alpha| \leq N} h_\alpha(x, y, t) t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying

$$f(x + t, y) \equiv \sum_{|\alpha| \leq N} h_\alpha(x, y, t) t^\alpha \partial^\alpha x f(x, y)$$  

with $\partial^\alpha x = \partial^\alpha / \partial x^\alpha$. Applying Theorem 9.1 of Stein and Street [6] to $g$, we conclude that there exist $N$ and $C^\omega$ functions $h_\alpha$ satisfying
in \(K \times B\). By applying \(\partial^\beta/\partial t^\beta\) to both sides of (4.6) for each multi-index \(\beta\) satisfying \(|\beta| = N + 1\), it follows that

\[
(4.7) \quad \sum_{|\beta| = N + 1} |\partial_{x,y}^\beta f(x,y)| \leq C \sum_{|\alpha| \leq N} |\partial_{x,y}^\alpha f(x,y)| \quad \forall (x,y) \in K.
\]

Consider the function

\[
(4.8) \quad u(x,y) = \sum_{|\alpha| \leq N} (\partial_{x,y}^\alpha f(x,y))^2.
\]

It follows from Cauchy-Schwarz and (4.7) that there exists \(C' < \infty\) such that

\[
(4.9) \quad |\nabla_{x,y} u(x,y)| \leq C'u(x,y) \quad \forall (x,y) \in K.
\]

The conclusion (4.4) follows from this differential inequality, using the path connectivity of a neighborhood of \(K\). \(\Box\)

**Lemma 4.4.** Let \(U, U'\) be connected open subsets of Euclidean spaces, and let \(K, K'\) be compact subsets of \(U, U'\), respectively. Let \((x,y) \mapsto F(x,y) \in \mathbb{R}\) be a \(C^\infty\) function in \(U \times U'\).

Define

\[
(4.10) \quad \Sigma = \{y \in U' : F(x,y) = 0 \quad \forall x \in U\}.
\]

Then \(\Sigma\) is a real analytic subvariety of positive codimension. Moreover, there exist \(C < \infty\) and \(\tau > 0\) such that for every \(\varepsilon > 0\) and every \(y \in K' \setminus \Sigma\),

\[
(4.11) \quad \left|\left\{x \in K : |F(x,y)| < \varepsilon\right\}\right| \leq C \text{ distance}(y, \Sigma)^{-C\varepsilon^\tau}.
\]

Lemma 4.4 is proved by combining the preceding lemmas with Lojasiewicz’s theorem. \(\Box\)

### 5. Preliminaries related to transversality and curvature

For each \(k \in \{1, 2, 3\}\) let \(W_k\) be a \(C^\infty\) nowhere vanishing vector field in \(\tilde{B}\) that satisfies \(W_k(\varphi_k) \equiv 0\). For each \(j \in \{1, 2\}\) define \(\psi_j : B \times \mathbb{R} \to \mathbb{R}^2\) by

\[
(5.1) \quad \psi_j(x,t) = (\varphi_j(x), tW_3(\varphi_j(x))).
\]

The linear independence of \(\nabla \varphi_j, \nabla \varphi_3\) at each \(x \in B\) guarantees that \(W_k\varphi_i\) vanishes nowhere whenever \(k \neq i\). Therefore \(\psi_j\) are submersions. By replacing \(\varphi_j\) by \(-\varphi_j\) for \(j = 1\) and/or for \(j = 2\) as necessary, we may assume without loss of generality that \(W_3(\varphi_j) > 0\). We make this assumption throughout the discussion.

For \(j \in \{1, 2\}\), let \(V_j\) be nowhere vanishing \(C^\infty\) vector fields in \(\tilde{B} \times (\mathbb{R} \setminus \{0\})\) that satisfy \(V_j(\psi_j) \equiv 0\).

**Lemma 5.1.** Let \(U \subset \mathbb{R}^2\) be a connected open set. Then \(V_1, V_2\) satisfy the bracket condition in \(U \times (\mathbb{R} \setminus \{0\})\) if and only if the curvature of the web defined by \(\Phi\) does not vanish identically in \(U\).

**Proof.** Choose coordinates in which

\[
(5.2) \quad \varphi_j(x_1, x_2) \equiv x_j \quad \text{for} \quad j = 1, 2.
\]

Then writing \(\{1, 2\} = \{i, j\}\), the vector fields

\[
(5.3) \quad V_j = \partial_{x_i} - (\partial_{x_i} \varphi_3)^{-1} \partial_{x_3} \varphi_3(x) \cdot t\partial_t
\]

satisfy \(V_j(\psi_j) \equiv 0\). Their Lie bracket is

\[
(5.4) \quad [V_1, V_2] = \theta(x) t\partial_t
\]
with
\begin{equation}
\theta(x) = \partial_{x_1} \left( \frac{\partial^2_{x_1 x_2} \varphi_3}{\partial x_2 \varphi_3} \right) - \partial_{x_2} \left( \frac{\partial^2_{x_1 x_1} \varphi_3}{\partial x_1 \varphi_3} \right).
\end{equation}

The vector fields \((\partial_{x_i} \varphi_3)^{-1} \partial^2_{x_i x_j} \varphi_3(x) t \partial_t\) commute with \(\theta(x) t \partial_t\). Therefore up to nonzero constant factors, arbitrary higher-order Lie brackets of \(V_1, V_2\) take the form \(\partial^\alpha \theta(x_1, x_2) \partial_t\) for arbitrary \(\alpha\). Thus for any \(\bar{x} \in \mathbb{R}^2\) and any \(t \neq 0\), the pair \(V_1, V_2\) satisfies the bracket condition at \((\bar{x}, t)\) if and only if \(\theta\) does not vanish to infinite order at \(\bar{x}\). Since \(\theta\) is real analytic, it follows that the bracket condition holds at \((\bar{x}, t)\) if and only if it holds at every point of \(U \times (\mathbb{R} \setminus \{0\})\), and that this is equivalent to \(\theta\) not vanishing identically in \(U\).

Both the bracket condition, and identical vanishing of the curvature of the web, are invariant under diffeomorphism. The bracket condition is also invariant under multiplication of \(V_j\) by nowhere vanishing smooth functions. If the curvature of the web vanishes identically in a neighborhood of some point \(\bar{x}\), then it is possible to change coordinates so that \(\varphi_j(x_1, x_2) \equiv x_j\) for \(j = 1, 2\) and so that \(\varphi_3\) is linear. In that situation, \(\nabla^2 \varphi_3 \equiv 0\), so \(\theta\) vanishes identically.

Suppose on the other hand that \(\theta \equiv 0\) in an open set. Working in coordinates in which \(\varphi_j(x_1, x_2) \equiv x_j\) for \(j = 1, 2\) and with the notations \(\kappa = \varphi_3\), \(\kappa_i = \partial_{x_i} \kappa\), and \(\kappa_{i,j} = \partial^2_{x_i x_j} \kappa\), a calculation gives
\begin{equation}
\theta(x) = (-\kappa_2 \partial_1 + \kappa_1 \partial_2)(\kappa^{-1}_2 \kappa^{-1}_1 \kappa_{1,2});
\end{equation}
recall that \(\kappa_i = W_j(\varphi_3)\) vanishes nowhere by hypothesis when \(\{1, 2\} = \{i, j\}\). Any function annihilated by the vector field \(-\kappa_2 \partial_1 + \kappa_1 \partial_2\) can be expressed in the form \(h \circ \kappa\) for some function \(h\). That is,
\begin{equation}
\frac{-\kappa_{1,2}}{\kappa_1 \kappa_2} = h \circ \kappa.
\end{equation}

Defining \(g\) to satisfy \(g''/g' = h\), this relation becomes
\begin{equation}
\partial^2_{1,2}(g \circ \kappa) \equiv 0.
\end{equation}
Therefore \((g \circ \kappa)(x_1, x_2)\) can locally be expressed in the form \(u(x_1) + v(x_2)\). Thus with \(x = (x_1, x_2)\),
\begin{equation}
\varphi_3(x) = \kappa(x) = g^{-1}(u(x_1) + v(x_2)).
\end{equation}

Thus upon changing variables in \(\mathbb{R}^2\) via the mapping \(x = (x_1, x_2) \leftrightarrow (u(x_1), v(x_2))\) and changing variables in the codomain \(\varphi_3(B)\) so that \(\varphi_3\) is replaced by \(g \circ \varphi_3\), all three mappings \(\varphi_i\) become linear. Thus the web has vanishing curvature. \(\square\)

**Lemma 5.2.** Let \(a_j, \varphi_j \in C^\omega\) for \(j = 2, 3\). Suppose that neither \(a_2, a_3\) vanishes identically, and that \(\alpha = a_2/a_3\) cannot be expressed in any nonempty open subset of \(B\) in the form \((h_2 \circ \varphi_2)/(h_3 \circ \varphi_3)\), with \(h_j \in C^\omega\). Suppose that \(\nabla \varphi_2, \nabla \varphi_3\) are pointwise linearly independent at each point of \(B\). Let
\begin{equation}
\Lambda = \{(x, x') \in B \times B : \varphi_2(x) = \varphi_2(x')\}.
\end{equation}
Define \(B(x, x')\) to be the matrix
\begin{equation}
B(x, x') = \begin{pmatrix}
\alpha(x) W_3 \varphi_2(x) & W_3 \alpha(x) \\
\alpha(x') W_3 \varphi_2(x') & W_3 \alpha(x')
\end{pmatrix}.
\end{equation}

Then the determinant \(\det B(x, x')\) does not vanish identically in any nonempty relatively open subset of \(\Lambda\).
The hypothesis is the special case \( \tau = 1 \) of the auxiliary hypothesis, for a particular permutation of the indices \( j \in \{1, 2, 3\} \).

**Proof.** The hypothesis that \( \nabla \varphi_2, \nabla \varphi_3 \) are pointwise linearly independent is equivalent to \( W_3 \varphi_2 \) vanishing nowhere. If \( \det B \) vanishes identically in some connected relatively open subset of \( \Lambda = \{(x, x') : \varphi_2(x) = \varphi_2(x')\} \) then the ratio \( \frac{W_3 \alpha(x)}{W_3 \varphi_2(x)} \) is locally a function of the form \( h \circ \varphi_2(x) \) with \( h \in C^\omega \). Let \( H \) be an antiderivative of \( -h \). Assuming without loss of generality that \( \alpha > 0 \), this relation can be written as

\[
W_3 \log(\alpha) = W_3(\varphi_2) \cdot (h \circ \varphi_2) = -W_3(H \circ \varphi_2).
\]

Thus \( W_3(\log(\alpha) + (H \circ \varphi_2)) \equiv 0 \) locally, whence \( \log(\alpha) + H \circ \varphi_2 \) is a function of the form \( G \circ \varphi_3 \), contradicting the hypothesis. \( \square \)

Recall that \( \pi : \mathbb{R}^2 \times \mathbb{R}^1 \rightarrow \mathbb{R}^2 \) denotes the mapping \((x, t) \mapsto x\). The auxiliary equation

\[
(5.12) \quad \sum_{j=1}^{2} (a_j \circ \pi) \cdot (g_j \circ \psi_j) = 0,
\]

and the equations obtained by permuting the indices in \( \{1, 2, 3\} \), will arise in the analysis below. The next lemma describes the structure of such nontrivial solutions, if any exist. Proposition [5.1] formulated below, describes the structure of approximate solutions when no exact nontrivial solutions exist.

**Lemma 5.3.** Let \( \varphi_j : \hat{B} \rightarrow \mathbb{R} \) be \( C^\omega \) submersions for \( j \in \{1, 2\} \). Let \( a_j, \beta_j : \hat{B} \rightarrow \mathbb{R} \) be \( C^\omega \). Let \( \psi_j(x, t) = (\varphi_j(x), t\beta_j(x)) \). Suppose that \( \psi_j : \hat{B} \times (0, \infty) \rightarrow \mathbb{R}^2 \) are submersions. Let \( V_j \) be nowhere vanishing \( C^\omega \) vector fields in \( B \times (0, \infty) \) that satisfy \( V_j(\psi_j) \equiv 0 \). Suppose that \( V_1, V_2 \) satisfy the bracket condition at each point of \( B \times (0, \infty) \).

Let \( g = (g_1, g_2) \) be a measurable \( \mathbb{R}^2 \)-valued function that satisfies (5.12) almost everywhere in some nonempty connected open subset of \( B \times (0, \infty) \). Then \( g \in C^\omega \), \( g \) is unique up to constant scalar multiples, and there exist \( \tau \in \mathbb{R} \) and \( h_j \in C^\omega \) such that \( g \) takes the form

\[
(5.13) \quad g_j(y, t) = h_j(y) e^\tau \quad \forall j \in \{1, 2\}.
\]

**Proof.** By induction on nonnegative integers \( n \), define mappings \( \Theta_{n, z} \), whose domains are small neighborhoods of \( 0 \) in \( \mathbb{R}^n \) and whose codomains are \( \mathbb{R}^3 \), as follows. \( \Theta_{0, z} = z \), and

\[
(5.14) \quad \Theta_{n, z}(t_1, \ldots, t_n) = \begin{cases} e^{t_n V_1} \Theta_{n-1, z}(t_1, \ldots, t_{n-1}) & \text{if } n \text{ is odd} \\ e^{t_n V_2} \Theta_{n-1, z}(t_1, \ldots, t_{n-1}) & \text{if } n \text{ is even} \end{cases}
\]

It follows from induction on \( n \) and the relation (5.12) that if \( n \) is odd, then with the notation \( \mathbf{t} = (t_1, \ldots, t_n) \),

\[
(5.15) \quad (g_2 \circ \psi_2)(\Theta_{n, z}(\mathbf{t})) = a b_n(\mathbf{t}, z)
\]

where \( \alpha = g_1(z) \) and

\[
(5.16) \quad b_n(\mathbf{t}, z) = -\frac{a_1(\Theta_{n, z}(t_1, \ldots, t_n))}{a_2(\Theta_{n, z}(t_1, \ldots, t_n))} \frac{a_2(\Theta_{n-1, z}(t_1, \ldots, t_{n-1}))}{a_1(\Theta_{n-1, z}(t_1, \ldots, t_{n-1}))} \cdots \frac{a_1(\Theta_{2, z}(t_1, \ldots, t_2))}{a_2(\Theta_{2, z}(t_1, \ldots, t_2))} \frac{a_1(\Theta_{1, z}(t_1))}{a_2(\Theta_{1, z}(t_1))}.
\]
If $n$ is even, then
\begin{equation}
(5.17)
(g_1 \circ \psi_1)(\Theta_{n,z}(t)) = \alpha b_n(t, z)
\end{equation}
where $\alpha$ is again equal to $g_1(z)$ and
\begin{equation}
(5.18)
b_n(t, z) = \frac{a_2(\Theta_{n,z}(t_1, \ldots, t_n))}{a_1(\Theta_{n,z}(t_1, \ldots, t_{n-1}))} \cdot \frac{a_1(\Theta_{n-1,z}(t_1, \ldots, t_{n-1}))}{a_2(\Theta_{n-1,z}(t_1, \ldots, t_{n-1}))} \cdot \frac{a_1(\Theta_{1,z}(t))}{a_2(\Theta_{1,z}(t))}.
\end{equation}
Each $b_n$ is a real analytic function of $(t, z)$.

Suppose that $g_j$ are measurable and that $g$ satisfies (5.12) almost everywhere in some connected open set $U \subset \mathbb{R}^3$, but that it is not the case that $g$ vanishes almost everywhere in $U$.

Since $V_1, V_2$ are real analytic and satisfy the bracket condition, for any $z \in U$ the set of all $t = (t_1, t_2, t_3)$ for which the Jacobian determinant $J_j(t)$ of the function $t \mapsto \Theta_{3,z}(t)$ satisfies $J_j(t) = 0$ is a proper $C^\infty$ variety. Therefore there exist arbitrarily small $t_0 = (t_{10}, t_{20}, t_{30})$ such that $J_j(t_0) \neq 0$.

Now fix any $z_0 \in U$, and let $t_0$ be as in the preceding paragraph. The Jacobian determinant, with respect to $t$ at $t = t_0$, of
\[ t \mapsto e^{-t_0,1}V_1 e^{-t_0,2}V_2 e^{-t_0,3}V_3 e^{t_1} \psi V_1 e^{t_2} V_2 e^{t_3} V_3 z_0, \]
is likewise nonzero. This mapping sends $t_0$ to $z_0$. Nonvanishing of the Jacobian determinant at $t_0$ is stable under small perturbation of $z_0$.

By the same reasoning as in the discussion above of $g_2 \circ \psi_2 \circ \Theta_{3,z}$, there exists $r > 0$ such that for almost every $z$ satisfying $|z - z_0| < r$, $g_2$ agreeing almost everywhere in $\psi_2(B(z, r))$ with a $C^\infty$ function. The corresponding assertion holds for $g_1$. Thus $(g_1, g_2)$ agrees, almost everywhere in $U$, with a $C^\infty$ $\mathbb{R}^2$-valued function. Moreover, this reasoning proves that $(g_1, g_2)$ is uniquely determined up to multiplication by the constant scalar $\alpha = g_1(z_0)$.

For any $r \in \mathbb{R} \setminus \{0\}$, $\tilde{g}_j(y, t) = g_j(y, rt)$ also defines a solution of (5.12). In combination with the uniqueness up to scalar multiplication, this implies that any smooth solution takes the form (5.13) for $y$ in the domain of $f_j$ and $t \in \mathbb{R}^+$, for some exponent $\tau \in \mathbb{R}$ and some $h_j \in C^\infty$, with a corresponding representation for $t \in \mathbb{R}^-$. Moreover, if a smooth nontrivial solution does exist, then the exponent $\tau$ is unique. \hfill \Box

6. Differences

In this section, we introduce a framework that will be used in the microscale and mesoscale analyses to follow. In the nondegenerate case of Theorem 1.1, we may assume with no loss of generality that $a_3 \equiv 1$, by replacing each coefficient $a_j$ in the definition of $S(f, \varepsilon)$ by $a_j/a_3$. For each $(x, s)$ satisfying $(x, e^{sW_3}x) \in S(f, \varepsilon) \times S(f, \varepsilon)$, we can subtract one inequality from another to cancel terms involving $f_3$, leaving
\begin{equation}
(6.1)
\sum_{j=1}^2 a_j(x) \left[(f_j \circ \varphi_j)(e^{sW_3}x) - (f_j \circ \varphi_j)(x)\right] = O(\varepsilon)
\end{equation}
since
\begin{equation}
(6.2)
a_j(e^{sW_3}x) - a_j(x) = O(|s|) = O(\varepsilon)
\end{equation}
and it is assumed that $f_j = O(1)$. Write (6.1) as
\begin{equation}
(6.3)
\sum_{j=1}^2 a_j(x) (\tilde{g}_j(x, s) = O(\varepsilon)
\end{equation}
with
\[
\begin{cases}
\tilde{\psi}_j(x, s) = (\varphi_j(e^{sW_3}x), \varphi_j(x)) \in \mathbb{R}^2 \\
\tilde{g}_j(y, z) = f_j(y) - f_j(z).
\end{cases}
\]

We will temporarily ignore the structural relation \( \tilde{g}_j(y, z) = f_j(y) - f_j(z) \), will analyze solutions of (6.3), for general functions \( \tilde{g}_j \) of two variables, and then will combine the results of that analysis with the structural information to recover information about \( f_j \). Permuting the indices will produce corresponding information concerning \( f_3 \).

It will be more useful to work in alternative coordinates adapted to the exploitation of asymptotic relations as \( \varepsilon \to 0^+ \). Change variables in domain and range, defining \( \psi_j : B \times [-C, C] \to \mathbb{R}^2 \) and functions \( g_j \) by
\[
\begin{cases}
\psi_j^\varepsilon(x, t) = (\varphi_j(x), \varepsilon^{-1} [\varphi_j(e^{tW_3}x) - \varphi_j(x)]) \\
g_j(y, t) = f_j(y + \varepsilon t) - f_j(y).
\end{cases}
\]

With these definitions,
\[
(g_j \circ \psi_j^\varepsilon)(x, t) = (f_j \circ \varphi_j)(e^{tW_3}x) - (f_j \circ \varphi_j)(x)
\]
and therefore
\[
\sum_{j=1}^{2} a_j(x) (g_j \circ \psi_j^\varepsilon)(x, t) = O(\varepsilon)
\]
for all \((x, t)\) satisfying \((x, e^{tW_3}x) \in S(f, \varepsilon) \times S(f, \varepsilon)\).

By the Cauchy-Schwarz inequality, the set of \((x, t)\) satisfying \((x, e^{tW_3}x) \in S(f, \varepsilon)^2\) has Lebesgue measure \( \gtrsim |S(f, \varepsilon)|^2 \). More generally, for any measurable subset \( S \subset S(f, \varepsilon) \), the set \( S^2 \) defined by
\[
S^2 = \{(x, t) : (x, e^{tW_3}x) \in S \times S\}
\]
satisfies \(|S^2| \gtrsim |S|^2\).

As \( \varepsilon \to 0^+ \), \( \psi_j^\varepsilon \) converges to the limiting mapping
\[
\psi_j(x, t) = \psi_j^0(x, t) = (\varphi_j(x), tW_3 \varphi_j(x))
\]
studied in §5 with
\[
\psi_j^\varepsilon(x, t) = \psi_j(x, t) + O(\varepsilon) \text{ in the } C^N \text{ topology, for every } N.
\]
The function \((x, t, \varepsilon) \mapsto \psi_j^\varepsilon(x, t)\) is well-defined for small \( \varepsilon < 0 \), also, and \((x, t, \varepsilon) \mapsto \psi_j^\varepsilon(x, t)\) is real analytic in a neighborhood of \( \varepsilon = 0 \).

Let \( V_j^\varepsilon \) be nowhere vanishing real analytic vector fields in a neighborhood of \( B \times (0, 1) \) that depend real analytically on \( \varepsilon \) in a neighborhood of \( \varepsilon = 0 \), that agree with those already defined for \( \varepsilon = 0 \), and that satisfy \( V_j^\varepsilon(\psi_j^\varepsilon) \equiv 0 \). Then \( V_j^\varepsilon, V_j^\varepsilon \) also satisfy the bracket condition, for every sufficiently small \( \varepsilon \), since that condition is stable under small perturbations in the \( C^\infty \) topology.

By induction on nonnegative integers \( n \), define mappings \( \Theta_{n, z}^\varepsilon \) for each \( z \in B \), whose domains are small neighborhoods of \( 0 \) in \( \mathbb{R}^n \) and whose codomains are \( \mathbb{R}^3 \), by replacing \( \psi_j, V_j \) by \( \psi_j^\varepsilon, V_j^\varepsilon \) respectively in the construction of \( \Theta_{n, z} \) in §5. Thus \( \Theta_{0, z}^\varepsilon = z \), and
\[
\Theta_{0, z}^\varepsilon(t_1, \ldots, t_n) = \begin{cases}
e^n V_j^\varepsilon \Theta_{n-1, z}^\varepsilon(t_1, \ldots, t_{n-1}) & \text{if } n \text{ is odd} \\
e^n V_j^\varepsilon \Theta_{n-1, z}^\varepsilon(t_1, \ldots, t_{n-1}) & \text{if } n \text{ is even}.
\end{cases}
\]
These satisfy
\begin{equation}
\Theta_{n,z}^\varepsilon = \Theta_{n,z}^0 + O(\varepsilon) \quad \text{in the } C^N \text{ topology for every } N.
\end{equation}

It follows from induction on \( n \) and the relation \((6.7)\) that if \( n \) is odd, and if \( \Theta_{k,z}^\varepsilon(t_1, \ldots, t_k) \in S^2 \) for every \( 0 \leq k \leq n \), then for \( t = (t_1, \ldots, t_n) \),
\begin{equation}
(g_2 \circ \psi_2^\varepsilon)(\Theta_{n,z}^\varepsilon(t)) = \alpha b_n^\varepsilon(t, z) + O(\varepsilon)
\end{equation}
where \( \alpha = g_1(z) \) and
\begin{equation}
b_n^\varepsilon(t, z) = \frac{a_1(\Theta_{n,z}^\varepsilon(t_1, \ldots, t_n))}{a_2(\Theta_{n,z}^\varepsilon(t_1, \ldots, t_n))} \cdots \frac{a_1(\Theta_{n-2,z}^\varepsilon(t_1, \ldots, t_n-2))}{a_2(\Theta_{n-2,z}^\varepsilon(t_1, \ldots, t_n-2))} \cdot \frac{a_1(\Theta_{n-1,z}^\varepsilon(t_1, \ldots, t_n-1))}{a_2(\Theta_{n-1,z}^\varepsilon(t_1, \ldots, t_n-1))}.
\end{equation}
Likewise, if \( n \) is even and \( \Theta_{k,z}^\varepsilon(t_1, \ldots, t_k) \in S^2 \) for every \( 0 \leq k \leq n \), then
\begin{equation}
(g_1 \circ \psi_1^\varepsilon)(\Theta_{n,z}^\varepsilon(t)) = \alpha b_n^\varepsilon(t, z) + O(\varepsilon)
\end{equation}
where \( \alpha = g_1(z) \) and
\begin{equation}
b_n^\varepsilon(t, z) = \frac{a_2(\Theta_{n,z}^\varepsilon(t_1, \ldots, t_n))}{a_1(\Theta_{n,z}^\varepsilon(t_1, \ldots, t_n))} \cdots \frac{a_1(\Theta_{n-2,z}^\varepsilon(t_1, \ldots, t_n-2))}{a_2(\Theta_{n-2,z}^\varepsilon(t_1, \ldots, t_n-2))} \frac{a_1(\Theta_{n-1,z}^\varepsilon(t_1, \ldots, t_n-1))}{a_2(\Theta_{n-1,z}^\varepsilon(t_1, \ldots, t_n-1))}.
\end{equation}
\( b_n^\varepsilon(t, z) \) is defined by \((6.16)\) and/or \((6.14)\) for all \( \varepsilon \) positive or nonpositive, in some neighborhood of \( 0 \). Each \( b_n \) is a real analytic function of \((t, z, \varepsilon)\) in a neighborhood of \( \varepsilon = 0 \), and
\begin{equation}
b_n^\varepsilon(t, z) = b_n^0(t, z) + O(\varepsilon)
\end{equation}
in \( C^N \) norm for every \( N \).

**Lemma 6.1.** Let \( f, \varepsilon, a, \Phi \) satisfy the hypotheses of Proposition \((7.2)\). Assume that there exists a \( C^\omega \) solution \( g^* = (g_{1,2}^*) \) of \( \sum_{j=1}^2 (a_j \circ \pi)(g_j^* \circ \psi_j^0) = 0 \) in some nonempty open subset of \( B \times (0, \infty) \) that does not vanish identically. Let \( S \subset B \times (0, 1) \) be measurable and satisfy \((x, t) \in S \Rightarrow (x, e^{\varepsilon t}V_3 x) \in S(f, \varepsilon)^2 \). Let \( \psi_j^\varepsilon, g_j \) be defined as in \((6.5)\). There exist a scalar \( \alpha \in \mathbb{R} \) and a measurable set \( E \subset S \) of Lebesgue measure \( |E| \geq |S|^C \) such that
\begin{equation}
(g_j \circ \psi_j^\varepsilon)(x, t) = \alpha \cdot (g_j^* \circ \psi_j^0)(x, t) + O(\varepsilon) \quad \forall (x, t) \in E.
\end{equation}

**Proof of Lemma \((6.1)\).** This is a combination of facts established above. Firstly, it was shown in the proof of Lemma \((3.3)\) that up to a nonzero constant factor,
\begin{equation}
b_n^0(t, z) = (g_j^* \circ \psi_j^\varepsilon)(\Theta_{n,z}^\varepsilon(t))
\end{equation}
with \( t = (t_1, \ldots, t_n) \) for odd \( n \) when \( j = 2 \), and for even \( n \) when \( j = 1 \). Secondly, \( b_n^\varepsilon(t, z) = b_n^0(t, z) + O(\varepsilon) \). Thirdly, with \( n = 3 \) for \( j = 2 \) and \( n = 4 \) for \( j = 1 \),
\begin{equation}
(g_j \circ \psi_j^\varepsilon)(\Theta_{n,z}^\varepsilon(t)) = g_1(z)b_n^\varepsilon(t, z) + O(\varepsilon)
\end{equation}
for any \((t, z)\) that satisfies \( \Theta_{k,z}^\varepsilon(t_1, \ldots, t_k) \in S^2 \) for every \( 0 \leq k \leq n \). Fourthly, by repeated application of the Cauchy-Schwarz inequality, there exists \( z \) such that the set of all \( t = (t_1, \ldots, t_4) \) that satisfy \( \Theta_{k,z}^\varepsilon(t_1, \ldots, t_k) \in S^2 \) for every \( 0 \leq k \leq 4 \) has Lebesgue measure \( |S|^C \geq |S|^{2C} \). Fifthly, as a consequence of the bracket condition, the image under \( \Theta_{4,z}^\varepsilon \) of the set of all such \( t \) has Lebesgue measure \( \geq |S|^C \), and likewise the image under \( \Theta_{3,z}^\varepsilon \) of
the set of all \( t' = (t_1, t_2, t_3) \) such there exists \( t_4 \) such that \( t = (t_1, \ldots, t_4) \) is in this same set, have Lebesgue measure \( \geq |S|^C \).

\[ \sum \sum 2 \] 

7. Microscale Analysis

The following hypotheses are in force throughout §7, §9 and §10: \( a_j \in C^\omega(\bar{B}) \) vanish nowhere. \( \varphi_j : \bar{B} \to \mathbb{R} \) are likewise \( C^\omega \) and have pairwise linearly independent gradients at each point of \( B \). The curvature of the web defined by \( \Phi \) does not vanish identically. The functions \( f_j \) are Lebesgue measurable and are \( O(1) \) almost everywhere. These hypotheses are omitted from formulations of lemmas and intermediate propositions. Either the auxiliary hypothesis, or a restricted variant of that hypothesis involving a particular permutation of the indices \( j \in \{1, 2, 3\} \), will be assumed, and this will be indicated. The main hypothesis and the pointwise lower bound \( |f_3| \gtrsim 1 \) will not be assumed until §10.

The microscale analysis works with certain pairs of points \( (x, x') \in S(f, \varepsilon) \times S(f, \varepsilon) \) that satisfy \( |x - x'| = O(\varepsilon) \). It is a hybrid microscale/macroscale analysis, in the sense that while \( x - x' \) is small (if \( \varepsilon \) is small), \( x \) varies over the macroscopic region \( B \). It is at the microscale that the coefficients \( a_j \) can be treated as constants, since

\[ \sum_j a_j(x)(f_j \circ \varphi_j)(x) = \sum_j a_j(x')(f_j \circ \varphi_j)(x) + O(\varepsilon) \text{ when } x - x' = O(\varepsilon). \]

For each \( j \neq k \), \( W_k \varphi_j \) vanishes nowhere; this is a restatement of the linear independence of the pair \( \nabla \varphi_j, \nabla \varphi_k \). Therefore \( W_k \varphi_j \) is either everywhere strictly positive, or everywhere strictly negative. One could replace \( \varphi_j \) by \(-\varphi_j \) to ensure positivity. Thus the absolute value signs on \( W_k \varphi_j \) in the auxiliary hypothesis (7.1), below, are of no real significance.

Lemma 7.1. Assume that for any \( \tau \in \mathbb{R} \), any \( C^\omega \) solution \((h_1, h_2)\) of the equation

\[ \sum_{j=1}^2 a_j |W_3 \varphi_j|^\tau (h_j \circ \varphi_j) = 0 \]

in any nonempty open subset of \( B \) vanishes identically. Then for any \( \sigma > 0 \) there exist \( c,C \in (0, \infty) \) with the following property.

Let \( \varepsilon > 0 \). For each \( j \in \{1, 2, 3\} \) let \( f_j : \varphi_j(B) \to \mathbb{R} \) be Lebesgue measurable and satisfy \( \|f_j\|_{L^\infty} = O(1) \). Let \( S \subset S(f, \varepsilon) \) be measurable. Either

\[ |S| \leq C \varepsilon^\sigma, \]

or there exist a function \( f_1^* \) that is constant on intervals of length \( \varepsilon \), and a measurable set \( S' \subset S \), satisfying

\[ |S'| \geq c|S|^C \]

\[ |f_1^*(\varphi_1(x)) - f_1(\varphi_1(x))| \leq C \varepsilon^{1-\sigma} \forall x \in S'. \]

In this formulation, the phrase “constant on intervals of length \( \varepsilon \)” means that there exists a partition of the domain of \( f_1^* \) into intervals, each of length \( \varepsilon \), such that the restriction of \( f_1^* \) to each such interval is constant. We use this language throughout the remainder of the analysis.

Recall the mappings \( \psi_j \), from a subset of \( \mathbb{R}^3 \) to \( \mathbb{R}^2 \), defined by \( \psi_j(x, t) = (\varphi_j(x), tW_3\varphi_j(x)) \). According to Lemma 5.3, any real analytic solution \( g = (g_1, g_2) \) of the associated equation

\[ \sum_{j=1}^2 a_j(x)(g_j \circ \psi_j)(x, t) = 0 \]

in any nonempty open subset of \( \mathbb{R}^2 \times \mathbb{R} \) takes the form \( g_j(y, t) = h_j(y)t^\tau \) for some \( \tau \in \mathbb{R} \) determined by \( \Phi \). Therefore the hypothesis (7.1) of Lemma 7.1 implies that any real analytic solution \( g \) of this associated equation, in any
nonempty open set, vanishes identically. It is this form of the hypothesis that will be used in the proof.

The hypotheses of Theorem 7.1 are invariant under permutations of the indices \( j \in \{1, 2, 3\} \). Lemma 7.1 is formulated in terms of a subset \( S \) of \( S(f, \varepsilon) \) so that it can be applied three times in succession, with the indices permuted in each application, and with \( S \) initially taken to be \( S(f, \varepsilon) \) and then replaced by a suitable subset upon each application. This successive application will establish the following proposition, which is the main result of this section.

**Proposition 7.2** (Microscale). Let \((a, \Phi)\) satisfy the auxiliary hypothesis. Then for any \( \sigma > 0 \) there exist \( c, C \in (0, \infty) \) with the following property.

Let \( 0 < \varepsilon \leq 1 \). For each \( j \in \{1, 2, 3\} \) let \( f_j : \varphi_j(B) \to \mathbb{R} \) be Lebesgue measurable and satisfy \( \|f_j\|_{L^\infty} = O(1) \). Either
\[
|S(f, \varepsilon)| \leq C\varepsilon^c,
\]
or there exist functions \( f_j^* \) that are constant on intervals of length \( \varepsilon \), and a measurable set \( S \subset S(f, \varepsilon) \), satisfying
\[
|S| \geq c|S(f, \varepsilon)|^C
\]
and
\[
|f_j^*(\varphi_j(x)) - f_j(\varphi_j(x))| \leq C\varepsilon^{1-\sigma} \quad \forall x \in S \quad \forall j \in \{1, 2, 3\}.
\]

To begin the proof of Lemma 7.1, let \( f, \varepsilon, S \) be given. Let \( \varrho_0 \) be a small positive constant. We say that \( y \in \varphi_j(B) \subset \mathbb{R}^1 \) is heavy if
\[
|\{x \in S : \varphi_1(x) = y\}| \geq \varrho_0|S|.
\]
In this definition, \(| \cdot |\) denotes the one-dimensional measure of a subset of \( \mathbb{R}^2 \). Define
\[
S' = \{x \in S : \varphi_1(x) \text{ is heavy}\}.
\]
If \( \varrho_0 \) is chosen to be a sufficiently small constant, depending only on \( B \) and \( \varphi_1 \), it follows that
\[
|S'| \geq \frac{1}{2}|S|.
\]
Indeed, the set of all \( x \in S \) such that \( \varphi_1(x) \) is not heavy has two-dimensional Lebesgue measure \( \leq C\varrho_0|S| \) by Fubini’s theorem, with \( C \) a finite constant that depends only on \( B, \varphi_1 \). Choose and fix such a constant \( \varrho_0 \).

Form
\[
S_0 = \{(x, s) \in B \times [-C\varepsilon, C\varepsilon] : (x, e^{sW_3}x) \in S' \times S'\}.
\]
The measure of \( S_0 \) satisfies the trivial bound \(|S_0| = O(\varepsilon|S|)\). On the other hand, by the Cauchy-Schwarz inequality, \(|S_0| \gtrsim |S'|^2\varepsilon \geq \frac{1}{4}|S|^2\varepsilon\). Let \( S_1 \) be the set of all \((x, s) \in S_0\) that satisfy
\[
\sum_{j=1}^2 |(f_j \circ \varphi_j)(e^{sW_3}x) - (f_j \circ \varphi_j)(x)| \leq \varepsilon^{1-\sigma}.
\]

**Lemma 7.3.** Under the hypotheses of Lemma 7.1, if \(|S_1| \geq \frac{1}{2}|S_0|\) then there exist a function \( f_1^* \) that is constant on intervals of length \( \varepsilon \), and a measurable set \( S_1' \subset S \) satisfying \(|S_1'| \gtrsim |S'|^C\), such that \(|f_1^* \circ \varphi_1 - f_1 \circ \varphi_1| \leq C\sigma\varepsilon^{1-\sigma}\) at each point of \( S_1' \).
Then \( x \) has measure \( |S| \geq \frac{1}{2} |S_1| \). Therefore the set of all \( (x, s) \) with \( |s| \leq \varepsilon \) and \( x \) rich has measure \( \geq \frac{1}{2} |S_1| \). By Fubini’s theorem, the set of all rich \( x \in B \) must have Lebesgue measure \( \geq \frac{1}{2} |S_1| \). Let the hypotheses of Lemma 7.1 be satisfied. For any \( \varepsilon > 0 \), \( \sigma > 0 \), and sufficiently large \( N \), we see that \( \varepsilon_1(x_I) \subseteq \varepsilon_2(x_I) \subseteq \varepsilon_3(x_I) \subseteq \varepsilon_4(x_I) \). Therefore by Fubini’s theorem, \( |(f_1 \circ \varphi_j)(e^{3W_3}x) - (f_1 \circ \varphi_j)(x)| \geq \varepsilon_1 - \varepsilon_2 \). For the remainder of the proof of Lemma 7.1, assume that \( |S_1| < \frac{1}{2} |S_0| \). Let \( S = S_0 \setminus S_1 \), which satisfies \( |S| \geq \frac{1}{2} |S_0| \geq |S|^2 \varepsilon \) and

\[
\sum_{j=1}^2 \left| (f_j \circ \varphi_j)(e^{3W_3}x) - (f_j \circ \varphi_j)(x) \right| > \varepsilon^{1-\sigma} \quad \forall (x, s) \in S.
\]

\[ (7.12) \]

**Lemma 7.4.** Let the hypotheses of Lemma 7.1 be satisfied. For any \( \sigma > 0 \) there exist \( c, C \in (0, \infty) \) with the following property. Let \( E \) be the set of all \( (x, s) \in B \times (0, \varepsilon) \) such that \( x \in S(f, \varepsilon) \), \( e^{3W_3}x \in S(f, \varepsilon) \), and

\[
\sum_{j=1}^2 \left| (f_j \circ \varphi_j)(e^{3W_3}x) - (f_j \circ \varphi_j)(x) \right| > \varepsilon^{1-\sigma}.
\]

\[ (7.13) \]

Then \( |E| \leq C \varepsilon^{c} \).
Before entering into the proof of Lemma 7.3, we introduce a device that will be employed. We aim to apply Lemma 4.11 which states that real analytic functions that do not vanish identically satisfy sublevel set bounds of the desired type. Let \( g_j \) and \( \psi_j^\varepsilon \) be defined as in (6.5). In order to apply Lemma 4.11 a \( C^\omega \) approximant of the function \( (x, t) \mapsto \sum_{j=1}^2 a_j(x)(g_j \circ \psi_j^\varepsilon)(x, t) \) is required. The functions \( b_n^\varepsilon(t, z) \) constructed in the framework introduced above and defined in (6.11) and (6.10) are real analytic, and provide good approximations to \( (g_j \circ \psi_j^\varepsilon)(\Theta_{n,z}^\varepsilon(t)) \), with \( n \) even for \( j = 1 \) and \( n \) odd for \( j = 2 \). But here a complication arises. While \( b_n^\varepsilon(t, z) \) is an analytic function of \( (t, z) \), we are not able to assert that \( b_n^\varepsilon(t, z) \) defines a \( C^\omega \) function of \( \Theta_{n,z}^\varepsilon(t) \), as we would like to do, since even in the case \( n = 3 \) when the domain and codomains of \( (t \mapsto \Theta_{n,z}^\varepsilon(t)) \) have equal dimensions, the mapping \( t \mapsto \Theta_{n,z}^\varepsilon(t) \) need not be invertible; indeed, its Jacobian determinant vanishes at \( t = (0, 0, 0) \). Worse yet, for \( n = 4 \) the dimensions of domain and codomain are unequal. To sidestep this difficulty we will lift \( \sum_{j=1}^2 a_j(x) (g_j \circ \psi_j^\varepsilon)(x, t) \) to a function of \( t = (t_1, \ldots, t_4) \).

To motivate the details of this lifting, suppose temporarily that \( \sum_{j=1}^2 (a_j \circ \pi) \cdot (g_j \circ \psi_j^\varepsilon) \) were to vanish identically on some open set, rather than merely being \( O(\varepsilon) \) on some measurable set. We could then conclude three exact functional relations. For \( t = (t_1, t_2, t_3, t_4) \in \mathbb{R}^4 \), let \( t' = t'(t) = (t_1, t_2, t_3) \in \mathbb{R}^3 \). First, just as for the limiting case \( \varepsilon = 0 \) in the proof of Lemma 5.3

\[
(7.14) \quad (a_1 \circ \pi)(\Theta_{3,z}^\varepsilon(t)) \cdot b_4^\varepsilon(t, z) + (a_2 \circ \pi)(\Theta_{3,z}^\varepsilon(t)) \cdot b_3^\varepsilon(t', z) = 0
\]

for all sufficiently small \( |t_4| \) whenever \( \Theta_{3,z}^\varepsilon(t') \in U \). Secondly, \( b_3^\varepsilon(t', z) \) would depend only on \( \psi_2(\Theta_{3,z}^\varepsilon(t')) \). Thirdly, \( b_4^\varepsilon(t, z) \) would likewise depend only on \( \psi_1(\Theta_{3,z}^\varepsilon(t)) \). Conversely, if these three relations were to hold for all \( t = (t', t_4) \) with \( |t_4| \) sufficiently small, and with \( t' \) in a connected open set in which \( D\Theta_{3,z}^0 \) was locally invertible, and if \( t \mapsto (b_3^\varepsilon(t', z), b_4^\varepsilon(t, z)) \) did not vanish identically in that open set, then there would exist a \( C^\omega \) solution \( (g_1, g_2) \) of \( \sum_{j=1}^2 (a_j \circ \pi) \cdot (g_j \circ \psi_j^\varepsilon) = 0 \) in a nonempty open set that did not vanish identically there.

The conjunction of these three conditions on \( (b_3^\varepsilon, b_4^\varepsilon) \) can be equivalently expressed as the vanishing on some nonempty open set in the \( \mathbb{R}_+^4 \) space of a real analytic function that is defined solely in terms of the datum \((a, \Phi)\), and is defined in an open set that depends only on this datum. Indeed, (7.14) already expresses the vanishing of such a function. To express the other two conditions in the same way, construct a family of \( C^\omega \) vector fields \( U_{3,z}^\varepsilon \) in an appropriate open subset of \( \mathbb{R}^3 \) that depend analytically on the parameters \( z, \varepsilon \) in a neighborhood of \( \varepsilon = 0 \), do not vanish identically in any nonempty open subset of \( \mathbb{R}^3 \) for any \( z, \varepsilon \), and have the property that \( b_3^\varepsilon(t', z) \) depends only on \( \psi_2(\Theta_{3,z}^\varepsilon(t')) \) if and only if \( U_{3,z}^\varepsilon(b_3^\varepsilon) \equiv 0 \). Such a family exists, and can be constructed using elementary row operations on matrices. Likewise, there exists an \( \mathbb{R}^2 \)-valued \( C^\omega \) vector field \( U_{4,z}^\varepsilon \), defined in an appropriate open subset of \( \mathbb{R}^4 \) depending analytically on \( (z, \varepsilon) \), such that \( b_4^\varepsilon(t, z) \) depends only on \( \psi_1(\Theta_{4,z}^\varepsilon(t)) \) if and only if \( U_{4,z}^\varepsilon(b_4^\varepsilon) \equiv 0 \) in the appropriate domain. By squaring the left-hand side of (7.14), \( U_{3,z}^\varepsilon(b_3^\varepsilon(t'), t') \), and \( U_{4,z}^\varepsilon(b_4^\varepsilon)(t) \), then adding these three squares, we obtain a real analytic function with the desired property.

A further reformulation is needed to discuss approximate relations for functions \( g_j \circ \psi_j^\varepsilon \circ \Theta_{n,z}^\varepsilon \) that are approximately equal to \( (b_3^\varepsilon, b_4^\varepsilon) \), but are not necessarily differentiable. \( U_{4,z}^\varepsilon(b_4^\varepsilon)(t) \) vanishes identically as a function of \( t \) if and only if \( b_4^\varepsilon(e^{\omega U_{3,z}^\varepsilon}(t)) - b_4^\varepsilon(t) \) vanishes identically as a function of \( (t, s') \), for \( s' \) in any small ball centered at 0 in \( \mathbb{R}^2 \). There is a corresponding statement for \( U_{3,z}^\varepsilon(b_3^\varepsilon) \). Thus the conjunction of the three conditions in question on \( (b_3^\varepsilon, b_4^\varepsilon) \) is equivalent to the identical vanishing, in a neighborhood of 0, of the
function of \((t, s, s')\) defined by

\[
(a_1 \circ \pi)(\Theta^\varepsilon_{4,z}(t)) \cdot b^\varepsilon_4(t, z) + (a_2 \circ \pi)(\Theta^\varepsilon_{2,z}(t)) \cdot b^\varepsilon_2(t', z)]^2 \]

\[
+ [b^\varepsilon_0(e^{sU^\varepsilon_{1,z}}(t', z)) - b^\varepsilon_0(t', z)]^2 + [b^\varepsilon_0(e^{sU^\varepsilon_{1,z}}(t, z)) - b^\varepsilon_0(t, z)]^2.
\]

This concludes the explanation for the somewhat artificial construction that follows.

**Proof of Lemma 7.4.** Let \(f, \varepsilon\) be given, and let \(S(f, \varepsilon) \subset B\) be the associated sublevel set. Form the functions \(g_j\) and mappings \(\psi_j\), as defined in (6.5). These satisfy

\[(g_j \circ \psi_j)(x, t) = (f_j \circ \varphi_j)(e^{\varepsilon t} x) - (f_j \circ \varphi_j)(x).
\]

Define

\[
S^\varepsilon = \{(x, t) \in B \times (0, 1] : x \in S(f, \varepsilon) \text{ and } e^{\varepsilon t} x \in S(f, \varepsilon)\},
\]

which satisfies \(|S^\varepsilon| \geq |S(f, \varepsilon)|^2\). Let \(I\) be a compact subinterval of \((0, 1]\). Introduce auxiliary variables \(s \in \mathbb{R}^3\) and \(s' \in \mathbb{R}^2\), varying over small neighborhoods of the origin in \(\mathbb{R}^3\) and in \(\mathbb{R}^2\), respectively. Write \(t = (t_1, t_2, t_3, t_4) \in \mathbb{R}^4\) and \(t' = t'(t) = (t_1, t_2, t_3)\) as above. Let \(S^\dagger\) be the set of all \((z, t, s, s')\), such that \(t \in I^4, |s|, |s'| \leq c,\) and

\[
\begin{cases}
z, \Theta^\varepsilon_{n,z}(t_1, \ldots, t_n) \in S^\varepsilon \text{ for } 1 \leq n \leq 4 \\
\Theta^\varepsilon_{4,z}(e^{sU^\varepsilon_{2,z}}(t)), \Theta^\varepsilon_{4,z}(e^{s'U^\varepsilon_{2,z}'}(t)) \in S^\varepsilon.
\end{cases}
\]

We claim that \(|S^\varepsilon| \geq |S^\varepsilon|^C \geq |S(f, \varepsilon)|^{2C}\) for a certain exponent \(C < \infty\). To justify this, consider the simpler claim that the set of all \((z, t_1, t_2)\) such that \(z \in S^\varepsilon, \Theta^\varepsilon_{1,z}(t_1) \in S^\varepsilon\), and \(\Theta^\varepsilon_{2,z}(t_2) \in S^\varepsilon\) has Lebesgue measure \(\geq |S^\varepsilon|^C\). Equivalently, \(z \in S^\varepsilon, w_1 = e^{t_1 V^\varepsilon_1}(z) \in S^\varepsilon,\) and \(w_2 = e^{t_2 V^\varepsilon_2}(w_1) \in S^\varepsilon\). We have already observed that by the Cauchy-Schwarz inequality, the set of all \((w_1, t_2)\) such that \(w_1 \in S^\varepsilon\) and \(e^{t_2 V^\varepsilon_2}(w_1) \in S^\varepsilon\) has Lebesgue measure \(\geq |S^\varepsilon|^2\). Therefore the set \(S\) of all \(w_1 \in S^\varepsilon\) with the property that \(e^{t_2 V^\varepsilon_2}(w_1) \in S^\varepsilon\) for a set of values of \(t_2\) having one-dimensional Lebesgue measure \(\geq |S^\varepsilon|^2\) satisfies \(\tilde{S} \geq |S^\varepsilon|^2\). By Cauchy-Schwarz, the set of all \((z, t_1)\) such that \(z \in \tilde{S}\) and \(e^{t_1 V^\varepsilon_1}w_1 \in \tilde{S}\) has Lebesgue measure \(\geq |\tilde{S}|^2 \geq |S^\varepsilon|^4\). By Fubini’s theorem, then, the set of all \((z, t_1, t_2)\) satisfying the desired three inclusions has Lebesgue measure \(\geq |S^\varepsilon|^4 \cdot |\tilde{S}| \geq |S^\varepsilon|^6\). This establishes the simpler claim. The original claim is proved by repeating this same argument multiple times.

For any \(z\), define

\[
S^\dagger_z = \{(t, s, s') : (z, t, s, s') \in S^\dagger\}.
\]

By Cauchy-Schwarz and Fubini’s theorem, there exists \(\tilde{\varepsilon}\) satisfying

\[
|S^\dagger_z| \geq |S(f, \varepsilon)|^{2C}.
\]

Choose and fix such a parameter \(\tilde{\varepsilon}\).

There exists a scalar \(\alpha \in \mathbb{R}\) such that for each \((t, s, s') \in S^\dagger_z,\)

\[
\begin{cases}
(g_1 \circ \psi^\varepsilon_j)(\Theta^\varepsilon_{4,z}(t')) = ab \tilde{\varepsilon} + O(\varepsilon) \\
(g_2 \circ \psi^\varepsilon_j)(\Theta^\varepsilon_{2,z}(t)) = ab \tilde{\varepsilon} + O(\varepsilon).
\end{cases}
\]

Since each \(f_j = O(1)\), likewise \(g_j = O(1)\). The functions \(b_n\) are products of the coefficients \(a_n\) and of their inverses, evaluated at certain points; these vanish nowhere. Consequently \(|\alpha| = O(1)\).
On the other hand, a lower bound \( \sum_{j=1}^{2} |(f_j \circ \varphi_j)(e^{W_3} x) - (f_j \circ \varphi_j)(x)| > \varepsilon^{1 - \sigma} \) is a hypothesis of Lemma 7.4. This bound ensures that
\[
(7.21) \quad |(g_1 \circ \psi_1^\dagger)(\Theta_{3,z}^\dagger(t'))| + |(g_2 \circ \psi_2^\dagger)(\Theta_{4,z}^\dagger(t))| \geq \varepsilon^{1 - \sigma} \quad \forall \ (t, s, s') \in S^4_z.
\]
Therefore
\[
(7.22) \quad |\alpha| \geq \varepsilon^{1 - \sigma}.
\]
The lower bound (7.22), and in particular its improvement by a factor \( \varepsilon^{-\sigma} \) over \( \varepsilon \) itself, will be essential.

Since
\[
(7.23) \quad \psi_1^\dagger(\Theta_{3,z}^\dagger(e^{U_z^{x,0}}(t'))) \equiv \psi_1^\dagger(\Theta_{3,z}^\dagger(t')),
\]
and since the corresponding statement holds for \( \psi_2^\dagger(\Theta_{4,z}^\dagger(e^{U_z^{x,0}}(t))) \), it follows that
\[
(7.24) \quad b_3^\dagger(e^{U_z^{x,0}}(t')) = b_3^\dagger(t') + O(\varepsilon) \quad \text{and} \quad b_4^\dagger(e^{U_z^{x,0}}(t)) = b_4^\dagger(t) + O(\varepsilon)
\]
for all \((z, t, s, s') \in S^4_z\). Since \( b_n^\dagger = b_n^0 + O(\varepsilon) \), and since \(|\alpha| = O(1)\), we conclude that the function \( F_\varepsilon \) defined by
\[
(7.25) \quad F_\varepsilon(t, s, s') = [(a_1 \circ \pi)(\Theta_{4,z}^0(t)) \cdot (b_1^0(\Theta_{4,z}^0(t)))
\]
\[
+ (a_2 \circ \pi)(\Theta_{4,z}^0(t)) \cdot (b_2^0(\Theta_{4,z}^0(t''))) \cdot (b_4^0(\Theta_{4,z}^0(t')))]^2
\]
\[
+ [b_3^0(e^{U_z^{x,0}}(t')) - b_3^0(t')]^2 + [b_3^0(e^{U_z^{x,0}}(t)) - b_4^0(t)]^2
\]
satisfies
\[
(7.26) \quad \alpha \cdot F_\varepsilon(t, s, s') = O(\varepsilon) \quad \forall \ (t, s, s') \in S^4_z.
\]
The quantity \( F_\varepsilon(t, s, s') \) is a \( C^\omega \) function of \((z, t, s, s')\). The parameter \( z \) can be taken to vary over a compact set. Recall that the hypothesis (7.1) of Lemma 7.4 ensures that any real analytic solution \( g^* = (g_1^*, g_2^*) \) of \( \sum_{j=1}^{2} (a_j \circ \pi) \cdot (g_j^* \circ \psi_j) \equiv 0 \) in any nonempty open set must vanish identically. Therefore there exists no \( z \) for which the function \( F_\varepsilon \) vanishes identically in a nonempty open set. By Lemma 4.1, these properties guarantee that there exist \( \tau, C \in (0, \infty) \) such that for every \( z \) and every \( \delta > 0 \),
\[
(7.27) \quad \left| \left\{ (t, s, s') : |F_\varepsilon(t, s, s')| < \delta \right\} \right| \leq C \delta^\tau.
\]
Applying this inequality with \( \delta = O(\varepsilon |\alpha|^{-1}) = O(\varepsilon^\sigma) \), we conclude that \(|S^4_z| = O(\varepsilon^\sigma)\) and therefore that \( |S(f, \varepsilon)| = O(\varepsilon^c) \) for an exponent \( c > 0 \) that depends on \( \sigma \). This concludes the proof of Lemma 7.4.

8. A TWO TERM SUBLEVEL SET INEQUALITY

We formulate an intermediate result, whose proof is implicit in that of Lemma 7.4 and which will be a key element of the mesoscale analysis, below. It concerns two term sublevel set inequalities, for functions of two variables, in a three-dimensional ambient space. We use the notation
\[
(8.1) \quad S(g, \varepsilon) = \{(x, t) \in B \times [-1, 1] : \left| \sum_{j=1}^{2} b_j(x) (g_j \circ \psi_j)(x, t) \right| < \varepsilon \}.
\]
Proposition 8.1. Let $B \subset \mathbb{R}^2$ be a closed ball of positive, finite radius, and let $\tilde{B}$ be an open neighborhood of $B$. Let $b_j : \tilde{B} \to \mathbb{R}$ be nowhere vanishing $C^\omega$ functions. Let $\varphi_1, \varphi_2 : \tilde{B} \to \mathbb{R}^1$ be $C^\omega$ submersions whose gradients are everywhere linearly independent. Let $\beta_j : B \to \mathbb{R}$ be nowhere vanishing $C^\omega$ functions. Let $\psi_j$ be the associated mappings $\psi_j(x,t) = (\varphi_j(x), t\beta_j(x))$. Assume that there exist nowhere vanishing $C^\omega$ vector fields $V_j$ that satisfy $V_j(\psi_j) \equiv 0$, and satisfy the bracket condition at each point of $\tilde{B} \times (\mathbb{R} \setminus \{0\})$.

Suppose that there exists no nonzero $C^\omega$ solution $g^* = (b_1^*, b_2^*)$ of the equation

$$
\sum_{j=1}^2 b_j(x)(g_j^* \circ \psi_j)(x,t) = 0
$$

in any nonempty open subset of $\tilde{B} \times (\mathbb{R} \setminus \{0\})$. Then for each $\sigma > 0$ there exist $C < \infty$ and $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0]$ and any measurable functions $g_j : B \times [-1,1] \to \mathbb{R}$ there exists $S \subset S(g, \varepsilon)$ satisfying $|S| \geq |S(g, \varepsilon)|^C$ such that for each $j \in \{1,2\} \cup \{j\}^*$, $|g_j(\psi(x))| \leq C\varepsilon^{1-\sigma}$ for every $x \in S$.

More generally, let $N \geq 0$, let $B' \subset \mathbb{R}^N$ be an open ball centered at 0, and let $b_j^* : \tilde{B} \times B' \to \mathbb{R}$ and $\varphi_j^* : \tilde{B} \times B' \to \mathbb{R}$ be $C^\omega$ functions. Set $b_j^*(x) = b_j(x,s)$ and $\varphi_j^*(x) = \varphi_j(x,s)$. Suppose that $((b_j^*, \varphi_j^*) : j \in \{1,2,3\})$ satisfies the hypotheses of Proposition 8.1. Then there exists a neighborhood $B'' \subset B'$ of 0 such that for each $s \in B''$, $((b_j^s, \varphi_j^s) : j \in \{1,2,3\})$ satisfies the conclusions of Proposition 8.1 with constants independent of $s$.

9. Mesoscale Analysis

In the author’s view, the mesoscale step is the decisive one in the analysis of sums $\sum_j a_j \cdot (f_j \circ \varphi_j)$ with variable coefficients $a_j$. It works with certain pairs of points $(x,x') \in S(f, \varepsilon) \times S(f, \varepsilon)$ that satisfy $|x - x'| = O(\delta)$, for a certain $\delta$ that is large relative to $\varepsilon$ but small relative to 1. It is a hybrid analysis, in the same sense as the microscale analysis. At the mesoscale, the variation of the coefficients $a_j$ creates significant effects.

Choose, and fix for the remainder of the proof of Theorem 8.1, an auxiliary exponent $\gamma \in (\frac{3}{2}, 1)$. Given any $\varepsilon > 0$, set $\delta = \varepsilon^\gamma$. Partition $\mathbb{R}^1$ into intervals $J$, each having length $\delta$.

Proposition 9.1 (Mesoscale). Let $(a, \Phi)$ satisfy the auxiliary hypothesis. For any $\rho > 0$ there exist $c, C \in (0, \infty)$ with the following property.

Let $\varepsilon > 0$. For each $j \in \{1,2,3\}$ let $f_j : \varphi_j(B) \to \mathbb{R}$ be Lebesgue measurable and satisfy $\|f_j\|_{L^\infty} \leq 1$. Let $S \subset S(f, \varepsilon)$ be measurable. Either

$$
|S| \leq C\varepsilon^\gamma,
$$

or there exist a measurable set $S \subset S$, and for each $j$, functions $f_j^*$ that are affine on intervals $J$ of length $\varepsilon^\gamma$, such that

$$
|S| \geq c|S|^C
$$

and

$$
|f_j^*(\varphi_j(x)) - f_j(\varphi_j(x))| \leq C\varepsilon^{1-\rho} \forall x \in S \forall j \in \{1,2,3\}.
$$

Moreover, the derivatives $(f_j^*)'$ satisfy

$$
|(f_j^*)'| \leq C\varepsilon^{-\rho}.
$$
The relation (9.3) asserts an approximation to $f_j$ on intervals of lengths $\delta$. If the parameters are chosen so that $\rho < 1 - \gamma$ then the upper bound $O(\varepsilon^{1-\rho})$ is of the form $O(\delta^\kappa)$ with $\kappa$ strictly greater than 1. Thus (9.3) provides a quantitative differentiable approximation to $f_j$ at scale $\delta$, on a set $S$ of significantly large Lebesgue measure.

Throughout the proof of Proposition 9.1 we maintain the relation $\delta = \varepsilon^{\gamma}$. The exponent $\gamma \in (\frac{1}{2}, 1)$ remains fixed throughout the discussion.

Let $\rho > 0$ be given. We begin by applying Proposition 7.2 to the datum $(f, S, \varepsilon)$, with a parameter $\sigma$ that depends on $\rho$ and is to be chosen below. Proposition 7.2 has two alternative conclusions. If the first conclusion holds then $|S| \leq C\varepsilon^c$, as desired. Therefore we may assume that the second conclusion holds. This provides a measurable set $S' \subset S$ satisfying $|S'| \gtrsim |S|^C$, and a function $f^* = (f_1^*, f_2^*, f_3^*)$ such that each $f_j^*$ is constant on intervals of length $\varepsilon$, and satisfies $|f_j^* \circ \varphi_j(x) - f_j \circ \varphi_j(x)| = O(\varepsilon^{1-\sigma})$ for each index $j$ and every point $x \in S'$. To conclude the proof of Proposition 9.1 it suffices to show that if $\sigma$ is chosen sufficiently small as a function of $\rho$ then there exist $c, C \in (0, \infty)$ such that for any measurable set $\tilde{S} \subset S(f^*, \varepsilon^{1-\sigma})$ there exist $\tilde{S}' \subset \tilde{S}$ and $\tilde{f} = (f_1^*, f_2^*, f_3^*)$ with each $f_j^*$ affine on intervals of length $\delta$, and with derivatives satisfying $|(f_j^*)'| = O(\varepsilon^{\rho})$, such that

\[ |\tilde{S}'| \geq c |\tilde{S}|^C \]

and for each index $j$,

\[ |f_j^* \circ \varphi_j(x) - f_j \circ \varphi_j(x)| \leq C\varepsilon^{1-\rho} \forall x \in \tilde{S}'. \]

Henceforth, write $f_j$ in place of $f_j^*$, and $S$ in place of $\tilde{S}$, to simplify notation. Thus the domain $\varphi_j(B)$ of each $f_j$ is partitioned into intervals, each of length $\varepsilon$, with $f_j$ constant on each of those intervals.

We claim that it suffices to treat the case in which the functions $f_j$ are nearly constant at the mesoscale, as well as at the microscale, that is,

\[ f_j(y) - f_j(y') = O(\delta^{1-\sigma}) \]

whenever $y, y'$ lie in a common interval $J$, among those intervals of length $\delta$ into which we have partitioned $\varphi_j(B)$. Indeed, since $S \subset S(f, \varepsilon^{1-\sigma}) \subset S(f, \delta)$ provided that $\sigma$ is chosen to be sufficiently small, Proposition 7.2 can be applied once more, but now with $\varepsilon$ replaced by $C\delta$. The conclusion is that either $|S| = O(\delta^c) = O(\varepsilon^{\gamma c})$, or there exist $\overline{S} \subset S$ satisfying $|\overline{S}| \gtrsim |S|^C$ and functions $f_j^{**}$ that are constant on intervals of length $\delta$ and satisfy

\[ f_j^{**} \circ \varphi_j(x) - f_j \circ \varphi_j(x) = O(\delta^{1-\sigma}) \forall x \in \overline{S} \]

for each $j \in \{1, 2, 3\}$. In the former case, in which $|S| = O(\varepsilon^{\gamma c})$, the first alternative conclusion (9.4) of Proposition 9.1 has been established, and the proof of that proposition is complete.

Consider the latter case. The local constancy of $f_j^{**}$, together with the inequality (9.8), imply that the functions $f_j$ satisfy (9.7) with $S$ replaced by its subset $\overline{S}$. We do not replace $f_j$ by $f_j^{**}$; the functions $f_j$ remain constant on intervals of length $\varepsilon$ but are not necessarily constant on any larger intervals. Note also that near constancy at the mesoscale, as expressed in (9.7), involves a bound $O(\delta^{1-\sigma})$, not $O(\varepsilon^{1-\sigma})$. Our present goal is to show that either $|\overline{S}| = O(\varepsilon^c)$, or there exist $S \subset \overline{S}$ satisfying $|S| \gtrsim |\overline{S}|^C$ and functions $f_j$ satisfying (9.3) and (9.4).
By dividing through by the nowhere vanishing analytic coefficient $a_3$, we may assume without loss of generality that $a_3 \equiv 1$. Continue to denote by $W_3$ some nowhere vanishing real analytic vector field in $\tilde{B}$ that annihilates $\varphi_3$. Form

$$S_1 = \{(x, s) \in B \times [-\delta, \delta] : (x, e^{sW_3}x) \in \tilde{S} \times \tilde{S}\},$$

which satisfies $|S_1| \gtrsim |\tilde{S}|^2 \delta$. Either the set of all points with $s > 0$ contributes at least half of the Lebesgue measure of $S_1$, or the set of all points with $s < 0$ does so. We discuss only the first case; the analysis will apply equally well to the second. Replace $\tilde{S}_1$ by $S_1 = \{(x, s) \in \tilde{S}_1 : s > 0\}$.

For each $(x, s) \in S_1$,

$$\sum_{j=1}^{2} \left[ a_j \cdot (f_j \circ \varphi_j)(e^{sW_3}x) - a_j \cdot (f_j \circ \varphi_j)(x) \right] = O(\varepsilon);$$

the terms involving $f_3$ cancel since $\varphi_3(e^{sW_3}x) = \varphi_3(x)$ and $a_3(e^{sW_3}x) = 1 = a_3(x)$. Writing

$$a_j(e^{sW_3}x) - a_j(x) = sW_3a_j(x) + O(s^2)$$

and dividing by $s$ gives

$$\sum_{j=1}^{2} \left[ (W_3a_j)(x)(f_j \circ \varphi_j)(x) + a_j(e^{sW_3}x) \left[ (f_j \circ \varphi_j)(e^{sW_3}x) - (f_j \circ \varphi_j)(x) \right] s^{-1} \right]$$

$$= O(\varepsilon|s|^{-1})$$

since $|s|^2 \leq \varepsilon^{2\gamma} \leq \varepsilon$ because $\gamma > \frac{1}{2}$. Writing $a_j(e^{sW_3}x) = a_j(x) + O(s) = a_j(x) + O(\delta)$ and invoking the upper bound (9.7) to obtain $(f_j \circ \varphi_j)(e^{sW_3}x) - (f_j \circ \varphi_j)(x) = O(\delta^{1-\sigma})$ gives

$$\sum_{j=1}^{2} \left[ (W_3a_j)(x)(f_j \circ \varphi_j)(x) + a_j(x) \left[ (f_j \circ \varphi_j)(e^{sW_3}x) - (f_j \circ \varphi_j)(x) \right] s^{-1} \right]$$

$$= O(\varepsilon|s|^{-1}) + O(\delta^{2-\sigma}|s|^{-1}),$$

which is again $O(\varepsilon|s|^{-1})$ provided that $\sigma$ is chosen to be sufficiently small, since $\delta^2 = \varepsilon^{2\gamma}$ and $\gamma > \frac{1}{2}$.

The quantity $(f_j \circ \varphi_j)(e^{sW_3}x)$ can be simplified for all $(x, s)$ in a relatively large set. Indeed,

$$\varphi_j(e^{sW_3}x) = \varphi_j(x) + sW_3\varphi_j(x) + O(s^2).$$

The domain of $f_j$ has been partitioned into intervals of length $\varepsilon$, with $f_j$ constant on each of these intervals. Therefore

$$f_j(\varphi_j(e^{sW_3}x)) = f_j(\varphi_j(x) + sW_3\varphi_j(x))$$

unless $\varphi_j(e^{sW_3}x)$ lies within distance $O(\delta^2) = O(\varepsilon^{2\gamma})$ of one of the endpoints of one of these intervals. The set of all $(x, s)$ for which this happens, has measure $O(\delta \cdot \varepsilon^{-1} \cdot \varepsilon^{2\gamma}) = O(\varepsilon^{2\gamma-1}\delta)$.

The exponent $2\gamma - 1$ is strictly positive by design. Let $c_0, \rho > 0$ be small constants that satisfy $2c_0 < \rho < \rho/2$, $1 - \gamma - \rho > 0$, and $2c_0 < 2\gamma - 1$. If $|\tilde{S}| \leq \varepsilon^{-c_0}$ then the proof of Proposition 9.1 is complete. If not, then recall that $|S_1| \gtrsim |\tilde{S}|^2 \delta \geq \varepsilon^{2c_0} \delta$. For small $\varepsilon > 0$, $\varepsilon^{2\gamma-1}\delta \ll \varepsilon^{2c_0} \delta$ is consequently negligible relative to $|S_1|$. Therefore for all $x$ in a set $S_2 \subset S_1$ of Lebesgue measure $\gtrsim \delta \varepsilon^{2c_0}$, (9.15) holds for both indices $j = 1, 2$. 


Define $S_3$ to be the set of all $(x, s) \in S_2$ satisfying $|s| \geq \varepsilon^6 \delta$. Then $|S_2 \setminus S_3| = O(\varepsilon^6 \delta)$, which is negligible relative to $|S_2| \gtrsim \varepsilon^{2c_0} \delta$. Thus $|S_3| \gtrsim |S_2| \gtrsim \varepsilon^{2c_0} \delta$, and (9.15) holds for every $(x, s) \in S_3$. Therefore

$$
(9.16) \quad \sum_{j=1}^{2} \left[ (W_3a_j)(x)(f_j \circ \varphi_j)(x) + \alpha_j(x) \left[ f_j(\varphi_j(x) + sW_3\varphi_j(x)) - (f_j \circ \varphi_j)(x) \right] s^{-1} \right] = O(\varepsilon |s|^{-1}) \forall (x, s) \in S_3.
$$

Defining

$$
(9.17) \quad \tilde{F}_j(y, s) = s^{-1}(f_j(y + s) - f_j(y)),
$$

this relation can be rewritten as

$$
(9.18) \quad \sum_{j=1}^{2} \left[ (W_3a_j)(x)(f_j \circ \varphi_j)(x) + \alpha_j(x) \cdot W_3\varphi_j(x) \cdot \tilde{F}_j(\varphi_j(x), sW_3\varphi_j(x)) \right] = O(\varepsilon^{1-\gamma-\varepsilon}) \forall (x, s) \in S_3.
$$

Substituting $F_j(x, t) = \tilde{F}_j(x, \delta t)$ and writing $\psi_j(x, t) = (\varphi_j(x), tW_3(\varphi_j)(x))$, as in the microscale analysis, this becomes

$$
(9.19) \quad \sum_{j=1}^{2} \left[ (W_3a_j)(x)(f_j \circ \varphi_j)(x) + \alpha_j(x) W_3\varphi_j(x) \cdot (F_j \circ \psi_j)(x, t) \right] = O(\varepsilon^{1-\gamma-\varepsilon}) \forall (x, t) \in S_4
$$

where $S_4 = \{(x, t) \in B \times (\varepsilon^6, 1) : (x, \delta t) \in S_3\}$ satisfies

$$
|S_4| = \delta^{-1}|S_3| \gtrsim \varepsilon^{2c_0}
$$

and also

$$
(9.21) \quad |S_4| \gtrsim |S|^2.
$$

In comparison to the framework introduced in §6 and exploited in §7, a simplification is that the variant mappings $\psi^0_j$ have been sidestepped so that only their limits $\psi_j = \psi^0_j$ need be dealt with. On the other hand, new terms $W_3a_j \cdot (f_j \circ \varphi_j)$ have arisen at the mesoscale, and these are not negligible. There are now effectively four unknown functions $f_1, F_1, f_2, F_2$, rather than the three with which the analysis began.

The terms $W_3a_j(x)(f_j \circ \varphi_j)(x)$ are independent of $t$. To exploit this we introduce

$$
(9.22) \quad S^2 = \{(x, t, t') \in B \times (0, 1] \times (0, 1] : (x, t) \in S_4 \text{ and } (x, t') \in S_4\}
$$

and subtract to obtain

$$
(9.23) \quad \sum_{j=1}^{2} a_j(x)W_3\varphi_j(x) \left[ (F_j \circ \psi_j)(x, t') - (F_j \circ \psi_j)(x, t) \right] = O(\varepsilon^{1-\gamma-\varepsilon}) \forall (x, t, t') \in S^2.
$$

Thus we arrive at a sublevel problem in the 4-dimensional $(x, t, t')$ space, associated to the mappings $(x, t, t') \mapsto \Psi_j(x, t, t') = (\psi_j(x, t'), \psi_j(x, t))$, and with each quantity $(F_j \circ \psi_j)(x, t') - (F_j \circ \psi_j)(x, t)$ regarded as a function of $\Psi_j(x, t, t')$. The terms in (9.19) involving $f_j \circ \varphi_j$ have been eliminated.

It is natural to seek to invoke a higher-dimensional analogue of Proposition 8.1. However, in contrast to the microscale analysis, this particular higher-dimensional sublevel set
problem is degenerate, in the sense that vector fields $U_j$ in $\mathbb{R}^4$ that annihilate $\Psi_j$ do not satisfy the bracket condition. Instead, the submanifolds defined by constancy of $t'/t$ form a foliation of $\mathbb{R}^4$ by three-dimensional leaves invariant under these vector fields.

To see this degeneracy, for each $0 \neq r \in \mathbb{R}$, define $\mathcal{H}_r$ to be the set of all $(x, t, t')$ satisfying $t' = rt$. Define

$$\Psi_j(x, t, t') = (\psi_j(x, t), \psi_j(x, t'))$$

and

$$= \left((\varphi_j(x), tW_3\varphi_j(x)), (\varphi_j(x), t'W_3\varphi_j(x))\right) \in \mathbb{R}^2 \times \mathbb{R}^2.$$

Since the first coordinate $\varphi_j(x)$ of $(\varphi_j(x), tW_3\varphi_j(x))$ is identically equal to the first coordinate of $(\varphi_j(x), t'W_3\varphi_j(x))$, each $\Psi_j$ takes values in a three-dimensional subspace of $\mathbb{R}^2 \times \mathbb{R}^2$. Thus $\Psi_j$ may be regarded as mappings from $\mathbb{R}^4$ to $\mathbb{R}^3$.

For any $r \neq r'$ and any $j \in \{1, 2\}$, $\Psi_j(\mathcal{H}_r) \cap \Psi_j(\mathcal{H}_{r'}) = \emptyset$. Indeed, for any $(x, t, t') \in \mathcal{H}_r$, the ratio of the second coordinate $t'W_3\varphi_j(x)$ to the second coordinate $tW_3\varphi_j(x)$ of $\psi_j(x, t')$ equals $t'/t = r$. Thus $t'/t$ is determined by $\Psi_j(x, t, t')$.

For each $r \in [\varepsilon^\theta, \varepsilon^{-\theta}]$ define

$$F_{j,r}^\sharp(y, t) = F_j(y, rt) - F_j(y, t)$$

and

$$S_r^\sharp = \{(x, t) \in \mathbb{R}^3 : (x, t, rt) \in S^\sharp_r\}.$$ 

With the notations

$$\tilde{a}_j(x) = a_j(x)W_3\varphi_j(x)$$

and $\eta = \varepsilon^{1-\gamma-\theta}$, one has

$$\sum_{j=1}^{2} \tilde{a}_j(x) (F_{j,r}^\sharp \circ \psi_j)(x, t) = O(\eta) \ \forall (x, t) \in S_r^\sharp.$$ 

The inequality (9.28) relates $F_{1,r}^\sharp$ to $F_{2,r}^\sharp$, only when $r' = r$; this expresses the degeneracy of this problem. Thus we have a one-parameter family of sublevel sets, parametrized by $r$, in a three-dimensional domain, of the type addressed by Proposition 8.1. Each of these sublevel sets will be analyzed in the same way as in the microscale analysis, with the parameter $r$ subsequently taken into account. While the functions $F_{j,r}^\sharp$ that appear in (9.28) depend on $r$, the coefficients $\tilde{a}_j$, the mappings $\psi_j$, and the inequality itself do not involve $r$. Thus any conclusions about $F_{j,r}^\sharp$ gleaned from (9.28) hold uniformly in $r$.

The linear form $\sum_{j=1}^{2} (\tilde{a}_j \circ \pi)(g_j \circ \psi_j)$ in (9.28) is not identical to the form $\sum_{j=1}^{2} (a_j \circ \pi)(g_j \circ \psi_j)$ that was encountered in §7 because the coefficients $\tilde{a}_j$ are different from $a_j$. However, since $W_3\varphi_j | W_3\varphi_j |^\tau = \pm |W_3\varphi_j |^{\tilde{\tau}}$ with $\tilde{\tau} = \tau + 1$, the hypothesis (7.1) for the form with coefficients $\tilde{a}_j$ is equivalent to the corresponding hypothesis for coefficients $\tilde{a}_j$. Therefore Proposition 8.1 can be applied to (9.28), to conclude that for each $r$, either $|S_r^\sharp| = O(\eta^\gamma)$ or there exists a measurable subset $S_r' \subset S_r^\sharp$ satisfying $|S_r'| \gtrsim |S_r^\sharp|^{C'}$ such that

$$F_{j,r}^\sharp(\psi_j(x, t)) = O(\eta^{1-\gamma}) \ \forall (x, t) \in S_r' \ \forall j \in \{1, 2\}.$$ 

Recall that $\eta = \varepsilon^{1-\gamma-\theta}$. Let $c_1 > 0$ be small. If

$$\left| \left\{ r : |S_r^\sharp| \gtrsim (\varepsilon^{1-\gamma-\theta})^{c_1} \right\} \right| \leq \varepsilon^{c_1}$$

then we conclude that

$$|S_4| = O(\varepsilon^{c_1} + \eta^c) = O(\varepsilon^{c_1} + \varepsilon^{(1-\gamma-\theta)c}).$$
By (9.21), this implies an upper bound $|S| = O(\varepsilon^{c'})$ for some $c' > 0$, completing the proof of Proposition 9.1.

Suppose instead that (9.29) does not hold. Assuming that $\varepsilon \leq \varepsilon_0$ and that $\varepsilon_0$ is sufficiently small, the set of all $(y, t, r)$ satisfying

$$F_j(y, rt) - F_j(y, t) = O(\eta^{1-\sigma})$$

for each $j \in \{1, 2\}$ consequently has Lebesgue measure $\geq \varepsilon^{c''}$ with $c_2 = (1 - \gamma - \varrho) \cdot c \cdot c_1 / 2$. Therefore

$$F_j(y, t') - F_j(y, t) = O(\eta^{1-\sigma})$$

for all $(y, t, t')$ in a set of measure $\gtrsim \varepsilon^{c}$ for a certain exponent $c > 0$. By Fubini’s theorem, there exists $\tilde{t}$ such that

$$F_j(y, t) = F_j(y, \tilde{t}) + O(\eta^{1-\sigma})$$

for all $(y, t)$ in a set of measure $\gtrsim \varepsilon^{c}$. That is, setting $F_j^*(y) = F_j(y, \tilde{t})$ and invoking the definition $F_j(y, t) = (\delta t)^{-1}(f_j(y + \delta t) - f_j(y))$,

$$f_j(y + \delta t) - f_j(y) = \delta t F_j^*(y) + O(\delta \eta^{1-\sigma})$$

for all $(y, t)$ in that same set. Now

$$\delta \eta^{1-\sigma} = \varepsilon^\gamma \varepsilon^{1-\gamma - \varrho}(1-\sigma) = \varepsilon^{1-\varrho - (1-\gamma)\sigma} = \varepsilon^{1-\varrho'}$$

where $\varrho' = \varrho + (1 - \gamma)\sigma$ can be made to be as small as desired, by choosing $\varrho, \sigma$ to be sufficiently small. Thus substituting $s = \delta t$,

$$f_j(y + s) = f_j(y) + s F_j^*(y) + O(\varepsilon^{1-\varrho'})$$

for all $(y, s)$ that satisfy $|s| \lesssim \delta$ and lie in a certain set of Lebesgue measure $\gtrsim \delta \cdot \varepsilon^{c}$.

As in the proof of Lemma 7.3, this means that for each $j$, upon partitioning the domain of each $f_j$ into intervals $I_{j,n}$ of lengths $\delta$, there exist affine functions $L_{j,n}$ so that if $f_j^*$ is defined to be $L_{j,n}$ on each $I_{j,n}$ then for all $x$ belonging to a subset $S^* \subset \overline{S}$ satisfying $|S^*| \gtrsim |S|C$,

$$||f_j \circ \varphi_j(x) - f_j^* \circ \varphi_j(x)|| \lesssim \varepsilon^{1-\varrho'} \quad \forall j \in \{1, 2\} \forall x \in S^*.$$  

To conclude the proof of Proposition 9.1 it remains to show that if $\varrho'$ is sufficiently small then the construction ensures that $(f_j^*)' = O(\varepsilon^{-\varrho})$. This is a matter of unraveling the above sequence of steps. For $y + s$ in each $\delta$-interval, $f_j^*(y + s)$ satisfies

$$f_j^*(y + s) = f_j^*(y) + s F_j^*(y, \tilde{t})$$

with $y = \varphi_j(x)$ with $(x, \tilde{t}) \in S_4$. Thus the derivative in question equals $F_j(\varphi_j(x), \tilde{t})$ for some $(x, t) \in S_4$. Now $(x, t) \in S_4$ if and only if $(x, \delta \tilde{t}) \in S_3$, whence $(x, \delta \tilde{t}) \in S_2$ and $|\tilde{t}| \geq \varepsilon^{\varrho}$. By the definition of $S_2$, this implies that $(x, e^{\delta W_3} x \in S \times \overline{S}$. This, in turn, implies that (9.7) holds, that is, $f_j \circ \varphi_j(x) - f_j \circ \varphi_j(x') = O(\delta \eta^{1-\sigma})$ and consequently

$$F_j(\varphi_j(x), \tilde{t}) = (\delta t)^{-1}((f_j \circ \varphi_j)(e^{\delta W_3} x) - (f_j \circ \varphi_j)(x))$$

$$= O((\delta \varepsilon^{\varrho})^{-1}\delta \eta^{1-\sigma})$$

$$= O(\varepsilon(1-\varrho')^{-1}).$$

By choosing $\sigma$ and $\varrho$, hence also $\varrho'$, to be sufficiently small, we ensure that this bound is $O(\varepsilon^{-\varrho})$. That concludes the proof of Proposition 9.1. \qed
10. Macroscale Analysis

The overarching strategy of the proof of Theorem 1.1 is to upgrade the smallness of $\sum_j a_j \cdot (f_j \circ \varphi_j)$ to smallness of its gradient. The gradient has two components, so a single scalar inequality is thereby transformed into two scalar inequalities, allowing the elimination of one of the three unknown functions $f_j$ and thus reducing matters to a potentially simpler problem. In this section we use the differentiability established through the mesoscale analysis to implement this strategy.

In the following proposition and its proof, $f'$ denotes the derivative, in the pointwise sense, of a piecewise differentiable function.

**Proposition 10.1 (Macroscale).** Let $(a, \Phi, f)$ satisfy all hypotheses of Theorem 1.1. Suppose that the curvature of the web in $B$ defined by $\Phi$ does not vanish identically. For any $\varrho > 0$ there exist $p, c, C \in (0, \infty)$ with the following property.

Let $\delta > 0$. Suppose that for each index $j$, the domain of $f_j$ is partitioned into intervals $I$ of common lengths $\delta$, and that the restriction of $f_j$ to each such interval $I$ is an affine function whose derivative satisfies $|f_j'| \leq \delta^{-\varrho}$. Then the sublevel set $S(f, \delta^{1+\varrho})$ satisfies $|S(f, \delta^{1+\varrho})| \leq C\delta^c$.

Only the case $\tau = 0$ of the auxiliary hypothesis (7.1) will be used in the proof of Proposition 10.1 but the main hypothesis and the lower bound on $|f_3|$ will come into play for the first time.

To simplify notation below we analyze the subset $S(f, \delta^{1+2\varrho})$ rather than $S(f, \delta^{1+2\varrho})$. Set $\eta = \delta^{1+2\varrho}$. We will begin the proof of Proposition 10.1 by establishing the following approximation result.

**Lemma 10.2.** Let $(a, \Phi) \in C^\omega(B)$. Suppose that $a_j$ vanish nowhere, and that the gradients of the mappings $\varphi_j$ are everywhere pairwise transverse. Suppose that the curvature of the web defined by $\Phi$ does not vanish identically, and that $(a, \Phi)$ satisfies the weak auxiliary hypothesis.

Let $k \in \{1, 2, 3\}$. There exist compact sets $T_k$ and $\Gamma_k$, a function $(y, t, \theta) \mapsto \alpha_{t, \theta}(y)$ that is $C^\omega$ in a neighborhood of $\varphi_k(B) \times T_k \times \Gamma_k$, and a function $(y, t) \mapsto \beta_t(y)$ that is $C^\omega$ in a neighborhood of $\varphi_k(B) \times T_k$ with the following property.

For any $\varrho > 0$ there exists $\rho_0(\varrho) > 0$ such that for any $\rho \in (0, \rho_0(\varrho))$ there exist $c, C \in (0, \infty)$ with the following property. Let $\delta > 0$. Let $f$ be Lebesgue measurable and satisfy $\|f_j\|_{L^\infty} = O(1)$ for each index $j$. Suppose that each $f_j$ is affine on intervals of lengths $\delta$ and satisfies $|f_j'| = O(\delta^{-\varrho})$. Let $S \subset S(f, \delta^{1+2\varrho})$ be measurable. Then either $|S| = O(\delta^c)$ or there exist a measurable subset $S' \subset S$ satisfying $|S'| \geq |S'|^C$, elements $t \in T_k$ and $\theta \in \Gamma_k$, and a scalar $\tilde{r} \in [c\delta^\varrho, 1]$ such that for every $y \in \varphi_k(S')$,

\begin{align}
(10.1) & \quad \tilde{r}\beta_t(y) f_k(y) = \alpha_{t, \theta}(y) + O(\delta^\varrho) \\
(10.2) & \quad |\beta_t(y)| \geq \delta^\varrho.
\end{align}

**Proof of Lemma 10.2.** It suffices to prove this for $k = 3$. The first step will be to establish smallness of $|\nabla(\sum_j a_j(f_j \circ \varphi_j))|$ on a significantly large subset of $S$. Consider any ordered triple $(I_j : j \in \{1, 2, 3\})$ of intervals $I_j$ of length $\delta$ for which $(\cap_{j=1}^3 \varphi_j^{-1}(I_j)) \cap S \neq \emptyset$. Choose any point $\bar{x} \in (\cap_{j=1}^3 \varphi_j^{-1}(I_j)) \cap S$. For any $x \in \cap_j \varphi_j^{-1}(I_j)$, the quantity $\sum_{j=1}^3 a_j(x) (f_j \circ
\( \varphi_j(x) \) can be expressed via Taylor expansion as

\[
\sum_{j=1}^{3} a_j(x) (f_j \circ \varphi_j)(x) = \sum_{j=1}^{3} a_j(\bar{x}) (f_j \circ \varphi_j)(\bar{x}) + (x - \bar{x}) \cdot \nabla (\sum_{j} a_j (f_j \circ \varphi_j))(\bar{x}) + O(\delta^{-\rho} |x - \bar{x}|^2)
\]

since \( f_j' = O(\delta^{-\rho}) \), \( f_j = O(1) \), the restriction of \( f_j \) to \( I_j \) has constant derivative, and \( a_j, \varphi_j = O(1) \) in \( C^2 \) norm. Therefore for any \( x \in (\bigcap_{j=1}^{3} \varphi_j^{-1}(I_j)) \cap S \),

\[
|(x - \bar{x}) \cdot \nabla (\sum_{j} a_j (f_j \circ \varphi_j)(\bar{x}))| = O(\eta + \delta^{-\rho}).
\]

Choose \( \rho \) sufficiently small that \( 2 - \rho \geq 1 + 2\eta \). Then \( \eta + \delta^{-\rho} = O(\delta^{1+2\rho} + \delta^{-\rho}) = O(\eta) \).

For any vector \( v \in \mathbb{R}^2 \), \( \{ x \in B(\bar{x}, C\delta) : |(x - \bar{x}) \cdot v| = O(\eta) \} \) has Lebesgue measure \( O(\min(\delta^2, \eta|v|^{-1})\delta) \). The set \( \bigcap_{j=1}^{3} \varphi_j^{-1}(I_j) \) is contained in \( B(\bar{x}, C\delta) \). Therefore by choosing \( v = \nabla (\sum_{j} a_j (f_j \circ \varphi_j)(\bar{x})) \), it follows that either

\[
\nabla (\sum_{j} a_j (f_j \circ \varphi_j)(\bar{x})) | \leq \delta^6
\]

or

\[
|\bigcap_{j} \varphi_j^{-1}(I_j) \cap S| = O(\eta|v|^{-1})\delta = O(\delta^{1+2\rho})\delta = O(\delta^{2+\rho}).
\]

If the former inequality holds then

\[
|\nabla (\sum_{j} a_j (f_j \circ \varphi_j)(x))| = O(\delta + \delta^{1-\rho} + \delta^6) = O(\delta^6)
\]

as well, since \( |x - \bar{x}| = O(\delta) \), \( \|a_j\|_{C^2} = O(1) \), \( f_j = O(1) \), \( f_j' \) is constant on \( I_j \), and we may assume \( \rho, \eta \) to be sufficiently small that \( 1 - \rho \geq \eta \).

There are \( O(\delta^{-2}) \) ordered triples \( (I_1, I_2, I_3) \) of intervals with \( \bigcap_{j} \varphi_j^{-1}(I_j) \neq \emptyset \). Summing over all of these, we conclude from the preceding dichotomy that

\[
|S| \leq \delta^6 + |\{ x \in S(f, \eta) : |\nabla (\sum_{j} a_j (f_j \circ \varphi_j)(x))| = O(\delta^6) \}|
\]

It remains to establish an upper bound for the measure of the set

\[
S = \{ x \in S(f, \eta) : |\nabla (\sum_{j=1}^{3} a_j (f_j \circ \varphi_j)(x))| = O(\delta^6) \}.
\]

We may assume that \( a_3 \equiv 1 \). Then

\[
S \subset \{ x \in S(f, \eta) : |\sum_{j=1}^{2} W_3(a_j (f_j \circ \varphi_j))(x)| = O(\delta^6) \}.
\]

By sacrificing one scalar inequality, we have eliminated the contribution of \( f_3 \) from the sum. Although the number of indices \( j \) in play has been reduced from three to two, there are now effectively two arbitrary functions for each index. Indeed, write

\[
W_3(a_j (f_j \circ \varphi_j)) = a_j \cdot W_3 \varphi_j \cdot (g_j \circ \varphi_j) + W_3 a_j \cdot (f_j \circ \varphi_j)
\]
with $g_j = f_j'$. The relationship between $g_j$ and $f_j$ provides limited information in this context, for the sublevel set, and hence its images under the mappings $\varphi_j$, could consist of many small connected components, so that information about $f_j'$ cannot be integrated to obtain useful macroscale information about $f_j$ itself. Thus for arbitrary measurable functions satisfying $f_j = O(1)$ and $g_j = O(\delta^{-\rho})$, we seek to analyze the Lebesgue measure of

$$ (10.12) \quad S_2 = \left\{ x \in B : \left| \sum_{j=1}^{2} \left[ a_j(x) \cdot W_3 \varphi_j(x) \cdot (g_j \circ \varphi_j)(x) + W_3 a_j(x) \cdot (f_j \circ \varphi_j)(x) \right] \right| = O(\delta^\rho) \right\} $$

without positing any relationship between $f_j$ and $g_j$. This increase in the number of unknown functions, from three to four, is the final difficulty to be overcome. A related issue arose in the mesoscale analysis, but we need to proceed differently here.

The reduction in the number of mappings $\varphi_j$ in the definition of the sublevel set, from three to two, is a crucial simplification. To begin to exploit it, change variables so that $\varphi_1, \varphi_2$ take the form $\varphi_j(x_1, x_2) \equiv x_j$. Let $S_3$ be the set of all $(t, y) = ((t, t'), y) \in \mathbb{R}^2 \times \mathbb{R}^1$ such that $((t, y), (t', y)) \in S_2 \times S_2$. Its three-dimensional Lebesgue measure satisfies $|S_3| \gtrsim |S_2|^2$, by the Cauchy-Schwarz inequality.

Let $\bar{c} > 0$ be a small constant and define $S_3^\ast$ to be the set of all $t$ such that

$$ |\{ y : (t, y) \in S_2 \}| \geq \bar{c}|S_2|. $$

If $\bar{c}$ is chosen to be sufficiently small then necessarily $|S_3^\ast| \gtrsim |S_3|$.

We now follow a path taken in §20 of [1]. For each $(t, y) \in S_3$, (10.12) can be regarded as a matrix equation

$$ (10.13) \quad B(t, y) \begin{pmatrix} g_2(y) \\ f_2(y) \end{pmatrix} = A(t, y) + O(\delta^\rho) $$

for an unknown quantity $\begin{pmatrix} g_2(y) \\ f_2(y) \end{pmatrix}$, with right-hand side $A(t, y)$ and coefficient matrix $B(t, y)$ defined by

$$ (10.14) \quad A(t, y) = - \begin{pmatrix} a_1(t, y)W_3 \varphi_1(t, y)g_1(t) + W_3 a_1(t, y)f_1(t) \\ a_1(t', y)W_3 \varphi_1(t', y)g_1(t') + W_3 a_1(t', y)f_1(t') \end{pmatrix} $$

and

$$ (10.15) \quad B(t, y) = \begin{pmatrix} a_2(t, y)W_3 \varphi_2(t, y) & W_3 a_2(t, y) \\ a_2(t', y)W_3 \varphi_2(t', y) & W_3 a_2(t', y) \end{pmatrix}. $$

We introduce a parameter $\theta = (\theta_1, \ldots, \theta_4) \in \mathbb{R}^4$ that satisfies $|\theta| = O(\delta^{-\rho})$ and define

$$ (10.16) \quad A_\ast(t, \theta, y) = - \begin{pmatrix} a_1(t, y)W_3 \varphi_1(t, y)\theta_1 + W_3 a_1(t, y)\theta_2 \\ a_1(t', y)W_3 \varphi_1(t', y)\theta_3 + W_3 a_1(t', y)\theta_4 \end{pmatrix}. $
The relation \((10.13)\) is thus rewritten as an instance of the more general relation

\[(10.17)\]

\[B(t, y) \begin{pmatrix} g_2(y) \\ f_2(y) \end{pmatrix} = A_\ast(t, \theta, y) + O(\delta^\rho).\]

Define

\[(10.18)\]

\[\beta_\ast(y) = \det(B(t, y)),\]

which is a real analytic function of \((t, y)\). According to Lemma \ref{lemma5.2}, the \(C^\omega\) function \((t, y) \mapsto \beta_\ast(y)\) does not vanish identically on any nonempty open set. This is a consequence of the weak auxiliary hypothesis; it does not rely on the full strength of the main hypothesis.

Since \((t, y) \mapsto \beta_\ast(y)\) is a real analytic function that does not vanish identically,

\[(10.19)\]

\[|\{(t, y) : |\beta_\ast(y)| < \delta^\rho\}| \leq C'\delta^A\rho\]

for some \(A < \infty\) and \(C' < \infty\). If \(|S_3| \leq 2C'\delta^A\rho\) then \(|S_2| = O(|S_3|^{1/2}) = O(\delta^\rho)\) with \(c = A\rho/2\). Therefore \(|S| = O(\delta^\rho)\), which is one of the alternative conclusions of Lemma \ref{lemma10.2}.

Therefore we may assume that \(|S_3| \geq 2C'\delta^A\rho\).

Multiplying by the cofactor matrix associated to \(B\) transforms the equation \((10.13)\) to

\[(10.20)\]

\[\beta_\ast(y) \begin{pmatrix} g_2(y) \\ f_2(y) \end{pmatrix} = A(t, \theta(t), y) + O(\delta^\rho)\]

with \(A(t, \theta, y)\) a \(C^\omega\) \(\mathbb{R}^2\)-valued function of \((t, \theta, y)\) and with \(\theta(t) = (f_1(t), g_1(t), f_1(t'), g_1(t'))\). \(A\) satisfies \(A(t, \theta, y) = O(\delta^{-\rho})\) uniformly in \((t, y)\), since \(f_1 = O(1)\) and \(g_1 = O(\delta^{-\rho})\). It depends linearly on \(\theta\). Therefore we may restrict \(\theta\) to a fixed bounded subset of \(\mathbb{R}^4\) and write \(A(t, R\theta, y)\) where \(0 \leq R \leq O(\delta^{-\rho})\).

Restricting attention to the component of the vector equation \((10.20)\) involving \(f_2\) gives

\[(10.21)\]

\[\beta_\ast(y) f_2(y) = R\alpha_{t, \theta}(y) + O(\delta^\rho) \quad \forall \,(t, y) \in S_3,\]

where

\[(10.22)\]

\[\left\{ \begin{array}{l}
R = 1 + |(f_1(t), g_1(t), f_1(t'), g_1(t'))| \in [1, O(\delta^{-\rho})], \\
\theta = R^{-1}(f_1(t), g_1(t), f_1(t'), g_1(t')),
\end{array} \right.\]

and \(\alpha_{t, \theta}(y)\) is a real analytic function of \((t, \theta, y)\).

There exists \(t\) such that \(S_3^+ = \{ y \in \mathbb{R} : (t, y) \in S_3 \}\) satisfies \(|S_3^+| \geq |S_3| \geq \delta^{A_{1}}\), where \(\cdot\) denotes one-dimensional Lebesgue measure on the left-hand side of this inequality, and three-dimensional measure on the right. There exists \(\theta\), which depends on \(t\) but not on \(y\), such that

\[(10.23)\]

\[\beta_\ast(y) f_2(y) = R\alpha_{t, \theta}(y) + O(\delta^\rho) \quad \forall \, y \in S_3^+\]

It will be more convenient in the analysis below to rewrite \((10.23)\) in the equivalent form

\[(10.24)\]

\[\tilde{r}\beta_\ast(y) f_2(y) = \alpha_{t, \theta}(y) + O(\delta^\rho) \quad \forall \, y \in S_3^+\]

with \(c\delta^\rho \leq \tilde{r} \leq 1\).

This is a form of the second alternative conclusion of Lemma \ref{lemma10.2} but the lemma asserts the stronger conclusion that there exists a subset \(S \subset S\) of measure \(\geq \delta^c\) such that \((10.24)\) holds for every \(y \in \varphi_2(S)\). The stronger conclusion is obtained from a simple modification of this proof. Indeed, let \(c_0 > 0\) be small and define \(y \in \varphi_2(B)\) to be rich if the set of all \(x \in S\) satisfying \(\varphi_2(x) = y\) has measure \(\geq c_0|S|\). If \(c_0\) is chosen sufficiently small then the set \(S' \subset S\) of all \(x \in S\) such that \(\varphi_2(x)\) is rich has measure \(\geq |S|\). Apply the above reasoning to \(S'\), rather than to \(S\). Define \(S \subset S'\) to be the set of all \(x \in S'\) such that

\[\cdot\]
\( \varphi_2(x) \in \mathcal{S}_3^+ \). Then \(|S| \gtrsim |\mathcal{S}_3^+| \cdot |\mathcal{S}| \gtrsim |\mathcal{S}|^c\) for some \( c > 0 \). Thus the second conclusion of Lemma \([10.2]\) has been established, completing the proof of the lemma. \( \square \)

We next use Lemma \([10.2]\) to complete the proof of Proposition \([10.1]\). By applying Lemma \([10.2]\) to the three indices \( k = 1, 2, 3 \) in succession, we conclude that either \(|S(f, \delta^{1+2\theta})| \leq \delta^{c_6}\), or there exist a measurable subset \( \mathcal{S} \subset S(f, \delta^{1+2\theta}) \) satisfying \(|\mathcal{S}| \gtrsim |S(f, \delta^{1+2\theta})|^C\) and three three-tuples \((t_k, \theta_k, \bar{r}_k)\) of parameters such that for each index \( j \in \{1, 2, 3\}\),

\[
\begin{align*}
(10.25) & \quad \bar{r}_j \beta^{\ast}_{t_j}(y) f_j(y) = \alpha^{\ast}_{t_j, \theta_j}(y) + O(\delta^\theta) \ \forall \ y \in \varphi_j(\mathcal{S}) \\
(10.26) & \quad |\beta^{\ast}_{t_j}(y)| \geq \delta^\rho \ \forall \ y \in \varphi_j(\mathcal{S}) \\
(10.27) & \quad c \delta^\rho \leq \bar{r}_j \leq 1.
\end{align*}
\]

We may assume that \(|S(f, \delta^{1+2\theta})| > \delta^{c_6}\) for a small constant \( c_6\), for otherwise the conclusion asserted by Proposition \([10.1]\) holds. Consequently \(|\mathcal{S}|\) satisfies a lower bound of the same type, with an exponent proportional to \( c_6\). Choose \((t_k, \theta_k, \bar{r}_k)\) satisfying the conclusions \([10.25]\), \([10.26]\), \([10.27]\). Write \( \mathbf{t} = (t_1, t_2, t_3), \bar{\theta} = (\theta_1, \theta_2, \theta_3), \) and \( \bar{r} = (\bar{r}_1, \bar{r}_2, \bar{r}_3)\). Then

\[
(10.28) \quad \sum_{k=1}^{3} \bar{r}_k \beta^{k}_{t_k}(\varphi_k(x)) \sum_{j=1}^{3} a_j(x) (f_j \circ \varphi_j)(x)
\]

\[
= \sum_{j=1}^{3} a_j(x) \prod_{k \neq j} \bar{r}_k \beta^{k}_{t_k}(\varphi_k(x)) \alpha^{\ast}_{t_j, \theta_j}(\varphi_j(x)) + O(\delta^\theta)
\]

for all \( x \in \mathcal{S}\).

Introduce the function

\[
(10.29) \quad F(x, \mathbf{t}, \theta, r) = F_{\mathbf{t}, \theta, r}(x) = \sum_{j=1}^{3} a_j(x) \prod_{k \neq j} r_k \beta^{k}_{t_k}(\varphi_k(x)) \alpha^{\ast}_{t_j, \theta_j}(\varphi_j(x)),
\]

where \( \mathbf{t} = (t_1, t_2, t_3), r = (r_1, r_2, r_3), \) and \( \bar{\theta} = (\theta_1, \theta_2, \theta_3)\). The parameter \((x, \mathbf{t}, \theta, r)\) varies freely over a certain compact domain, in a neighborhood of which \( F \) is an analytic function of \((x, \mathbf{t}, \theta, r)\). For the particular parameter \((\mathbf{t}, \bar{\theta}, \bar{r})\) that appears in \((10.28)\),

\[
(10.30) \quad |F(x, \mathbf{t}, \bar{\theta}, \bar{r})| = O(\delta^\theta) \ \forall \ x \in \mathcal{S}.
\]

We claim that \( F \) does not vanish identically on any nonempty open set in the space of the vector \((x, \mathbf{t}, \theta, r)\). It is through this claim that the main hypothesis of Theorem \([1.1]\) first comes into play. If \( F \equiv 0 \) on an open set \( U \) of the parameter space, then since none of the factors \( \beta^{k}_{t_k} \) vanish identically, \( \sum_{j=1}^{3} a_j(x) \alpha^{\ast}_{t_j, \theta_j}(\varphi_j(x)) \equiv 0 \) would vanish identically in \( U \). By the main hypothesis, \( \alpha^{\ast}_{t_j, \theta_j} \circ \varphi_j \) would vanish identically in a certain nonempty open subset of \( B \), for each \( j \in \{1, 2, 3\} \), for all \(( (t_k, \theta_k) : k \in \{1, 2, 3\} )\).

This is a contradiction. Indeed, \((10.16)\) expresses \( \alpha_{t_2, \theta_2}^2(y) \) as the second component of the vector \( B(t_2, y)^{-1}A_\ast(t_2, \theta_2, y) \), where \( \theta_2 = (\theta_{2,1}, \ldots, \theta_{2,4}) \in \mathbb{R}^4 \). As \( \theta_2 \) varies over \( \mathbb{R}^4 \), the vector \( A_\ast(t_2, \theta_2, y) \) varies over all of \( \mathbb{R}^2 \) for generic \((t_2, \theta_2)\) since \( a_2(t, y) \) and \( W_{2,2} \varphi_2(t, y) \) vanish nowhere. Moreover, for generic \((t_2, y)\) \( B(t_2, y) \) is invertible and hence does not map \( \mathbb{R}^2 \) into the subspace of \( \mathbb{R}^2 \) defined by the vanishing of the second component. Thus it is not the case that \( \alpha_{t_2, \theta_2}^2 \circ \varphi_2 \) vanishes identically in a certain nonempty open set for all \(( (t_k, \theta_k) : k \in \{1, 2, 3\} )\).
Choose and fix \( x_0 \in B \). Define
\[
G(t, \theta, r) = \sum_{|\alpha| \leq N} |\partial_x^\alpha F(x_0, t, \theta, r)|^2,
\]
with \( N \) chosen so that there exists \( C \) such that
\[
C^{-1} \sum_{0 \leq |\alpha| \leq N} |\partial_x^\alpha F(x, t, \theta, r)|^2 \leq G(t, \theta, r) \leq C \sum_{0 \leq |\alpha| \leq N} |\partial_x^\alpha F(x, t, \theta, r)|^2
\]
uniformly for all \( (x, t, \theta, r) \). Lemma 10.3 guarantees that such an \( N \) exists.

Let
\[
\Sigma = \{(t, \theta, r) : G(t, \theta, r) = 0\}.
\]
In this definition, \( r \) is allowed to vary freely over \([0, 1]^3\), while \( t, \theta \) vary over closed balls in Euclidean spaces.

\( \Sigma \) is a \( C^\omega \) variety of positive codimension. By Lojasiewicz’s theorem, there exist \( c, \kappa > 0 \) such that
\[
G(t, \theta, r) \geq c \text{ distance}((t, \theta, r), \Sigma)^\kappa \quad \forall (t, \theta, r).
\]

**Lemma 10.3.** If \( 0 < \rho \leq \rho/4 \) then there exists a constant \( c > 0 \) such that for all data satisfying the hypotheses of Proposition 10.1, the parameter \( (t, \theta, \bar{r}) \) that appears in (10.28) satisfies
\[
\text{distance}((\bar{t}, \bar{\theta}, \bar{r}), \Sigma) \geq c \delta^{2\rho}.
\]

**Proof.** We claim that for any \((t, \theta, r) \in \Sigma\), at least one of the following three vanishing properties holds.

1. \( \alpha_t^j \circ \varphi_j \) vanishes identically on \( B \) for every \( j \in \{1, 2, 3\} \).
2. There exists \( k \in \{1, 2, 3\} \) such that \( \beta_t^j \circ \varphi_k \) vanishes identically on \( B \).
3. There exists \( k \in \{1, 2, 3\} \) such that \( r_k = 0 \).

Indeed, for \((t, \theta, r) \in \Sigma\) the function \( \sum_{|\alpha| \leq N} |\partial_x^\alpha F(x, t, \theta, r)|^2 \) vanishes identically for \( x \in B \). In particular, \( F(x, t, \theta, r) = 0 \) for every \( x \in B \). If none of the three factors \( \beta_t^j \circ \varphi_k \) vanish identically on \( B \), and if no \( r_k \) vanishes, then there exists a nonempty open subset \( V \subset B \) on which none of these three vanish. Multiply by \( \prod_{k=1}^3 (\beta_t^j \circ \varphi_j)^{-1} \) in (10.29) to conclude that
\[
\sum_{j=1}^3 a_j(x) \cdot [(\beta_t^j)^{-1} \alpha_t^j] \circ \varphi_j(x) = 0 \quad \forall x \in V.
\]
The main hypothesis of Theorem 1.1 guarantees that for each \( j \in \{1, 2, 3\} \), \((\beta_t^j)^{-1} \alpha_t^j \circ \varphi_j \) vanishes identically on \( \varphi_j(V) \). Therefore \( \alpha_t^j \circ \varphi_j \) vanishes identically in \( \varphi_j(V) \), hence also in \( \varphi_j(B) \) by analytic continuation. This establishes the claim.

Now consider the parameter \((t, \bar{\theta}, \bar{r}) \) in (10.28). According to (10.26), \( |\beta_t^j(\varphi_j(y))| \geq \delta^\rho \) for every \( y \in S \), for each index \( j \). Therefore \((t, \bar{\theta}, \bar{r}) \) lies at distance \( \geq \delta^\rho \) from the subset of \( \Sigma \) on which any \( \beta_t^j \circ \varphi_j \) vanishes identically.

Likewise, the parameter \( \bar{r} \) in (10.28) belongs to \([c \delta^\rho, 1]^3\) and therefore \((\bar{t}, \theta, \bar{r}) \) lies at distance at least \( c \delta^\rho \) from the subset of \( \Sigma \) on which \( r = 0 \).
For any \( y \in \varphi_3(S) \), the relation \( \tilde{r}_3 \beta_3^2(y) f_3(y) = \alpha_3^3 \delta_3(y) + O(\delta^\rho) \), together with the hypothesis that \( |f_3(y)| \geq 1 \) — used here for the first time — and the lower bounds established in the preceding paragraphs, imply that

\[
|\alpha_3^3 \delta_3(y)| \geq \delta^\rho \cdot \delta^\rho \cdot 1 - O(\delta^\rho).
\]

Since \( \rho \) was chosen to satisfy \( 4 \rho \leq \rho \), \( |\alpha_3^3 \delta_3(y)| \geq \delta^2 \rho \) for every \( y \in \varphi_3(S) \) provided that \( \delta = \varepsilon^\gamma \) is sufficiently small; this holds if \( \varepsilon_0 \) is chosen to be sufficiently small. Therefore the parameter \( (t, \theta, \bar{r}) \) in (10.28) lies at distance \( \geq \delta^2 \rho \) from the subset of \( \Sigma \) on which \( \alpha_3^3 \delta_3 \) vanishes identically.

By Lemma 10.3 and (10.34), \( G(t, \bar{r}, \theta, \bar{r}) \geq \delta^2 \rho \). Therefore by (10.32),

\[
\sum_{|a| \leq N} \left| \partial^a \bar{F}(x, t, \bar{r}, \bar{r}) \right|^2 \geq \delta^C \rho \quad \forall x \in \mathcal{S}.
\]

Moreover, the \( C^{N+1} \) norm of \( \bar{F} \), as a function of \( x \), is bounded above by a finite constant, uniformly in \( (t, \theta, \bar{r}) \). By Lemma 1.2 the measure of the sublevel set on which \( |\bar{F}| = O(\delta^\rho) \) is \( O(\delta^\rho - C \rho) \) for certain constants \( c, C \in (0, \infty) \). Choosing \( \rho \) to be sufficiently small, as a function of \( \varrho \), ensures that \( cg - C \rho > 0 \), completing the proof of Proposition 10.1.

11. Conclusion of Proof for the Nondegenerate Case

We are now in a position to complete the proof of Theorem 1.1 in the main case.

**Proof of Theorem 1.1 in the nondegenerate case.** Let \((a, \Phi)\) satisfy the hypotheses of the nondegenerate case of the theorem. As was explained in the introduction, we may also assume that the curvature of the web in \( B \) defined by \( \Phi \) does not vanish identically. Choose small positive quantities \( \varrho, \rho, \varepsilon_0 \), with \( \rho \) depending on the choice of \( \varrho \) and \( \varepsilon_0 \) depending on both. Let \( \varepsilon \in (0, \varepsilon_0] \). Choose and fix any \( \gamma \in (\frac{1}{2}, 1) \).

Let \( f \) be measurable and satisfy \( \|f\|_{L^\infty} = O(1) \) and \( |f_3(y)| \geq 1 \) for every \( y \in \varphi_3(B) \). By Proposition 9.1 either \( |S(f, \varepsilon)| \leq \varepsilon \), in which case the proof is complete, or there exist \( \mathcal{S} \subset \mathcal{S}(f, \varepsilon) \) satisfying \( |\mathcal{S}| \geq |S(f, \varepsilon)|^C \) and functions \( g_j \) that are affine on intervals of length \( \varepsilon^\gamma \) and satisfy \( |g_j(\varphi_j(x)) - f_j(\varphi_j(x))| \leq \varepsilon^{1-\rho} \) for every \( x \in \mathcal{S} \) for each index \( j \), and also \( |g_3(\varphi_3(x))| \geq 1 \) for every \( x \in \mathcal{S} \). Moreover, according to (9.4) these functions can be constructed so that their derivatives satisfy \( |g_j'(y)| \leq C \varepsilon^{-\rho} \) for every \( y \). The parameters \( c, C \) also depend on \( \rho \) but not on \( \varepsilon, f \). In the case in which the proof is not already complete, it suffices to show that \( |S(g, \varepsilon^{1-\rho})| \leq \varepsilon^c \).

Choose \( \rho \) to satisfy \( (1 + \rho) \gamma < 1 \). The piecewise affine function \( g \) satisfies the hypotheses of Proposition 10.1, whose conclusion is that if \( \rho \) is chosen to be sufficiently small as a function of \( \varrho \) in the preceding step then \( |S(g, (\varepsilon \rho)^{1+\rho})| \leq C \varepsilon^c \). Moreover, if \( \rho, \varrho \) are both sufficiently small then \( (1 + \rho) \gamma < (1 - \rho) \). Since \( \varepsilon \leq \varepsilon_0, \varepsilon (1 + \rho) \gamma > \varepsilon^{1-\rho} \) and therefore \( S(g, \varepsilon^{1-\rho}) \subset S(g, (\varepsilon \rho)^{1+\rho}) \). Thus \( |S(g, \varepsilon^{1-\rho})| \leq C \varepsilon^c \), and the proof of Theorem 1.1 is complete in the nondegenerate case.

12. The Degenerate Case

We next turn to the degenerate case of Theorem 1.1 showing how the degenerate case is a consequence of the nondegenerate case via a rather general reduction procedure. Define \( \Sigma \subset \tilde{B} \) to be the degenerate locus. \( \Sigma \) is the union of the real analytic varieties defined by the vanishing of some \( a_j(x) \), and the linear dependence of gradients \( \nabla \varphi_i(x), \nabla \varphi_j(x) \) for some pair of indices \( i \neq j \).
To avoid steps in which we divided by some $a_k$ or implicitly divided by some component of some gradient $\nabla \varphi_j$, we modify various definitions in the analysis, as follows. Define the vector fields
\begin{equation}
W_k = \frac{\partial \varphi_k}{\partial x_2} \cdot \frac{\partial}{\partial x_1} - \frac{\partial \varphi_k}{\partial x_1} \cdot \frac{\partial}{\partial x_2}
\end{equation}
so that $W_k(\varphi_k) \equiv 0$, $W_k \in C^\omega$, and $W_k$ does not vanish identically. The mappings $\psi_j^\varepsilon$ are now defined in terms of this modified $W_3$. One can likewise construct real analytic vector fields $V_j^\varepsilon$, for $j = 1, 2$, that satisfy $V_j(\psi_j^\varepsilon) \equiv 0$ but do not vanish identically.

The definitions of the quantities $b_n^\varepsilon$ should likewise be modified to eliminate denominators. For odd $n$, define $b_n^\varepsilon$ by changing its definition \((6.14)\) to
\begin{equation}
b_n^\varepsilon(t, z) = -a_1(\Theta_{n,z}^\varepsilon(t_1, \ldots, t_n)) \cdot a_2(\Theta_{n-1,z}^\varepsilon(t_1, \ldots, t_{n-1})) \\
\quad \cdot a_1(\Theta_{n-2,z}^\varepsilon(t_1, \ldots, t_{n-2})) \cdots a_1(\Theta_{1,z}^\varepsilon(t_1)),
\end{equation}
the product of the numerators only in \((6.14)\), incorporating also the minus sign; and define $\tilde{b}_n^\varepsilon$ to be the product of the denominators in \((6.14)\). For even $n$ make the corresponding definitions, based on the numerators and denominators in \((6.16)\).

The relation \((6.19)\) was $(g_j \circ \psi_j^\varepsilon)(\Theta_{n,z}^\varepsilon(t)) = g_1(z)b_n^\varepsilon(t, z) + O(\varepsilon)$; with the revised definition of $b_n^\varepsilon$ it becomes
\begin{equation}
\tilde{b}_n^\varepsilon(t, z)(g_j \circ \psi_j^\varepsilon)(\Theta_{n,z}^\varepsilon(t)) = g_1(z)b_n^\varepsilon(t, z) + O(\varepsilon),
\end{equation}
with $n = 3$ for $j = 2$ and $n = 4$ for $j = 1$. Likewise, modify the definition \((7.25)\) of $F_z(t, s, s')$ by multiplying by a real analytic function of $(z, t, s, s')$ that does not vanish identically on any open set, and that suffices to cancel all denominators implicitly present. With this modification, there exist $C, \tau \in (0, \infty)$ satisfying
\[
\left\{ (z, t, s, s') : |F_z(t, s, s')| < \delta \right\} \leq C \delta^\tau \forall \delta > 0.
\]

Let $c_1, c_2 > 0$ be small parameters. The set of all $x \in B$ lying within distance $\varepsilon^{c_1}$ of the positive codimension real analytic variety $\Sigma$ has Lebesgue measure $O(\varepsilon^{c_3})$ for some $c_3 = c_3(c_1) > 0$. Thus $B$ can be expressed as the union of a set of measure $O(\varepsilon^{c_3})$ and of $O(\varepsilon^{-c_2})$ balls $B_\nu$, with each ball having radius $\varepsilon^{c_2}$ and lying at distance $\geq \varepsilon^{c_1}$ from $\Sigma$. By Lojasiewicz’s theorem, each coefficient $a_j$ satisfies $|a_j| \geq c_4^\varepsilon$ at every point in each $B_\nu$, where $c_4 = Cc_1$. Lower bounds of this type also hold for $|\nabla \varphi_1|$ and $|\det(\nabla \varphi_1, \nabla \varphi_2)|$ on each $B_\nu$ for any pair of distinct indices $i, j$.

Choosing $c_2$ to be $C_5c_1$ for a sufficiently large constant $C_5$ and using these bounds, the proof of Proposition \((7.2)\) can be repeated on each ball $B_\nu$ with a loss of a factor that is $O(\varepsilon^{-C_6c_1-Cc_2})$ for some constant $C < \infty$. A typical point is that an upper bound $O(1)$ for the $C^2$ norm of each $\varphi_k$ together with a lower bound $c_5\varepsilon^{c_3}$ for $\det(\nabla \varphi_1, \nabla \varphi_j)$ for distinct indices $i, j$ allow the conclusion that the restriction of each mapping $x \mapsto (\varphi(x), \varphi_j(x))$ to $B_\nu$ is a diffeomorphism, for all sufficiently small $\varepsilon > 0$, if $c_2 < 2c_5$. This legitimizes passing from a lower bound for the Lebesgue measure of a set $S \subset B$ to a lower bound of the form $c|S|^2$ for the Lebesgue measure of $\tilde{S} = \{(x, x') \in S \times S : \varphi_j(x) = \varphi_j(x')\}$, as was done repeatedly in the analysis of the nondegenerate case.

Summing over all balls $B_\nu$ and taking the exceptional set of points near $\Sigma$ into account, the conclusion is that if $c_1, c_2$ are chosen sufficiently small relative to the exponent $\tau$ in the upper bound for $|F_z(t, s, s')|$, then either
\begin{equation}
|S(f, \varepsilon)| \leq C\varepsilon^{c_6-c_1-Cc_2} + C\varepsilon^{c_3},
\end{equation}
or there exist a measurable set $S$ contained in $S(f,\varepsilon) \cap B_\nu$ for some index $\nu$ and functions $f_j^\nu$ that satisfy the conclusions (7.5) and (7.6) of Proposition 7.2. Choose $c_1, c_2$ to also be sufficiently small that the net exponent $c - Cc_1 - Cc_2$ in (12.4) is positive.

The same reasoning applies to the mesoscale and macroscale analyses, completing the proof of Theorem 1.1 in the general case.

Part 3. Complements and further results

13. Reformulation of the main hypothesis

We establish an equivalent reformulation of the main hypothesis. This provides a straightforward extension of that hypothesis to $C^\infty$ data $(a, \Phi)$. To any $f = (f_1, f_2, f_3)$ associate the function

$$G_f(x) = \sum_{j=1}^3 a_j(x)(f_j \circ \varphi_j)(x).$$

Proposition 13.1. Let $(a, \Phi)$ be real analytic in a neighborhood $\tilde{B}$ of a closed ball $B \subset \mathbb{R}^2$. Assume that none of the mappings $\varphi_j$ are constant on $B$, and that there exists no pair of distinct indices $i \neq j \in \{1, 2, 3\}$ for which $\nabla \varphi_i$ and $\nabla \varphi_j$ are everywhere linearly dependent. Then the following are equivalent.

1. $(a, \Phi)$ satisfies the main hypothesis.
2. For each point $\bar{x} \in B$ at which $(a, \Phi)$ is nondegenerate, there exists a positive integer $M$ such that for any $f \in C^M$ defined in a neighborhood of $\Phi(\bar{x})$, if $G_f$ vanishes to order $M$ at $\bar{x}$ then every $f_j$ vanishes to order $M$ at $\varphi_j(\bar{x})$.
3. There exist $M \in \mathbb{N}$, $C < \infty$, and $\tau > 0$ and a real analytic variety $\Sigma \subset \tilde{B}$ of positive codimension such that for every $\bar{x} \in B$ and every tuple of functions $f \in C^M$ defined in a neighborhood of $\Phi(\bar{x})$,

$$\text{distance}(\bar{x}, \Sigma)^\tau \sum_{j=1}^3 \sum_{0 \leq k \leq M} \left| \frac{d^k}{dy^k} (f_j \circ \varphi_j)(\bar{x}) \right| \leq C \sum_{0 \leq |\alpha| \leq M} \left| \frac{\partial^\alpha}{\partial x^\alpha} G_f(\bar{x}) \right|.$$

The proof developed below has subsequently been applied in Proposition 7.1 of [4] to prove a corresponding result for inequalities involving two indices $j$, rather than three.

Condition (3) of the Proposition directly implies (2), which directly implies (1). Conversely, (1) implies (3) by the next lemma and Lojasiewicz’s theorem. We say that a function vanishes to order $N$ at a point if the function, and all of its partial derivatives of orders less than or equal to $N$ vanish at that point.

Lemma 13.2. Let $(a, \Phi)$ satisfy the main hypothesis and be nondegenerate. There exists a positive integer $M$ such that for any $\bar{x} \in B$ and any $f \in C^M$ in a neighborhood of $\Phi(\bar{x})$, if $G_f$ vanishes to order $M$ at $\bar{x}$ then every $f_j$ vanishes to order $M$ at $\varphi_j(\bar{x})$.

Conversely, if $(a, \Phi)$ satisfies the conclusion of Lemma 13.2, then the main hypothesis of Theorem 1.1 is satisfied.

Proof of Lemma 13.2. Let $U \subset B$ be open and nonempty. Let $f = (f_1, f_2, f_3)$, with each $f_j \in C^2(\varphi_j(U))$. Suppose that $G_f$ vanishes identically on some open set. With $b_j = a_j/a_3$,

$$0 = W_3(a_3^{-1}G_f) = \sum_{j=1}^2 \left[ W_3(b_j)(f_j \circ \varphi_j) + b_j W_3(\varphi_j)(f'_j \circ \varphi_j) \right].$$
By dividing by \( b_2 W_3(\varphi_2) \) and then applying the vector field \( W_2 \), we obtain a linear relation, with \( C^\infty \) coefficients, between \( f_j \circ \varphi_j \) and \( f_j' \circ \varphi_j \) for \( j = 1, 2 \) and \( f_j'' \circ \varphi_1 \). The coefficient of \( f_j'' \circ \varphi_1 \) vanishes nowhere, since \( W_2(\varphi_1) \) vanishes nowhere. Thus we may express \( f_j'' \circ \varphi_1 \) as a \( C^\infty \)-linear combination of those \( f_j \circ \varphi_j \) and \( f_j' \circ \varphi_j \) with \( j \in \{1, 2\} \).

Since the roles of the indices can be freely permuted, there is an expression of this type for \((f_j'' \circ \varphi_j : j \in \{1, 2, 3\})\) as a \( C^\infty \) linear combination of \(((f_j \circ \varphi_j, f_j' \circ \varphi_j) : j \in \{1, 2, 3\})\). If each \( f_j \in C^\infty \), this relation can be differentiated arbitrarily many times. We conclude that if \( G_f \) vanishes identically on some open set \( U \), then for each \( x \in U \) and for every \( n \geq 2 \), the \( n \)-th derivative \( f_j^{(n)}(\varphi_j(x)) \) is determined as a linear combination of \(((f_k(\varphi_k(x)), f_k'(\varphi_k(x)) : k \in \{1, 2, 3\})\). The coefficients in these linear combinations are real analytic functions of \( x \).

Likewise, if \( G_f \) vanishes to order \( N \) at some point \( x \), and if each \( f_j \in C^N \), then the \( N \)-jet of each \( f_j \) at \( \varphi_j(x) \) is uniquely determined as a linear function of the 1-jets of the three functions \( f_k \) at \( \varphi_k(x) \). If \( f \) is a formal power series at \( \Phi(x) \), by which we mean that each \( f_j \) is a formal power series at \( \varphi_j(x - \bar{x}) \), then as a formal power series in \( x - \bar{x} \), \( G_f \) vanishes identically if and only if the full formal power series of each \( f_j \) at \( \varphi_j(\bar{x}) \) equals a certain linear function of the 1-jets of the three functions \( f_k \) at \( \varphi_k(\bar{x}) \).

Composition, multiplication, and addition of formal power series are well-defined. We claim that if \( f \) is a formal power series at \( \Phi(\bar{x}) \), with the property that \( G_f = 0 \) as a formal power series in \( x - \bar{x} \), then \( f \) has positive radius of convergence. That is, any formal power series solution of \( G_f = 0 \) is real analytic. This claim will be shown below, and assumed here.

Consider any \( \bar{x} \in B \). Define \( V_N \) to be the vector space of all \( N \)-jets \( f \) at \( \Phi(\bar{x}) \) such that the associated \( N \)-jet \( G \) vanishes. These are nested in the natural sense. That is, the image of \( V_N \) under the natural map from \( N \)-jets to \( N - 1 \)-jets is contained in \( V_{N-1} \). This map is injective for every \( N \geq 2 \), since as we have shown, partial derivatives of order \( N \) are determined as linear functions of lower-order partial derivatives. Thus \( \dim(V_N) \leq \dim(V_{N-1}) \) for every \( N \geq 2 \). Therefore either there exists \( M \) such that \( V_M = \{0\} \), or there exist \( M \) and \( D \) such that \( V_N \) has dimension \( D \) for every \( N \geq M \).

Let \( \bar{x} \in B \), and \( N \geq 2 \). If \( V_N(\bar{x}) = \{0\} \) then there exists a neighborhood \( V \) of \( \bar{x} \) such that \( V_N(x) = \{0\} \) for every \( x \in V \). Indeed, \( V\phi_k(x) = \{0\} \) if and only if a certain linear mapping is injective, and injectivity is stable under small perturbations.

On the other hand, if \( \dim(V_N(\bar{x})) = \dim(V_M(\bar{x})) > 0 \) for every \( N \geq M \) then for each \( N \geq M \), every \( N \)-jet \( g \) satisfying \( G_g(\bar{x}) = 0 \) in the sense of \( N \)-jets, extends in a unique way to an \( N + 1 \)-jet with the corresponding property. Beginning with a nonzero element of \( V_M(\bar{x}) \) and iterating this process, produces a nonzero formal power series solution of \( G_{g^*} = 0 \). By the claim announced above, the formal power series \( g^* \) defines a \( C^\infty \) solution \( f \) of \( G_f \equiv 0 \) in some neighborhood of \( \bar{x} \). Since the \( M \)-jet of \( f \) at \( \bar{x} \) is nonzero, \( f \) does not vanish identically. Therefore \( f \) violates the main hypothesis of Theorem. Thus only the case in which \( V_N(\bar{x}) = \{0\} \) for all sufficiently large \( N \), can arise.

We have seen that triviality of \( V_M(x) \) is an open condition in \( x \). Therefore by compactness, there exists a single \( M \) satisfying this conclusion for every point \( x \) in the closed ball \( B \). \( \square \)

It remains to prove the claim. Choose any real analytic curve \((-1, 1) \ni t \mapsto \gamma(t) \in B \) with \( \gamma(0) = \bar{x} \) and \((\varphi_j \circ \gamma)'(t) \neq 0 \) for every \((j, t)\). Consider any function

\[
u(t) = ((g_k \circ \varphi_k \circ \gamma(t), g_k' \circ \varphi_k \circ \gamma(t) : k \in \{1, 2, 3\})\]

where each \( g_k \) is an unknown function defined in a neighborhood of \( \varphi_k(\bar{x}) \). By the above discussion, if \( f \) is smooth and \( G_f \equiv 0 \) in a neighborhood of \( \bar{x} \) then the \( \mathbb{R}^2 \)-valued function \( u \) defined by
\[
u(t) = ((f_k \circ \varphi_k \circ \gamma(t), f'_k \circ \varphi_k \circ \gamma(t)) : k \in \{1,2,3\})
\]
satisfies a certain linear ordinary differential equation
\[
\frac{d}{dt} u(t) = H(u(t))
\]
in a neighborhood of \( t = 0 \). The coefficients of this equation are real analytic functions of \( t \), and are determined by \( (a, \Phi) \) and the choice of \( \gamma \).

If \( f \) is a formal power series at \( \Phi(\bar{x}) \), and if \( G_f \equiv 0 \) in the sense of formal power series, then the formal power series \( v \) defined by
\[
v(t) = ((f_k \circ \varphi_k \circ \gamma, f'_k \circ \varphi_k \circ \gamma) : k \in \{1,2,3\})
\]
is a solution of the same differential equation (13.3), in the sense of formal power series.

Since (13.3) is a linear ordinary differential equation with analytic coefficients and non-vanishing leading-order coefficient, there exists a real analytic solution \( u \) of (13.3) in a neighborhood of \( t = 0 \), with the initial condition \( u(0) = ((f_k \circ \varphi_k \circ \gamma(0), f'_k \circ \varphi_k \circ \gamma(0)) : k \in \{1,2,3\}) \). This solution is unique. Its Taylor series at \( t = 0 \) must coincide with the formal power series \( v \). Therefore the formal power series \( t \mapsto f_k \circ \varphi_k \circ \gamma \) is convergent in a neighborhood of \( t = 0 \), for each \( k \in \{1,2,3\} \). Since \( \varphi_k \circ \gamma \) is a locally invertible mapping from \( \mathbb{R}^1 \) to \( \mathbb{R}^1 \) with real analytic inverse, the formal power series \( f_k \) are likewise convergent. \( \Box \)

We next formulate a variant, which bears the same relationship to a hypothesis used in [1] as does Lemma 13.2 to the main hypothesis of Theorem 1.1. In Lemma 13.3 the notation is \( G_f(x) = \sum_{j=1}^{3} (f_j \circ \varphi_j)(x) \); the coefficients \( a_j \) are identically equal to 1.

**Lemma 13.3.** Let \( \varphi_j \) be real analytic mappings from \( \tilde{B} \) to \( \mathbb{R}^1 \). Suppose that for any \( i \neq j \), \( \nabla \varphi_i \) and \( \nabla \varphi_j \) are linearly independent at each point of \( B \). Assume that for any connected open set \( U \subset \tilde{B} \) and any \( f \in C^\infty(\Phi(U)) \) satisfying \( G_f \equiv 0 \) in \( U \), the functions \( f_j \circ \varphi_j \) are all constant in \( U \). There exist a positive integer \( M \) and a positive real number \( C \) such that for any \( \bar{x} \in B \), for any \( f \in C^M \) in a neighborhood of \( \Phi(\bar{x}) \) satisfying \( f_j(\varphi_j(\bar{x})) = 0 \) for each index \( j \),
\[
(13.4) \quad \sum_{j=1}^{3} \sum_{0 < k \leq M} |f_j^{(k)}(\varphi_j(\bar{x}))| \leq C \sum_{0 < |\alpha| \leq M} \left| \frac{\partial^\alpha G_f}{\partial x^\alpha}(\bar{x}) \right|
\]

Here \( f_j^{(k)} \) is the \( k \)-th order derivative of \( f_j \). The sum on the right-hand side extends over the indicated multi-indices \( \alpha \), with \( |\alpha| = |(\alpha_1, \alpha_2)| = \alpha_1 + \alpha_2 \). The conclusion for a fixed \( M, \bar{x} \) is equivalent to the assertion that if \( G_f - G_{\tilde{f}}(\bar{x}) \) vanishes to order \( M \) at \( \bar{x} \) then each \( f_j - f_j(\varphi_j(\bar{x})) \) vanishes to order \( M \) at \( \varphi_j(\bar{x}) \). With \( M \) held fixed, validity of this assertion at \( \bar{x}_0 \) implies its validity at all \( \bar{x} \) in some neighborhood of \( \bar{x}_0 \), and implies the inequality with \( C \) independent of \( \bar{x} \) in that neighborhood.

The proof of Lemma 13.3 is almost identical to that of Lemma 13.2 and is consequently omitted. \( \Box \)

**14. Supplementary remarks**

We formulate a more general version of Proposition 8.1, which is proved in the same way as that proposition. It will not be used in this paper, so details of its proof are omitted.
Proposition 14.1. Let $B \subset \mathbb{R}^3$ be a closed ball of positive, finite radius, and let $\hat{B}$ be an open neighborhood of $B$. Let $b_j : \hat{B} \to \mathbb{R}$ be nowhere vanishing $C^\omega$ functions. Let $\psi_1, \psi_2 : \hat{B} \to \mathbb{R}$ be $C^\omega$ submersions whose gradients are everywhere linearly independent. Assume that there exist nowhere vanishing $C^\omega$ vector fields $V_j$ in $\hat{B}$ that satisfy $V_j(\psi_j) \equiv 0$, and satisfy the bracket condition at each point of $B$. Let sublevel sets $S(g, \varepsilon)$ be defined as in (14.1).

There exist $h = (h_1, h_2) \in C^\omega$ satisfying $\sum_{j=1}^2 b_j(x)(h_j \circ \psi_j)(x) \equiv 0$ in $B \times (0, \infty)$, and $C < \infty$, with the following property. For any $\varepsilon > 0$ and any measurable functions $g_j : \psi_j(B) \to \mathbb{R}$ there exist $\alpha \in \mathbb{R}$ and a measurable subset $S \subset S(g, \varepsilon)$ satisfying $|S| \geq C^{-1}|S(g, \varepsilon)|^C$

such that

\[
|g_j(y) - \alpha h_j(y)| \leq C\varepsilon \quad \forall y \in \psi_j(S) \quad \forall j \in \{1, 2\}.
\]

In particular, the vector space of all solutions of $\sum_{j=1}^2 b_j \cdot (g_j \circ \psi_j) \equiv 0$ in any connected nonempty open set is a vector space of dimension $\leq 1$. If the dimension equals 0 then (14.1) states that $|g_j(y)| = O(\varepsilon)$ for every $y \in \psi_j(S)$.

The next result may be useful in the treatment, in future work, of those rare data $(a, \Phi)$ for which the main hypothesis holds, but the auxiliary hypothesis does not.

Observation 14.2. Let $a, \Phi$ be $C^\omega$. Assume that each $a_j$ vanishes nowhere, and that $\nabla \varphi_1, \nabla \varphi_2$ are everywhere linearly independent. Let $B$ be connected and simply connected. Let $U \subset B \times (0, \infty)$ be a nonempty connected open set. Let $g_j \in C^\omega(\psi_j(U))$ satisfy $\sum_{j=1}^2 (a_j \circ \pi) \cdot (g_j \circ \psi_j) \equiv 0$ in $U$. Then each function $g_j$ extends to a real analytic function in $\varphi_j(B) \times (0, \infty)$.

By analytic continuation, then, the extensions satisfy $\sum_{j=1}^2 (a_j \circ \pi) \cdot (g_j \circ \psi_j) \equiv 0$ in $B$. For equations of this type, existence of local solutions is equivalent to existence of global solutions.

Proof. Let $U' \subset \mathbb{R}^2$ be a nonempty connected open subset such that $U'' \times I \subset U$ for some nonempty open interval $I \subset \mathbb{R}$. We have shown that there exists $\tau \in \mathbb{R}$ such that for $j = 1, 2$, $g_j(y, t) = h_j(y) e^{t\tau}$ on $U'' \times I$ for some $h_j \in C^\omega(\varphi_j(U'))$. These functions $h_j$ satisfy $\sum_{j=1}^2 b_j(x)(h_j \circ \varphi_j)(x) \equiv 0$ in $U''$, with $b_j(x) = a_j(x)(W_3 \varphi_j(x))^\tau$ being nowhere vanishing real analytic functions.

It suffices to show that each $h_j$ extends to a $C^\omega$ function $\tilde{h}_j : \varphi_j(B) \to \mathbb{R}$; by the connectivity of $B$, it then follows that $\sum_{j=1}^2 b_j \cdot (\tilde{h}_j \circ \varphi_j)$ vanishes identically on $B$.

For $j \in \{1, 2\}$ let $W_j$ be nowhere vanishing $C^\omega$ vector fields satisfying $W_j(\varphi_j) \equiv 0$. Let $(i, j)$ be any permutation of $(1, 2)$. Consider any curve $\Gamma$ parametrized by $s \mapsto e^{sW_j x}$ with $x \in U'$. Suppose that $e^{sW_j x} \in B$ for every $s$ belonging to some open interval $J \subset \mathbb{R}$ that contains 0. Then

\[
(h_i \circ \varphi_i)(x) \equiv -a_j(x)a_i(x)^{-1}(h_j \circ \varphi_j)(x)
\]

for all $x \in \Gamma$ in some neighborhood of $x$. That is,

\[
(h_i \circ \varphi_i)(e^{sW_j x}) \equiv -a_j(e^{sW_j x})a_i(e^{sW_j x})^{-1}(h_j \circ \varphi_j)(e^{sW_j x})
\]

for all $s$ in some neighborhood of 0. The right-hand side is a well-defined real analytic function of $s$ on the full interval $J$, since $\varphi_j(e^{sW_j x})$ is constant and consequently $(h_j \circ \varphi_j)(e^{sW_j x})$ is likewise constant. Thus the function $s \mapsto (h_i \circ \varphi_i)(e^{sW_j x})$ extends to a real
analytic function on the full interval $J$. The mapping $\varphi_i$ restricts to an invertible real analytic mapping from $J$ to the interval $\{e^{sW_2}: s \in J\}$. Therefore $h_i$ extends to a real analytic function on that same interval.

For any two points in $B$ there exists a continuous path in $B$ from one point to the other, consisting of a concatenation of integral curves of $W_1$ and $W_2$. Since each $h_k \circ \varphi_i$ is constant along integral curves of $W_k$, it follows that $h_i$ extends to a real analytic function on the full interval $\varphi_j(B)$.

The equation $\sum_{j=1}^3 a_j \cdot (f_j \circ \varphi_j) = 0$ shares this extension property.

**Observation 14.3.** Let $a, \Phi$ be $C^\infty$. Assume that each $a_j$ vanishes nowhere, and that $\nabla \varphi_i, \nabla \varphi_j$ are everywhere linearly independent for all $i \neq j \in \{1, 2, 3\}$. Let $B \subset \mathbb{R}^2$ be connected and simply connected, and let $U \subset B$ be nonempty, connected, and open. Let $f_j : \varphi_j(U) \to \mathbb{R}$ be $C^\infty$ and satisfy $\sum_{j=1}^3 a_j \cdot (f_j \circ \varphi_j) \equiv 0$ in $U$. Then each $f_j$ extends to a real analytic function in $\varphi_j(B)$.

**Proof.** Assume without loss of generality that $a_3 \equiv 1$. Apply $W_3$ to obtain

$$\sum_{j=1}^2 \left[W_3 a_j \cdot (f_j \circ \varphi_j) + a_j \cdot W_3 \varphi_j \cdot (f'_j \circ \varphi_j)\right] = 0 \text{ in } U.$$

On any integral curve of $W_2$ in any open subset of $B$ in which both functions $f_j \circ \varphi_j$ are analytic, the two terms with $j = 2$ are constant, and this equation becomes an inhomogeneous linear ordinary differential equation for the restriction of $f_1$ to an open interval. The coefficients of this differential equation are analytic, and the leading coefficient, which is $a_1 \cdot W_3 \varphi_1$ composed with an invertible analytic mapping, vanishes nowhere. Therefore $f_1$ continues analytically to the image under $\varphi_1$ of any integral curve of $W_2$ that passes through $U$.

The proof of analytic extendability for $j = 1, 2$ is completed as in the final paragraph of the preceding argument. For $j = 3$, it suffices to permute the indices. □

**Observation 14.4.** Under the hypotheses of Observation 14.3, the vector space of all real analytic solutions $f$ of the equation $\sum_{j=1}^3 a_j \cdot (f_j \circ \varphi_j) \equiv 0$ in $B$ has finite dimension.

**Proof.** The vector space in question has dimension less than or equal to 3. Indeed, consider any point $\bar{x} \in B$, and consider the integral curve $\Gamma$ of $W_2$ that passes through $\bar{x}$. The restriction of $f_1 \circ \varphi_1$ to $\Gamma$ — equivalently, the restriction of $f_1$ to the subinterval $\varphi_1(\Gamma)$ of its domain — is uniquely determined by the three scalar quantities $f_1(\varphi_1(\bar{x})), f_2(\varphi_2(\bar{x})), f'_2(\varphi_2(\bar{x}))$, as well as by the fixed data $a, \Phi$, through the ordinary differential equation that it satisfies. The restriction of $f_2 \circ \varphi_2$ to $\Gamma$ is constant, so both functions $f_j \circ \varphi_j|_\Gamma$ are determined by the initial three quantities.

Consider the integral curve $\Gamma'$ of $W_1$ that passes through any point $\bar{x}' \in \Gamma$. The restriction of $f_2 \circ \varphi_2$ to $\Gamma'$ is uniquely determined by the three scalar quantities $f_2(\varphi_2(\bar{x}')) = f_2(\bar{x}), f_1(\varphi_1(\bar{x}'))$, and $f'_1(\varphi_1(\bar{x}'))$, and thus by $f_1(\varphi_1(\bar{x})), f_2(\varphi_2(\bar{x})), f'_2(\varphi_2(\bar{x}))$. Thus both functions $f_j \circ \varphi_j$ are uniquely determined at $e^{s_1 W_1} e^{s_2 W_2} \bar{x}$ for all $(s_1, s_2)$ in an open neighborhood of $(0, 0)$. This reasoning can be iterated. Thus for $j = 1, 2$, $f_j$ is uniquely determined in a neighborhood of $\varphi_j(\bar{x})$, and hence by analytic continuation in the connected domain $\varphi_j(B)$.

The same conclusion is reached for $f_3$ by using the relation $\sum_{j=1}^3 a_j \cdot (f_j \circ \varphi_j) = 0$ to solve for $f_3(\varphi_3(\bar{x}))$, then repeating the above analysis with the indices permuted. □
15. Approximation by regular functions

The proof of Theorem 15.1 has implicitly established the following intermediate result, in which the main hypothesis of Theorem 15.1 is not assumed; nor is any lower bound assumed for $|f|$. We write $\alpha_{j,\omega}(y) = \alpha_j(y, \omega)$ and $\beta_{j,\omega}(y) = \beta_j(y, \omega)$ to emphasize that these are regarded as functions of $y$ that depend on a parameter $\omega$.

**Theorem 15.1.** Let $B \subset \mathbb{R}^2$ be a closed ball, and let $\tilde{B}$ be an open neighborhood of $B$. Let $a, \Phi \in C^\infty(\tilde{B})$. Assume that none of the mappings $\varphi_j$ are constant on $B$, and that for each $i \neq j \in \{1, 2, 3\}$, $\nabla \varphi_i(x)$ and $\nabla \varphi_j(x)$ are not everywhere linearly dependent on $B$. Assume that the curvature of the web defined by $\Phi$ does not vanish identically. Let $(a, \Phi)$ be nondegenerate and satisfy the auxiliary hypothesis.

There exist $N < \infty$, a compact set $\Omega \subset \mathbb{R}^N$, open neighborhoods $U$ of $B$ and $\tilde{\Omega}$ of $\Omega$, and $C^\infty$ functions $\alpha_j, \beta_j : \varphi_j(U) \times \tilde{\Omega} \to \mathbb{R}$, such that for each $\rho > 0$ there exist $c, C, \tau \in (0, \infty)$ with the following property.

Let $\varepsilon > 0$ be arbitrary. Let $f$ be an ordered triple of Lebesgue measurable functions satisfying $\|f_j\|_{L^\infty} \leq 1$. Either

\begin{equation}
|S(f, \varepsilon)| \leq C' \varepsilon^\tau \tag{15.1}
\end{equation}

or there exist a measurable set $S \subset S(f, \varepsilon)$ satisfying

\begin{equation}
|S| \geq c|S(f, \varepsilon)|^C \tag{15.2}
\end{equation}

and a parameter $\omega \in \Omega$ such that for each $j \in \{1, 2, 3\}$,

\begin{equation}
\begin{cases}
|f_j(y) - \frac{\alpha_{j,\omega}(y)}{\beta_{j,\omega}(y)}| \leq C \varepsilon^{1-\rho} \quad \forall y \in \varphi_j(S) \\
|\beta_{j,\omega}(y)| \geq \varepsilon^\rho \quad \forall y \in \varphi_j(S).
\end{cases} \tag{15.3}
\end{equation}

Results of this type may play a role in the removal of auxiliary hypotheses such as those in Theorems 15.1 and 16.1.

16. A special case with arbitrarily many summands

In this section we consider sublevel sets $S(f, \varepsilon)$ for sums $\sum_{j=1}^n a_j(x) (f_j \circ \varphi_j)(x)$ with an arbitrary number $n$ of terms, in the special case in which all mappings $\varphi_j$ are linear. Such a situation is specified by a datum

$$\mathcal{D} = (n, \Phi, a) = (n, \{\varphi_j : 1 \leq j \leq n\}, \{a_j : 1 \leq j \leq n\}).$$

Below, we define a family of derived data $\mathcal{D}^*$ associated to a datum $\mathcal{D}$. This family is finite, but has cardinality very roughly of size $2^{2^2}$ with an exponential tower of height $n - 2$. To $\mathcal{D}$ is associated the linear functional $\sum_{j=1}^n a_j(x) (f_j \circ \varphi_j)(x) = 0$; to each derived datum $\mathcal{D}^*$ is associated a corresponding functional.

16.1. Formulation of the inequality. The following result applies to sublevel set inequalities with arbitrarily many summands.

**Theorem 16.1.** Let $B \subset \mathbb{R}^2$ be a closed ball of positive, finite radius. Let $\tilde{B} \subset \mathbb{R}^2$ be an open neighborhood of $B$. Let $n \geq 1$. For each $j \in \{1, 2, \ldots, n\}$ let $\varphi_j : \mathbb{R}^2 \to \mathbb{R}$ be a surjective linear mapping, and let $a_j : \tilde{B} \to \mathbb{R}$ be real analytic. Let $\mathcal{D} = (n, \{\varphi_j\}, \{a_j\})$.

Suppose that for any two distinct indices $i \neq j \in J$, $\nabla \varphi_i, \nabla \varphi_j \in \mathbb{R}^2$ are linearly independent. Suppose that none of the coefficients $a_j$ vanish identically in $B$. Suppose that for any nonempty open set $U \subset \tilde{B}$, and for $\mathcal{D}$ as well as for any datum $\mathcal{D}^*$ associated to $\mathcal{D}$,
any real analytic solution \( f \) of the linear equation associated to \( \mathcal{D}^* \) vanishes identically in \( \Phi(U) \).

Then there exist \( C < \infty \) and \( \tau > 0 \) such that for any Lebesgue measurable \( f \) and any \( \varepsilon > 0 \),
\[
(16.1) \quad |\{ x \in S(f, \varepsilon) : \sum_j |f_j \circ \varphi_j(x)| \geq 1 \}| \leq C\varepsilon^\tau.
\]

The proof of this theorem combines a simplification of the proof of Theorem 1.1 with a recursion. The complexity of this recursion increases quite rapidly as the number of summands increases. We conjecture that the conclusion holds without any auxiliary hypotheses involving associated data. Since each datum \( \mathcal{D}^* \) gives rise to its own auxiliary hypothesis, and since the number of data associated to a given datum \( \mathcal{D} \) grows so rapidly as the order of \( \mathcal{D} \) increases, it would be desirable to eliminate these auxiliary hypotheses.

By a ratio of real analytic functions we will mean such a ratio, with denominator not vanishing identically in any nonempty open set. Such a ratio is well-defined on the complement of an analytic variety of positive codimension.

16.2. Associated data and associated linear equations. In order to complete the statement of Theorem 16.1, we next introduce the associated data and their associated linear equations. Let \( n \geq 1 \). Let \( J \) be an index set of cardinality \( n \). For each \( j \in J \) let \( \varphi_j : \mathbb{R}^2 \to \mathbb{R}^1 \) be linear and surjective, and let \( a_j : \tilde{B} \to \mathbb{C} \) be a ratio of real analytic functions that does not vanish identically in any nonempty open set.

Let \( W_j \) be nonzero constant-coefficient real vector fields that satisfy \( W_j(\varphi_j) \equiv 0 \). We are interested in sublevel sets
\[
S(f, \varepsilon) = \{ x \in B : |\sum_j a_j(x)(f_j \circ \varphi_j)(x)| < \varepsilon \}.
\]

Our main hypothesis is that there are no nonzero \( C^\infty \) solutions, in any nonempty open set, of the equation
\[
(16.2) \quad \sum_{j \in J} a_j(x)(f_j \circ \varphi_j)(x) = 0.
\]

However, the analysis developed below requires certain auxiliary hypotheses, which are formulated in terms of certain associated equations that generalize (16.2). These associated equations take the general form
\[
(16.3) \quad \sum_{j=1}^m \sum_{k_j=1}^{N_j} a_{j,k_j}(x)(f_{j,k_j} \circ \varphi_j)(x) = 0.
\]

They are constructed recursively from \( n, \{ \varphi_j \}, \{ a_j \} \) as follows.

Begin with a more general datum
\[
\mathcal{D} = (n, \{ \varphi_j : 1 \leq j \leq n \}, \{ N_j : 1 \leq j \leq n \}, \{ a_{j,k_j} : 1 \leq k_j \leq N_j \})
\]
for any \( n, N_j \in \mathbb{N} \). Construct an associated datum, as follows.

If \( N_n = 1 \) set \( m = n - 1 \) and \( \varphi_j^* = \varphi_j \) for each \( j \leq m \). Set \( N_j^* = 2N_j \) for each \( j \leq m \).

For each \( j \leq m \) define
\[
a_{j,k_j}^* = a_{j,k_j}/a_{n,k_{N_n}} \quad \forall 1 \leq k_j \leq N_j
\]
\[
a_{j,k_j}^* = W_n(a_{j,k_j}/a_{n,k_{N_n}}) \quad \forall N_j < k_j \leq 2N_j.
\]
If \( N_n > 1 \) set \( m = n \). Set \( \varphi_j^* = \varphi_j \) for each \( j \leq m \). Set \( N_j^* = 2N_j \) for each \( j \leq n - 1 \), and \( N_m^* = N_n - 1 \). For each \( j \in \{1, 2, \ldots, n - 1\} \) and each \( k_j \in \{1, 2, \ldots, N_j\} \), define \( a_{j,k_j}^* \) as in the preceding paragraph. Set
\[
a_{n,k_n}^* = a_{n,k_n}/a_{n,k_{N_n}} \quad \forall k_n < N_n.
\]
The datum \( D^* = (m, \{\varphi_j^*\}, \{N_j^*\}, \{a_{j,k_j}^*\}) \) is said to be associated to \( D = (n, \{\varphi_j\}, \{N_j\}, \{a_{j,k_j}\}) \).

Applying this association rule recursively, and making association a transitive relation, generates a finite collection of data associated to \( D \). We apply this association rule with \( (1, 2, \ldots, n) \) replaced by an arbitrary permutation of itself, and with the indices \( k_j \) also arbitrarily permuted for each index \( j \), so that any associated datum \( D^* \) for any such permutation is also associated to \( D \).

### 16.3. Formulation of the theorem for general data \( D^* \).

Theorem 16.1 is the special case in which \( N_j = 1 \) for every index \( j \) of the following result. We write
\[
f = (f_{j,k_j} : j \in J \text{ and } 1 \leq k_j \leq N_j)
\]
and
\[
(16.4) \quad S(f, \varepsilon) = \{x \in B : \left| \sum_{j=1}^{n} \sum_{k_j=1}^{N_j} a_{j,k_j}(x)(f_{j,k_j} \circ \varphi_j)(x) \right| < \varepsilon \}.
\]

**Theorem 16.2.** Let \( B \subset \mathbb{R}^2 \) be a closed ball of positive, finite radius. Let \( \tilde{B} \subset \mathbb{R}^2 \) be an open neighborhood of \( B \). Let \( n \geq 1 \). For each \( j \in \{1, 2, \ldots, n\} \) let \( N_j \in \mathbb{N} \). For each \( j \in \{1, 2, \ldots, n\} \) let \( \varphi_j : \mathbb{R}^2 \to \mathbb{R}^1 \) be a surjective linear mapping. For each \( j \in \{1, 2, \ldots, n\} \) and each \( k \in \{1, 2, \ldots, N_j\} \) let \( a_{j,k} : \tilde{B} \to \mathbb{C} \) be a ratio of real analytic functions. Let \( J = \{1, 2, \ldots, n\} \) and \( D = (J, \{\varphi_j\}, \{N_j\}, \{a_{j,k}\}) \).

Suppose that for any two distinct indices \( i \neq j \in J \), \( \nabla \varphi_i, \nabla \varphi_j \in \mathbb{R}^2 \) are linearly independent. Suppose that none of the coefficients \( a_{j,k} \) vanish identically in \( B \). Suppose that for any nonempty open set \( U \subset \tilde{B} \), and for \( D \) as well as for any datum \( D^* = (n^*, \{\varphi_j^*\}, \{N_j^*\}, \{a_{j,k}^*\}) \) associated to \( D \), any real analytic solution \( f \) of
\[
\sum_{j=1}^{n^*} \sum_{k_j=1}^{N_j^*} a_{j,k_j}^*(x)(f_{j,k} \circ \varphi_j^*)(x) = 0 \quad \text{in } U
\]
satisfies \( f_{j,k} \equiv 0 \) in \( \varphi_j(U) \) for all \( (j, k) \).

Then there exist \( C < \infty \) and \( \tau > 0 \) such that for any Lebesgue measurable \( f \) and any \( \varepsilon > 0 \),
\[
(16.5) \quad |\{x \in S(f, \varepsilon) : \sum_{j=1}^{n} \sum_{k_j=1}^{N_j} |f_{j,k} \circ \varphi_j(x)| \geq 1\}| \leq C\varepsilon^\tau.
\]

It can be shown that in the sense of jets of sufficiently high order at a point, generic data \( D \) do satisfy all of the hypotheses of the theorem.

### 16.4. The case \( n = 1 \).

In this subsection we prove Theorem 16.2 for \( n = 1 \). With only one mapping \( \varphi = \varphi_1 \), the linearizability hypothesis is tautologous wherever \( \nabla \varphi \neq 0 \). The recursion on \( N_1 \) is now straightforward, and no auxiliary hypotheses are needed. The statement is as follows.
Proposition 16.3. Let \( B \subset \mathbb{R}^2 \) be a closed ball of finite, positive radius. Let \( \tilde{B} \) be an open neighborhood of \( B \). Let \( N \in \mathbb{N} \). For each \( k \in \{1, 2, \ldots, N\} \) let \( a_k : \tilde{B} \to \mathbb{R} \) be real analytic. Let \( \varphi : \tilde{B} \to \mathbb{R} \) be real analytic, and be nonconstant on \( B \). Suppose that for any nonempty open set \( U \subset \tilde{B} \) and any \( f = (f_1, \ldots, f_N) \in C^2(U) \), if \( \sum_{k=1}^{N} a_k(x)(f_k \circ \varphi)(x) = 0 \) in \( U \) then \( f \) vanishes identically in \( \varphi(U) \). Then there exist \( \tau > 0 \) and \( C < \infty \) such that for any \( \varepsilon > 0 \) and any Lebesgue measurable \( f \),

\[
\{ x \in B : \left| \sum_{k=1}^{N} a_k(x)(f_k \circ \varphi)(x) \right| < \varepsilon \text{ and } |f \circ \varphi(x)| \geq 1 \} \leq C \varepsilon^\tau.
\]

Proof. Consider the case in which \( \varphi(x, y) \equiv y \). Write \( x = (x_1, \ldots, x_N) \in \mathbb{R}^N \). Define

\[
A(x, y) = (a_k(x_j, y))_{j,k=1}^{N}
\]

for those \((x, y)\) with each \((x_j, y) \in \tilde{B} \). Define

\[
\alpha(x, y) = \det(A(x, y)).
\]

We claim that \( \alpha \) does not vanish identically in \( B \). The conclusion of Proposition 16.3 follows from this claim, by reasoning developed above. Indeed, if

\[
S = \{ x \in B : \left| \sum_{k=1}^{N} a_k(x)(f_k \circ \varphi)(x) \right| < \varepsilon \text{ and } |f \circ \varphi(x)| \geq 1 \}
\]

then

\[
S^* = \{ (x, y) : (x_k, y) \in S \forall k \in \{1, 2, \ldots, N\} \}
\]

satisfies

\[
|S^*| \geq c|S|^C.
\]

If \((x, y) \in S^*\) then

\[
A(x, y) \begin{pmatrix} f_1(y) \\ \vdots \\ f_N(y) \end{pmatrix} = O(\varepsilon).
\]

Since \( |f(y)| \geq 1 \), it follows that \( \alpha(x, y) = \det(A((x, y))) \) satisfies

\[
|\alpha(x, y)| = O(\varepsilon).
\]

But if \( \alpha \) does not vanish identically, then since it is a real analytic function, \( \alpha \) satisfies

\[
|\{(x, y) : |\alpha(x, y)| \leq \delta \}| = O(\delta^\tau)
\]

for some exponent \( \tau > 0 \).

To prove the claim, suppose to the contrary that \( \det(A((x, y))) \) were to vanish identically. Then in some nonempty open set, possibly after permuting the indices \( k \), there would exist real analytic functions \( c_k(x, y) \) that did not vanish identically and that satisfied

\[
a_N(x_j, y) = \sum_{k=1}^{N-1} c_k(x, y)a_k(x_j, y)
\]

for all \( j \in \{1, 2, \ldots, N\} \) for every \((x, y) \in U \). Differentiating with respect to \( x_i \) for any \( i \neq j \) gives

\[
\sum_{k=1}^{N-1} \frac{\partial c_k(x, y)}{\partial x_i} a_k(x_j, y) \equiv 0
\]
in $U$.

We proceed by induction on $N$, the case $N = 1$ being the quintessence of triviality. For the inductive step, there are two cases. If there exists a pair of indices $i \neq j$ for which there exists $k$ such that $\frac{\partial c_k(x,y)}{\partial x_i}$ does not vanish identically in $U$, then the conclusion follows by induction on $N$. On the other hand, if $\frac{\partial c_k(x,y)}{\partial x_i}$ vanishes identically for any $i$ then each $c_k(x,y)$ is a function $c_k(y)$ of $y$ alone, and we have

$$a_N(x,y) = \sum_{k=1}^{N-1} c_k(y)a_k(x,y).$$

The main hypothesis of the Proposition is contradicted upon defining $f_N \equiv -1$ and $f_k(y) = c_k(y)$ for $k < N$.

This completes the discussion of the special case in which $\varphi_1(x,y) = \varphi(x,y) \equiv y$. The nondegenerate case, in which $\nabla \varphi$ vanishes nowhere, is reducible to this special case by a change of variables. The general case can be reduced to the nondegenerate case by the same reasoning that was used in §12 to reduce the general case of Theorem 1.1 to the nondegenerate case of that theorem. \hfill \Box

16.5. Sketch of the proof of Theorem 16.2 for $n > 1$. The proof is a variant of the proof of Theorem 1.1 with some significant differences. We will sketch it, indicating the points at which these differences arise.

The proof is by a double induction, with the outer induction on $n$ and the inner induction on $N_n$ with $n$ and $(N_j : j < n)$ held fixed. Let $W_j$ be a nonzero real analytic vector field, with constant coefficients, satisfying $W_j(\varphi_j) \equiv 0$. Let $N^*_j = N_j$ for $j < n$ and $= N_n - 1$ for $j = n$.

By the permutation invariance of the hypotheses, we may assume that

$$|f_{1,1}(y)| \geq 1 \text{ for each } y \in \varphi_1(B)$$

and that $f_{j,k} = O(1)$ for all $j,k$.

Let $\varepsilon, f$ be given. Let $\mathcal{S}$ be the set of all $x \in S(f, \varepsilon)$ such that $\sum_j \sum_{k_j} |f_{j,k} \circ \varphi_j(x)| \geq 1$. We may assume without generality that $\sum_j \sum_{k_j} |f_{j,k} \circ \varphi_j(x)| \geq 1$ for every $x \in B$. We may also assume without loss of generality that $a_{n,N_n} \equiv 1$.

For the microscale step, let a small quantity $\sigma > 0$ be given. Consider the set $S_1$ of all $(x,s)$ with $x \in B$ and $|s| \leq \varepsilon$ such that $x \in \mathcal{S}$ and $e^{sW_n}x \in \mathcal{S}$. This set satisfies $|S_1| \geq \varepsilon |\mathcal{S}|^C$. For any $(x,s) \in S_1$,

$$\sum_{j=1}^{n} \sum_{k_j=1}^{N^*_j} a_{j,k_j}(x)F_{j,k_j}(\varphi_j(x), sW_n(\varphi_j)) = O(\varepsilon)$$

where

$$F_{j,k}(y,t) = f_{j,k}(y+t) - f_{j,k}(y).$$

This is the first point at which the analysis for linear mappings $\varphi_j$ diverges from the analysis for the case of webs with nonvanishing curvature. Since each $W_n \varphi_j$ is a constant independent of $x$, (16.14) is a family of inequalities for the functions $x \mapsto F_{j,k_j}(\varphi_j(x), sW_n(\varphi_j))$; we are dealing with a one-parameter family of sublevel set inequalities for $x \in B$ parametrized by $s$, rather than with a single irreducible sublevel set inequality for $(x,s) \in B \times \mathbb{R}^1$. Moreover, the datum $(n, \{\varphi_j\}, \{N_j\}, \{a_{j,k_j}\})$ defining each inequality is independent of $s$, and is the same datum with which we began, except that $N_n$ has been replaced by $N_n - 1,$
effectively equivalent to requiring that \( f_{n,N_n} \) vanishes identically. Thus associated data do not arise in this step.

For each \( s \in [-\varepsilon, \varepsilon] \) we may apply the induction hypothesis to the tuple of functions \( \tilde{f}_{j,k}(y) = F_{j,k}(y, sW_n(\varphi_j)) \) indexed by \((j, k)\) satisfying \( 1 \leq j \leq n, 1 \leq k \leq N_j \) for \( n < N \), and \( 1 \leq k_n \leq N_n - 1 \). We conclude that the set of all \( x \in S_1 \) satisfying

\[
\sum_{j=1}^{n} \sum_{k_j = 1}^{N_j} |F_{j,k}(\varphi_j(x), sW_n \varphi_j)| \geq \varepsilon^{1-\sigma}
\]

has Lebesgue measure \( \lesssim \varepsilon^{\sigma} \), uniformly for all parameters \( s \).

Therefore either \( |S_1| \lesssim \varepsilon^{1+\sigma} \), in which case the proof is complete, or there exists \( S_2 \subset S_1 \) whose Lebesgue measure satisfies \( \varepsilon^{-1}|S_2| \gtrsim (\varepsilon^{-1}|S_1|)^C \) such that for every \( j \) and every \( k_j \leq N_j \),

\[
|f_{j,k}(\varphi_j(x) + sW_n \varphi_j) - f_{j,k}(\varphi_j(x))| = O(\varepsilon^{1-\sigma}) \quad \forall (x, s) \in S_2.
\]

By permuting the indices and repeating this argument, the same conclusion follows for \( f_{n,k,N} \). We assume henceforth that the second possibility arises.

As in the proof of Proposition \([7.2]\) it follows that there exist functions \( f^*_{j,k} \) that are constant on intervals of length \( \varepsilon \), in the same sense as in the discussion in \([47]\) such that

\[
|f_{j,k} \circ \varphi_j(x) - f^*_{j,k} \circ \varphi_j(x)| = O(\varepsilon^{1-\sigma})
\]

for every \((j, k)\), for every \( x \in S_3 \), where \( S_3 \subset S \) satisfies \( |S_3| \gtrsim |S|^C \). We may, and do, replace \( f_{j,k} \) by \( f^*_{j,k} \) henceforth.

We next adapt the mesoscale analysis. Let a small quantity \( \rho > 0 \) be given. Fix \( \gamma \in (\frac{1}{2}, 1) \) and set \( \delta = \varepsilon^\gamma \). The above reasoning can be repeated at scale \( \delta \). Therefore either the proof is again complete, or there exists \( S_4 \subset S_3 \) satisfying \( |S_4| \gtrsim |S|^C \) such that

\[
|f_{j,k} \circ \varphi_j(x) - f_{j,k} \circ \varphi_j(x')| = O(\delta^{1-\sigma})
\]

whenever \( x, x' \in S_4 \) satisfy \( |x - x'| = O(\delta) \).

Let \( \varrho > 0 \) be another small parameter, to be specified below. Consider the set \( S_5 \) of all ordered pairs \((x', x)\) in \( S_4 \times S_4 \) of the form \( x' = e^{sW_n}x \) with \( \delta^{1+\varrho} \leq |s| \leq \delta \). Define

\[
F_{j,k}(y, t) = t^{-1}(f_{j,k}(y + t) - f_{j,k}(y))
\]

for \( t \neq 0 \) and

\[
\psi_j(x, t) = (\varphi_j(x), tW_n \varphi_j).
\]

Provided that \( |S_4| \gg \delta^\varrho \), as we may assume, \( S_5 \) satisfies \( |S_5| \gtrsim \delta|S_4|^2 \). For \((x', x)\) in \( S_5 \),

\[
\sum_{j=1}^{n} \sum_{k_j = 1}^{N_j} \left[ W_n a_{j,k}(x)f_{j,k}(x) + a_{j,k}(x)W_n \varphi_j(F_{j,k} \circ \psi_j)(x, s) \right] = O(\varepsilon^{1-\gamma-\varrho})
\]

for every \((x', x) = (e^{sW_n}x, x) \in S_5 \), provided that \( \sigma, \varrho \) are sufficiently small. Note that each quantity \( W_n \varphi_j \) is a constant. We have used the bound \( (16.16) \) to ensure that

\[
|a_{j,k}(x + sW_n \varphi_j) - a_{j,k}(x)| \cdot |F_{j,k} \circ \psi_j(x, s)| = O(\delta \cdot \delta \cdot \delta^{1-\sigma} (\delta \cdot \delta^{1-\sigma})^{-1}) = O(\delta^{1-\sigma-\varrho}),
\]

which is negligible relative to \( \varepsilon^{1-\gamma-\varrho} = \varepsilon^{1-\varrho} \delta^{-1} \) for sufficiently small \( \varepsilon, \varrho, \sigma \) since \( \delta^2 = \varepsilon^{2\gamma} \) is negligible relative to \( \varepsilon^1 \) because \( \gamma > \frac{1}{2} \).
Let \( S_6 \) be the set of all \((x, s, s')\) such that \((x, s) \in S_5\) and \((x, s') \in S_5\). By the Cauchy-Schwarz inequality, \(\delta^{-2}|S_6| \gtrsim (\delta^{-1}|S_6|)^2\). For any \((x, s, s') \in S_6\),

\[
\sum_{j=1}^{n} \sum_{k_j=1}^{N_j} a_{j,k_j}(x) \left[ (W_n \varphi_j \cdot F_{j,k_j} \circ \psi_j)(x, s') - (W_n \varphi_j \cdot F_{j,k_j} \circ \psi_j)(x, s) \right] = O(\varepsilon^{1-\gamma-\theta}).
\]

Because of the special form \(\psi_j(x, s) = (\varphi_j(x), sW_n \varphi_j)\) with \(W_n \varphi_j\) constant, for each fixed ordered pair \((s', s)\), this is a sublevel set inequality in the variable \(x \in B \subset \mathbb{R}^2\), of the same form as that with which we began, except that \(N_n\) has been replaced by \(N_n^* = N_n - 1\).

Invoking the induction hypothesis for each \((s, s')\), we conclude, as in the macroscale step above, that either \(\delta^{-2}|S_6| = O(\varepsilon^c)\) for some \(c > 0\), or there exists \(S_7 \subset S_6\) satisfying \(\delta^{-2}|S_7| \gtrsim (\delta^{-2}|S_6|)^C\) such that

\[
(16.17) \quad |(F_{j,k_j} \circ \psi_j)(x, s') - (F_{j,k_j} \circ \psi_j)(x, s)| = O(\varepsilon^{(1-\gamma-\theta)(1-\sigma)}) \quad \forall j \leq n \quad \forall k_j \leq N_j^*
\]

for every \((x, s', s) \in S_7\). The same reasoning can be repeated with the roles of the indices permuted to yield the same conclusion for the index \((n, N_n)\) as well, on a set \(S_8 \subset S_7\) whose Lebesgue measure satisfies \(\delta^{-2}|S_8| \gtrsim (\delta^{-2}|S_7|)^C\).

Choose \(\sigma, \rho\) so that \((1-\gamma-\theta)(1-\sigma) > \gamma\). It follows from (16.17), as in the proof of (13.37), that there exist \(S_9 \subset S\) of measure \(|S_9| \gtrsim |S|^C\) and functions \(f_{j,k_j}^*\) that are affine on intervals of lengths \(\varepsilon^\gamma\) and satisfy

\[
|f_{j,k_j}^*(\varphi_j(x)) - f_{j,k_j}^*(\varphi_j(x))| \leq C\varepsilon^{-\rho} \quad \forall x \in S \quad \forall j \in \{1, 2, 3\}.
\]

Moreover, the derivatives \((f_{j,k_j}^*)'\) satisfy

\[
|(f_{j,k_j}^*)'| \leq C\varepsilon^{-\rho}.
\]

These constants \(C\) depend on \(\rho\). We may replace each \(f_{j,k_j}\) by \(f_{j,k_j}^*\) henceforth.

We next carry out the macroscale step. It is here that associated data arise. Let \(\varrho > 0\) be another sufficiently small parameter, to be chosen below. By the same reasoning as in (10) it suffices to show that the set \(S\) defined to be

\[
S = \left\{ x \in B : \left| \sum_{j=1}^{n} \sum_{k_j=1}^{N_j} W_n a_{j,k_j}(x) (f_{j,k_j} \circ \varphi_j)(x) + a_{j,k_j}(x) W_n \varphi_j(x) (g_{j,k_j} \circ \varphi_j)(x) \right| < \varepsilon^\varrho \right\}
\]

satisfies \(|S| = O(\varepsilon^c)\) for some \(c = c(\varrho) > 0\). Here \(g_{j,k_j}\) arises as the derivative of \(f_{j,k_j}\), but as in (10) we regard it as an independent function, not necessarily related to \(f_{j,k_j}\) in any way. This sum is interpreted under the convention that if \(N_n = 1\) then the summation extends only over \(j \in \{1, 2, \ldots, n-1\}\).

Rewriting this sum by setting \(N_j^* = 2N_j\) for \(j < n\), and for \(j < n\), \(f_{j,k_j} = g_{j,k_j} - N_j\) for \(N_j < k_j \leq N_j^*\). For \(j = n\), \(W_n \varphi_j = 0\) so the contribution of \(j = n\) is \(\sum_{k_n \leq N_n-1} a_{n,k_n}(x) \left( f_{n,k_n} \circ \varphi_n \right)(x) \). Thus the sum becomes

\[
\sum_{j=1}^{n} \sum_{k_j=1}^{N_j\ast} a_{j,k_j}(x) \left( f_{j,k_j} \circ \varphi_j \right)(x)
\]

if \(N_n > 1\), and becomes

\[
\sum_{j=1}^{n-1} \sum_{k_j=1}^{N_j\ast} a_{j,k_j}(x) \left( f_{j,k_j} \circ \varphi_j \right)(x)
\]
if $N_n = 1$. This is a sum for the datum
\[
D^* = (n, \{ \varphi_j : j \leq n \}, \{ N_j \}, \{ a_{j,k_j} : j \leq n \text{ and } k_j \leq N_j \})
\]
if $N_n > 1$, and for the datum
\[
D^* = (n - 1, \{ \varphi_j : j \leq n - 1 \}, \{ N_j \}, \{ a_{j,k_j} : j \leq n - 1 \text{ and } k_j \leq N_j \})
\]
if $N_n = 1$. In each case, the datum $D^*$ is associated to $D$.

Thus we have arrived at a sublevel set inequality, not for the given datum but rather for an associated datum $D^*$ under one step of the recursion defined above. By the inductive hypothesis, a sublevel set inequality holds for this associated datum. Since $|f_{1,1}(y)| \geq 1$ for every $y \in \varphi_1(B)$, the inductive hypothesis ensures that $|S|$ satisfies such a upper bound, completing the proof of Theorem 16.2. \hfill \Box

It is the author’s hope that the analysis developed in this paper provides an outline for a treatment of sums $\sum_{j=1}^n a_j \cdot (f_j \circ \varphi_j)$ with general mappings $\varphi_j$ and with an arbitrary number of summands, by a recursive argument, albeit under increasingly complicated auxiliary hypotheses as $n$ increases. Moreover, it would be desirable to remove the auxiliary hypotheses, which for linear mappings are encoded by associated data, and which would take a still more complicated form for general mappings. As of this writing, work in this direction is underway.

17. ANOTHER FORM OF DEGENERACY

In §11 we stated that an excluded degenerate case of Theorem 11.1, in which two or more mappings $\varphi_j$ have gradients that are everywhere linearly dependent, can be treated by a modification of the main argument. No auxiliary hypothesis is needed for those cases. In this section, we justify that statement.

There are two subcases, of which the first, the subcase in which the gradients of all three mappings $\varphi_j$ are everywhere linearly dependent, was treated above in §16.4. The remaining subcase, in which $\nabla \varphi_2, \nabla \varphi_3$ are everywhere linearly dependent but $\nabla \varphi_1$ is linearly independent of these at generic points, is a special case of the case $n = 2$ of Theorem 16.2 if it is possible to change variables so that all three mappings become linear.

In the general case of this remaining subcase, a change of notation we have a sum of the type appearing in (16.4) with $n = 2$, $N_1 = 1$, and $N_2 = 2$. Assuming for simplicity that $\nabla \varphi_j$ vanish nowhere, and that $\{ \nabla \varphi_1, \nabla \varphi_2 \}$ is everywhere linearly independent, we may change variables so that $\varphi_j(x_1, x_2) \equiv x_j$. By fixing $N_1$ values of $x_2$ and allowing $x_1$ to vary freely, we obtain a linear system of $N_1$ approximate equations for the $N_1$ quantities $f_{1,k_1}(x_1)$. This can be exploited by proceeding as in the discussion beginning with the matrix equation (10.13) in the macroscale analysis of §10. See also §16.4.

The more general case, in which no $\nabla \varphi_j$ vanishes identically but each may vanish on some variety of positive codimension and $\nabla \varphi_1, \nabla \varphi_2$ are linearly independent at generic points but not necessarily at every point, can be reduced to the case of the preceding paragraph by the arguments in §12.

References

[1] M. Christ, On trilinear oscillatory integral inequalities and related topics, preprint. [arXiv:2007.12753]
[2] ———, On implicitly oscillatory quadrilinear integrals, preprint
[3] M. Christ, P. Durcik, and J. Roos, Trilinear smoothing inequalities and a variant of the triangular Hilbert transform, Adv. Math. 390 (2021), Paper No. 107863
[4] M. Christ and Z. Zhou, A class of singular bilinear maximal functions, preprint. [arXiv:2203.16725]
[5] A. Nagel, E. M. Stein, and S. Wainger, *Balls and metrics defined by vector fields. I. Basic properties*, Acta Math. 155 (1985), no. 1-2, 103-147

[6] E. M. Stein and B. Street, *Multi-parameter singular Radon transforms III: Real analytic surfaces*, Adv. Math. 229 (2012), no. 4, 2210–2238

Michael Christ, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA

*Email address*: mchrist@berkeley.edu