Whitney's Extension Problems and Interpolation of Data

by

*Charles Fefferman**

Department of Mathematics
Princeton University
Fine Hall
Washington Road
Princeton, New Jersey 08544

Email: cf@math.princeton.edu

Abstract

Given a function $f : E \to \mathbb{R}$ with $E \subset \mathbb{R}^n$, we explain how to decide whether f extends to a C^m function F on $\mathbb{R}^n$. If E is finite, then one can efficiently compute an F as above, whose C^m norm has the least possible order of magnitude (joint work with B. Klartag).

Keywords: Whitney extension problem, interpolation

*Supported by grants DMS-0601025 & ONR-N00014-08-1-0678.

Let $f : E \to \mathbb{R}$ be a function defined on a given (arbitrary) set $E \subset \mathbb{R}^n$, and let $m \geq 1$ be a given integer. How can we decide whether f extends to a function $F \in C^m(\mathbb{R}^n)$? If such an F exists, then how small can we take its C^m norm? What can we say about the derivatives of F up to order m at a given point? Can we take F to depend linearly on f?

These questions go back to work of H. Whitney [33,34,35] in 1934. In the decades since Whitney's seminal work, fundamental progress was made by G. Glaeser, [23] Y. Brudnyi and P. Shvartsman [4...9 and 28...30], and E. Bierstone-P. Milman-W. Pawłucki [1]. (See also N. Zobin [36,37] for the solution of a closely related problem.) Building on this work, our recent papers [11...16] gave a complete solution to the above problems. Along the way, we solved the analogous problems with $C^m(\mathbb{R}^n)$ replaced by $C^{m,\omega}(\mathbb{R}^n)$, the space of functions whose m^{th} derivatives have a given modulus of continuity ω. (See [14,15].) It is natural also to consider Sobolev spaces $W^{m,p}(\mathbb{R}^n)$, for which work on the above problems is just beginning (see Shvartsman [31]).

The finite, effective versions of the above problems are basic questions about interpolation of data: Let $E \subset \mathbb{R}^n$ be a finite set, and let $f : E \to \mathbb{R}$ be given. Fix $m \geq 1$. We want to find an interpolant, i.e., a function $F \in C^m(\mathbb{R}^n)$ such that $F = f$ on E. How small can we take the C^m norm of an interpolant? How can we compute an interpolant whose C^m norm is close to least possible? What if we require only that F and f agree approximately on E? What if we are allowed to discard a few points from E? An efficient solution to these problems would likely have practical applications. Joint work [19,20] of B. Klartag and the author shows how to compute efficiently an interpolant F whose C^m norm is within a factor C of least possible, where C is a constant depending only on m and n. Unfortunately, we don't know whether that constant C is absurdly large; we suspect that it is. To remedy this defect, and hopefully obtain results with practical applications, we pose the following sharper version of the interpolation problem.

Let $m, n \geq 1$, and let $f : E \to \mathbb{R}$, where $E \subset \mathbb{R}^n$ consists of N points. Let $\epsilon > 0$ be given. Compute an interpolant F whose C^m norm is within a factor $(1 + \epsilon)$ of least possible.

Work on this difficult problem is just beginning. (See [17,18].)

The goal of this expository paper is to state our main results on the classical Whitney problems for $C^m(\mathbb{R}^n)$, and our joint results with Klartag on the problem of interpolation. We omit here our results on $C^{m,\omega}(\mathbb{R}^n)$, and e.g. the recent work of Shvartsman on $W^{m,p}(\mathbb{R}^n)$. Also, we make no attempt to explain here the ideas in our proofs and algorithms, and merely content ourselves with stating theorems. We hope to present those ideas in a more detailed exposition, to be written someday.

This article is an expanded version of a talk given at a conference celebrating the 25^{th} anniversary of the founding of MSRI. It was a pleasure and an honor to participate in that conference. I am grateful to Anna Tsao for bringing to my attention the practical problem of fitting a smooth surface to data. I am grateful also to Gerree Pecht, for expertly LaTeXing this article.

Let us state our problems in more detail.

Fix integers $m, n \geq 1$. We will work in $C^m(\mathbb{R}^n)$, the space of C^m functions $F : \mathbb{R}^n \to \mathbb{R}$ for which the norm

$$\| F \|_{C^m(\mathbb{R}^n)} = \max_{|\alpha| \leq m} \sup_{x \in \mathbb{R}^n} |\partial^\alpha F(x)|$$

is finite. For $F \in C^m(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, we write $J_x(F)$ (the "jet" of F at x) to denote the m^{th} degree Taylor polynomial of F at x, i.e.,

$$[J_x(F)](y) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!} [\partial^\alpha F(x)] \cdot (y - x)^\alpha \quad \text{for } y \in \mathbb{R}^n .$$

Thus, $J_x(F)$ belongs to $\mathcal{P}$, the vector space of all real-valued m^{th} degree polynomials on $\mathbb{R}^n$. We answer the following four questions.

Suppose we are given a compact[†] set $E \subset \mathbb{R}^n$ and a function $f : E \to \mathbb{R}$.

<u>Question 1:</u> How can we decide whether there exists a function $F \in C^m(\mathbb{R}^n)$ such that $F = f$ on E?

<u>Question 2:</u> Let $x \in \mathbb{R}^n$ be given. Compute

[†] In the introduction, we took E to be arbitrary. Without loss of generality, we can take E to be compact.

$$ATP(x) = \{J_x(F) : F \in C^m(\mathbb{R}^n), \ F = f \text{ on } E\}.$$

(Here "ATP" stands for "Allowed Taylor Polynomials".)

Note that $ATP(x)$ is a (possibly empty) affine subspace of $\mathcal{P}$.

<u>QUESTION 3:</u> Compute the order of magnitude of

$$\| f \|_{C^m(E)} = \inf\{\| F \|_{C^m(\mathbb{R}^n)} : F \in C^m(\mathbb{R}^n), F = f \text{ on } E\}.$$

We owe the reader a precise definition of the phrase "order of magnitude". Suppose $X, Y \geq 0$ are real numbers determined by m, n and other data (e.g. the set E and the function f). Then we say that X and Y have "the same order of magnitude", and we write $X \sim Y$, provided $cX \leq Y \leq CX$ for constants c and C depending only on m and n. To "compute the order of magnitude" of X is to compute some Y such that $X \sim Y$.

Define a Bauach space

$C^m(E) = \left\{F\big|_E : F \in C^m(\mathbb{R}^n)\right\}$, equipped with the norm $\| f \|_{C^m(E)}$ from Question 3.

<u>QUESTION 4:</u> Is there a bounded linear map

$$T : C^m(E) \to C^m(\mathbb{R}^n)$$

such that

$$Tf\big|_E = f \text{ for all } f \in C^m(E)?$$

There are analogues of the above questions with $C^m(\mathbb{R}^n)$ replaced by other function spaces, e.g. Sobolev spaces $W^{m,p}(\mathbb{R}^n)$, or the space $C^{m,\omega}(\mathbb{R}^n)$ of all C^m functions whose m^{th} derivatives have modulus of continuity ω.

For $C^{m,\omega}(\mathbb{R}^n)$, these problems are completely solved, although we do not discuss them here; see [13,14,15]. In the setting of $W^{m,p}(\mathbb{R}^n)$, work is just beginning; see Shvartsman [31].

Next we pose our effective, finite problems.

Suppose we are given:

- $E \subset \mathbb{R}^n$ finite;
- $f : E \to \mathbb{R}$; and
- $\sigma : E \to [0, \infty)$.

<u>QUESTION 5:</u> Decide whether there exists $F \in C^m(\mathbb{R}^n)$ such that

$\| F \|_{C^m(\mathbb{R}^n)} \lesssim 1$

and

$|F(x) - f(x)| \lesssim \sigma(x)$ for all $x \in E$.

Here and below, we write $X \lesssim Y$ to indicate that $X \leq CY$, for a constant C depending only on m and n. To appreciate Question 5, think of an experimenter measuring $F(x)$ for $x \in E$, and obtaining the experimental result $f(x)$ with uncertainty $\sigma(x)$. See Figure 1 for two examples.

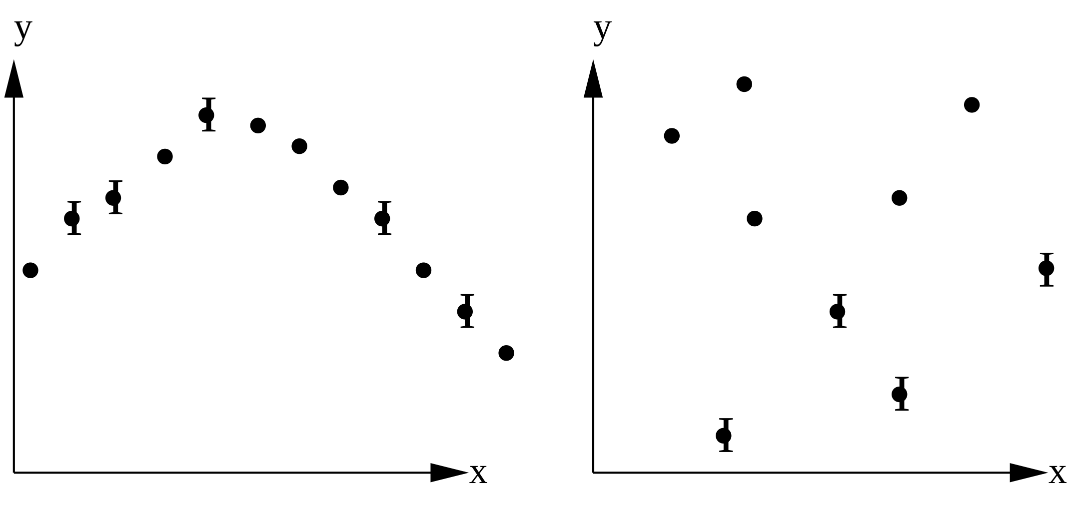

Figure 1(a)

Figure 1(b)

Figure 1: Graphs of two different sets of experimental data.

In the following sharper version of Question 5, we are asked to compute a more-or-less optimal interpolant $F \in C^m(\mathbb{R}^n)$.

<u>QUESTION 6:</u> Compute a function $F \in C^m(\mathbb{R}^n)$ and a real number $M \geq 0$, such that

$$\| F \|_{C^m(\mathbb{R}^n)} \leq M$$

and

$$|F(x) - f(x)| \leq M\sigma(x) \text{ for all } x \in E,$$

with M having the least possible order of magnitude.

We will explain later what it means to "compute a function F". We write $\| f \|_{C^m(E,\sigma)}$ to denote the infimum of M over all pairs F, M as in Question 6.

Perhaps some of the data gathered by our experimenter are "outliers", and should be discarded. Which data points should we discard?

<u>QUESTION 7:</u> Suppose we are given $E \subset \mathbb{R}^n$ (finite), $f : E \to \mathbb{R}$, $\sigma : E \to [0, \infty)$, and an integer Z, less than the number of points in E. How can we find a subset $S \subset E$ with at most Z points, in such a way that the norm $\| f \|_{C^m(E \setminus S, \sigma)}$ is as small as possible?

We regard Questions 5,6,7 as computer-science problems. We seek algorithms, and we ask how many computer operations are needed to carry out those algorithms.

Let us briefly review some of the history of Questions 1 through 4. The subject starts with the work of H. Whitney [33,34,35] in 1934. Whitney answered Questions 1 through 4 for $C^m(\mathbb{R}^1)$ (i.e., for the case $n = 1$), and he proved the classical Whitney Extension Theorem, which gives a simple, complete answer to the following question.

<u>QUESTION 0:</u> Let $E \subset \mathbb{R}^n$ be compact. Suppose that for each $x \in E$ we are given an m^{th} degree polynomial $P^x \in \mathcal{P}$. How can we decide whether there exists a function $F \in C^m(\mathbb{R}^n)$ such that $J_x(F) = P^x$ for all $x \in E$?

See also [25] and [32] for the classical Whitney theorem.

In 1958, G. Glaeser [23] answered Questions 1 and 2 for $C^1(\mathbb{R}^n)$, i.e., for the case $m = 1$.

He gave a geometrical solution based on his notion of the "interated paratangent space". This notion is defined in terms of iterated limits. Examples given by Glaeser [23] and Klartag-Zobin [24] suggest strongly that any complete answer to Question 1 must bring in such iterated limits.

In a series of papers [4...9 and 28...30] from the 1970's to the present decade, Y. Brudnyi and P. Shvartsman studied the analogues of Questions 1 through 4 for the space $C^{m,\omega}(\mathbb{R}^n)$ (the space of functions whose m^{th} derivatives have modulus of continuity ω).

They conjectured the following crucial

FINITENESS PRINCIPLE:

To decide whether a given $f : E \to \mathbb{R}$ extends to a function $F \in C^{m,\omega}(\mathbb{R}^n)$, it is enough to look at all restrictions $f\big|_S$, where $S \subset E$ is an arbitrary k-element subset. Here, k is an integer constant depending only on m and n.

More precisely, if $f\big|_S$ extends to a function $F^S \in C^{m,\omega}(\mathbb{R}^n)$ of norm at most 1 (for each $S \subset E$ with at most k elements), then f extends to a function $F \in C^{m,\omega}(\mathbb{R}^n)$, whose norm is bounded by a constant depending only on m and n.

Brudnyi and Shvartsman proved the above Finiteness Principle in the first non-trivial case, $m = 1$. Their proof introduces the elegant idea of "Lipschitz selection", and produces for $m = 1$ the optimal k, namely $k = 3 \cdot 2^{n-1}$. See [7]. The papers of Brudnyi and Shvartsman contain additional results and conjectures relevant to Questions 1 through 4.

The next significant progress on Questions 1 and 2 was the paper [1] of E. Bierstone, P. Milman and W. Pawłucki. They found an analogue of Glaeser's iterated paratangent space for $C^m(\mathbb{R}^n)$. Using this notion, they answered a variant of Questions 1 and 2 in the case where E is a subanalytic set. Their version of the iterated paratangent space is very close to the dual of the "Glaeser-stable bundle" defined later in this exposition. See [1,2,12].

Question 4 was answered affirmatively in the case $m = 1$ by Bromberg [3], and in the case $n = 1$ by Merrien [26]. See also [8].

Let us now prepare to answer Questions 1 through 4.

It is convenient to view these questions from the broader context of "bundles". With apologies to geometers, we start by defining "bundles" as follows.

Let $E \subset \mathbb{R}^n$ be a compact set. For each $x \in E$, suppose we specify a (possibly empty) affine subspace $H_x \subset \mathcal{P}$.

Then we call the family $\mathcal{H} = (H_x)_{x \in E}$ a "bundle" over E.

We refer to H_x as the "fiber" of $\mathcal{H}$ at x. If $\mathcal{H} = (H_x)_{x \in E}$ and $\mathcal{H}'_x = (H'_x)_{x \in E}$ are bundles over the same compact set E, then we say that $\mathcal{H}'$ is a "sub-bundle" of $\mathcal{H}$ provided we have $H'_x \subseteq H_x$ for all $x \in E$. Note that H_x need not vary continuously with x; in fact, $\dim H_x$ may vary wildly.

The crucial question regarding "bundles" is as follows.

<u>Question 1′</u>: Let $\mathcal{H} = (H_x)_{x \in E}$ be a bundle over a compact set. How can we decide whether there exists a function $F \in C^m(\mathbb{R}^n)$ such that $J_x(F) \in H_x$ for all $x \in E$?

We call such a function F a "section" of the bundle $\mathcal{H}$, and we write $F \in \Gamma(\mathcal{H})$.

Note that if any of the fibers H_x is empty, then obviously the bundle $\mathcal{H}$ has no sections.

It is easy to see that Question 1′ generalizes Question 1. In fact, given $f : E \to \mathbb{R}$, we define a bundle $\mathcal{H} = (H_x)_{x \in E}$ by setting

$H_x = \{P \in \mathcal{P} : P(x) = f(x)\}$ for all $x \in E$.

A section of the bundle $\mathcal{H}$ is precisely a C^m function F such that $F = f$ on E. Consequently, Question 1 is a particular case of Question 1′. The corresponding generalizations of Questions 2 and 3 are as follows.

<u>Question 2′</u>: Let $\mathcal{H} = (H_x)_{x \in E}$ be a bundle over a compact set E. For $x \in \mathbb{R}^n$, compute the affine space $ATP(x, \mathcal{H}) = \{J_x(F) : F \in \Gamma(\mathcal{H})\} \subseteq \mathcal{P}$.

<u>Question 3′</u>: Let $\mathcal{H} = (H_x)_{x \in E}$ be a bundle over a compact set E. Compute the order of magnitude of the infimum of $\| F \|_{C^m(\mathbb{R}^n)}$ over all sections $F \in \Gamma(\mathcal{H})$.

We put off the generalization of Question 4 until later. Note that $ATP(x, \mathcal{H})$ may be empty.

Our initial plan is to answer Questions 1,2,3 by answering the more general Questions 1′, 2′, 3′. Unfortunately, Questions 1′, 2′, 3′ as stated are absurdly general. For instance, any variable-coefficient partial differential equation with boundary conditions on a bounded domain Ω may be written as

$$(*)\quad LF = g \text{ on } \Omega\,,\ \tilde{L}F = h \text{ on } \partial\Omega\,.$$

Here, F is the unknown function; g and h are given; and L, $\tilde{L}$ are variable-coefficient linear partial differential operators.

Taking $E=$ closure (Ω), we note that the condition $LF(x) = g(x)$ for a particular $x \in \Omega$, or $\tilde{L}F(x) = h(x)$ for a particular $x \in \partial\Omega$, asserts simply that the jet $J_x(F)$ belongs to a particular affine subspace $H_x \subseteq \mathcal{P}$. Thus, the boundary-value problem $(*)$ has a C^m solution F if and only if the bundle $(H_x)_{x\in E}$ has a section.

Consequently, Question 1′ contains as a special case every linear boundary-value problem (in which the unknown is a single real-valued function). Particular cases of $(*)$ are already highly non-trivial open problems. We are unlikely to get a complete answer to Question 1′ anytime soon.

The cure is to change our definition of a "bundle". To prepare the way, let $x \in \mathbb{R}^n$ be given. Then the vector space $\mathcal{P}$ becomes a ring under a natural multiplication $\odot_x$ ("jet multiplication"), uniquely specified by demanding that $J_x(FG) = J_x(F) \odot_x J_x(G)$ for any smooth functions F, G. More explicitly, if

$$P(y) = \sum_{|\alpha|\leq m} A_\alpha \cdot (y-x)^\alpha \text{ and } Q(y) = \sum_{|\beta|\leq m} B_\beta \cdot (y-x)^\beta\,,$$

then $P \odot_x Q$ is the polynomial

$$(P \odot_x Q)(y) = \sum_{|\alpha|+|\beta|\leq m} A_\alpha \cdot B_\beta \cdot (y-x)^{\alpha+\beta}\,.$$

That is, we simply multiply P by Q in the usual way, and then discard powers of $(y-x)$ of order greater than m. Now we can give our revised definition of a "bundle".

Let $E \subset \mathbb{R}^n$ be a compact set. For each $x \in E$, let H_x be either the empty set, or else a coset of an ideal in the ring $(\mathcal{P}, \odot_x)$. Then we call $\mathcal{H} = (H_x)_{x \in E}$ a "bundle". Thus, rather than a translate of a vector subspace of $\mathcal{P}$, we demand that a non-empty fiber must be a translate of an ideal.

From now on, we adopt the above revised definition of "bundles". It is now reasonable to pose Questions $1'$, $2'$, $3'$. As before, these questions generalize Questions 1,2,3. In the next few pages, we answer our (appropriately general) Questions $1'$, $2'$, $3'$.

A key idea regarding Questions $1'$, $2'$, $3'$ is that of "Glaeser refinement". Given a bundle $\mathcal{H} = (H_x)_{x \in E}$, we will produce another bundle $\tilde{\mathcal{H}} = (\tilde{H}_x)_{x \in E}$, called the "Glaeser refinement" of $\mathcal{H}$, with three crucial properties:

(A) $\tilde{\mathcal{H}}$ is a sub-bundle of $\mathcal{H}$.

(B) Any section of $\mathcal{H}$ is already a section of $\tilde{\mathcal{H}}$.

(C) $\tilde{\mathcal{H}}$ can be computed from $\mathcal{H}$ by doing linear algebra and taking a limit.

Before giving the definition of the Glaeser refinement, we explain how to exploit properties (A), (B), (C). The idea is to iterate the Glaeser refinement: Suppose we want to understand the sections of a given bundle $\mathcal{H}_0$. According to properties (A) and (B), we may without loss of generality replace $\mathcal{H}_0$ by its Glaeser refinement $\mathcal{H}_1$. Again applying properties (A) and (B), we may similarly replace $\mathcal{H}_1$ by its own Glaeser refinement $\mathcal{H}_2$. Continuing by induction, we define each $\mathcal{H}_k$ to be the Glaeser refinement of $\mathcal{H}_{k-1}$; and we know from repeated applications of (A) and (B) that

($\dagger$1) $\mathcal{H}_0 \supseteq \mathcal{H}_1 \supseteq \mathcal{H}_2 \supseteq \dots$

while

($\dagger$2) $\Gamma(\mathcal{H}_0) = \Gamma(\mathcal{H}_1) = \Gamma(\mathcal{H}_2) = \dots .$

Without loss of generality, we may replace our original bundle $\mathcal{H}_0$ by its k^{th} iterated Glaeser refinement $\mathcal{H}_k$. Each $\mathcal{H}_k$ arises from $\mathcal{H}_{k-1}$ as in property (C) above. Thus, we pass from our given bundle $\mathcal{H}_0$ to the sub-bundle $\mathcal{H}_k$ by taking a k-fold iterated limit.

The above construction, taken from [12], is a slight variant of a construction given in Bierstone-Milman-Pawłucki [1]. A simple ingenious Lemma of Glaeser [23], adapted by Bierstone-Milman-Pawłucki [1], shows that the above process stabilizes. In fact, let $k_* = 2 \cdot \dim \mathcal{P} + 1$. Then

$\mathcal{H}_{k_*} = \mathcal{H}_{k_*+1} = \mathcal{H}_{k_*+2} = \dots$.

That is, the bundle $\mathcal{H}_{k_*}$ is "Glaeser stable", i.e., it is its own Glaeser refinement. Since $\mathcal{H}_0$ and $\mathcal{H}_{k_*}$ have the same sections, and since $\mathcal{H}_{k_*}$ is in principle computable from $\mathcal{H}_0$ thanks to (C), we arrive at the following conclusion.

In order to answer Questions 1′, 2′, 3′, we may assume without loss of generality that the bundle $\mathcal{H}$ is Glaeser stable.

Before proceeding further, we give the definition of the Glaeser refinement, and check its essential properties (A), (B), (C). This is the most technical part of our (otherwise painless) exposition.

We begin by preparing the way with a few simple remarks.

(#1) Let us fix a large enough integer constant k depending only on m and n.

(#2) Given points $x_0, x_1, \dots, x_k \in \mathbb{R}^n$, and given polynomials $P_0, P_1, \dots, P_k \in \mathcal{P}$, we define

$$Q(P_0, P_1, \dots, P_k; x_0, x_1, \dots, x_k) = \sum_{\substack{i,j=0 \\ (x_i \neq x_j)}}^{k} \sum_{|\alpha| \leq m} \left[\frac{\partial^\alpha (P_i - P_j)(x_j)}{|x_i - x_j|^{m-|\alpha|}} \right]^2 .$$

(#3) Suppose $F \in C^m(\mathbb{R}^n)$.

Given $x_0, x_1, \dots, x_k \in \mathbb{R}^n$, define $P_i = J_{x_i}(F)$ for $i = 0, 1, \dots, k$.

Then

$\lim_{x_1, \dots, x_k \to x_0} Q(P_0, P_1, \dots, P_k; x_0, x_1, \dots, x_k) = 0$ (by Taylor's theorem).

Armed with the above useful remark, we can (finally) give the definition of the Glaeser refinement. Let $\mathcal{H} = (H_x)_{x\in E}$ be a bundle over a compact set. The Glaeser refinement $\tilde{\mathcal{H}} = (\tilde{H}_x)_{x\in E}$ is defined as follows.

Fix $x_0 \in E$ and $P_0 \in H_{x_0}$. Then $P_0 \in \tilde{H}_{x_0}$ if and only if

$$(*)\quad \min\{Q(P_0, P_1, \ldots, P_k; x_0, x_1, \ldots, x_k) : P_1 \in H_{x_1}, P_2 \in H_{x_2}, \ldots, P_k \in H_{x_k}\} \to 0$$

as $x_1, x_2, \ldots, x_k \in E$ tend to x_0.

Note that $\tilde{H}_x$ may be empty, even if all the H_x are non-empty. That is why it is convenient to allow the case of empty fibers in the definition of "bundle".

Again, our definition of the Glaeser refinement, taken from [12], is a slight variant of a definition given in [1]. Let us check that properties (A), (B), (C) hold for the Glaeser refinement as we have just defined it.

First of all, (A) holds trivially, since we have defined $\tilde{\mathcal{H}}$ as a sub-bundle of $\mathcal{H}$. Next, (B) follows easily from observation (#3). In fact, let F be section of $\mathcal{H}$, and let $P_0 = J_{x_0}(F)$. Then by taking $P_i = J_{x_i}(F)$ for $i = 1, 2, \ldots, k$, we see from (#3) that $(*)$ holds, and thus $P_0 \in \tilde{H}_{x_0}$. Consequently, F is also a section of $\tilde{\mathcal{H}}$, proving (B). Finally, to prove (C), fix $x_0 \in E$ and $P_0 \in H_{x_0}$. For fixed $x_1, \ldots, x_k \in E$, we compute

$$\wedge(x_0, P_0; x_1 \ldots x_k) = \min\{Q(P_0, \ldots, P_k; x_0, \ldots, x_k) : P_1 \in H_{x_1}, \ldots, P_k \in H_{x_k}\}$$

by minimizing a quadratic form over a finite-dimensional affine space; this is routine linear algebra. By definition, $P_0 \in \tilde{H}_{x_0}$ if and only if $\wedge(x_0, P_0; x_1, \ldots, x_k) \to 0$ as $x_1, \ldots, x_k \in E$ tend to x_0. Thus, to decide whether $P_0 \in \tilde{H}_{x_0}$, we carry out routine linear algebra and then take a limit. This proves (C), concluding our discussion of the definition of the Glaeser refinement.

We are now at the heart of the matter, namely Questions 1′, 2′, 3′ for a Glaeser-stable bundle. The answers are given by the following result.

Theorem 1. *Let $\mathcal{H} = (H_x)_{x\in E}$ be a Glaeser-stable bundle over a compact set E. Then*

(A) *$\mathcal{H}$ has a section if and only if all the fibers H_x $(x \in E)$ are non-empty.*

Moreover, suppose $\mathcal{H}$ has a section. Then the following hold.

(B) *For each* $x \in E$, *we have* $ATP(x, \mathcal{H}) = H_x$. *(Trivially, for each* $x \in \mathbb{R}^n \setminus E$, *we have* $ATP(x, \mathcal{H}) = \mathcal{P}$.*)*

(C) *Let* k *be the integer constant from (#1). Then there exists* $M \in (0, \infty)$ *such that, for any* $x_1, \dots, x_k \in E$ *there exist polynomials* $P_1 \in H_{x_1}, \dots, P_k \in H_{x_k}$ *such that*

$$|\partial^\alpha P_i(x_0)| \le M \text{ for } |\alpha| \le m \text{ and } i = 1, \dots, k; \text{ and}$$

$$|\partial^\alpha (P_i - P_j)(x_j)| \le M |x_i - x_j|^{m-|\alpha|} \text{ for } |\alpha| \le m-1 \text{ and } i, j = 1, \dots, k.$$

(D) *The infimum of* $\| F \|_{C^m(\mathbb{R}^n)}$ *over all sections* $F \in \Gamma(\mathcal{H})$ *has the same order of magnitude as the least possible* M *in (C).*

The proof of this theorem is given in [12]; it builds on the earlier papers [11,14].

We note in passing that every bundle over a finite set E is Glaeser stable; in this case, conclusions (A), (B), (C) above are obvious, but (D) has non-trivial content.

This concludes our discussion of Questions 1,2,3 and 1′, 2′, 3′.

Regarding Question 4, we have the following result.

Theorem 2. *Let* $E \subset \mathbb{R}^n$ *be any subset. Then there exists a linear map* $T : C^m(E) \to C^m(\mathbb{R}^n)$, *such that*

- $Tf\big|_E = f$ *for all* $f \in C^m(E)$*; and*
- *The norm of* T *is bounded by a constant depending only on* m *and* n.

Our generalization of Question 4 in the context of bundles is as follows.

Let $E \subset \mathbb{R}^n$, and for each $x \in E$, let $I(x) \subset \mathcal{P}$ be an ideal in the ring $(\mathcal{P}, \odot_x)$. We write $\mathcal{J}$ to denote the ideal in $C^m(\mathbb{R}^n)$, given by

$$\mathcal{J} = \{F \in C^m(\mathbb{R}^n) : J_x(F) \in I(x) \text{ for all } x \in E\} .$$

Thus, $C^m(\mathbb{R}^n)/\mathcal{J}$ is a Banach space. We write $\pi : C^m(\mathbb{R}^n) \to C^m(\mathbb{R}^n)/\mathcal{J}$ to denote the natural projection.

Theorem 3. *Given any* E *and* $(I(x))_{x\in E}$ *as above, there exists a linear map* $T : C^m(\mathbb{R}^n)/\mathcal{J} \to C^m(\mathbb{R}^n)$*, such that*

- $\pi T =$ *identity on* $C^m(\mathbb{R}^n)/\mathcal{J}$*; and*
- $\| T \| \leq C$*, with* C *depending only on* m *and* n*.*

The last two results are proven in [16].

We recover Theorem 2 from Theorem 3 by taking $I(x) = \{P \in \mathcal{P} : P(x) = 0\}$ for each $x \in E$.

Thus, we have answered the original Questions 1,...,4 arising from Whitney's work.

We turn now to the effective finite version of the problem, and prepare to address Questions 5,6 and 7.

To do so, we must first explain what it means to "compute a function" $F \in C^m(\mathbb{R}^n)$ from data m, n, E, f, σ. We have in mind the following dialogue with a computer.

We first enter the data m, n, E, f, σ. The computer performs "one-time work", and then signals to us that it is ready to accept "queries". We may then repeatedly "query" our computer. A "query" consists of a point $\underline{x} \in \mathbb{R}^n$. Once we enter the query $\underline{x}$, the computer responds by performing a calculation (the "work at query time"), and then printing out $\partial^\alpha F(\underline{x})$ for each $|\alpha| \leq m$.

In the above definition, we demand that the function F be determined uniquely by the input data m, n, E, f, σ. We disallow "adaptive algorithms" in which the function F depends on our queries.

The resources used to "compute a function" $F \in C^m(\mathbb{R}^n)$ are:

- The number of computer operations needed for the one-time work;
- The number of computer operations needed to answer a query; and
- The number of memory cells needed for the whole computation.

We call these simply the "one-time work", the "query work", and the "storage", respectively.

The "computer" we have in mind uses standard von Neumann architecture (see [27]), but it works with exact real numbers. We assume that an arbitrary real number can be stored in a single memory cell, and that elementary arithmetic operations on real numbers may be performed without round-off errors. This model of computation is called "real RAM". (In [20], we also give a rigorous discussion of the effects of round-off errors, but we now omit all further mention of this issue.)

We are looking for algorithms that work well for arbitrary inputs m, n, E, f, σ. If we were willing to make favorable assumptions on the geometry of the set E (e.g. if E looks something like a lattice), then our problems would be much easier. An example of a delicate case arises e.g. if every point of the set E lies near the zero set of a polynomial Q of degree less than m. For such E, the restriction of $F \in C^m(\mathbb{R}^n)$ to E changes little when we change F by adding a multiple of Q. Consequently, a naïve algorithm will carry out an ill-conditioned calculation and produce an awful result.

Now we are ready to answer Questions 5 and 6, by stating the following results, due to B. Klartag and the author [19,20]. We write $\#(E)$ to denote the number of elements of a finite set E.

Theorem 4. *Given* m, n, E, f, σ *(with* $\#(E) = N$*), we can compute the order of magnitude of* $\| f \|_{C^m(E,\sigma)}$ *using at most* $CN \log N$ *operations, and at most* CN *storage. Here,* C *depends only on* m *and* n*.*

Theorem 5. *Given* m, n, E, f, σ *(with* $\#(E) = N$*), we can compute a function* $F \in C^m(\mathbb{R}^n)$*, such that*

- $\| F \|_{C^m(\mathbb{R}^n)} \leq M$
- $|F(x) - f(x)| \leq M\sigma(x)$ *for all* $x \in E$
- $M \leq C \| f \|_{C^m(E,\sigma)}$.

The computation uses one-time work at most $CN \log N$*, query work at most* $C \log N$*, and storage at most* CN*. Here,* C *depends only on* m *and* n*.*

Very likely these results are best possible. Already it takes CN computer operations merely to read the data.

Regarding Question 7, we have the following result. See [22].

Theorem 6. *Given* m, n, E, f, σ, *with* $\#(E) = N$, *there is an enumeration* $x_1, x_2, \dots, x_N$ *of the set* E, *with the following properties:*

(A) *Let* $1 \leq Z \leq N$, *and let* $S \subset E$ *be any subset with at most* cZ *elements. (Here,* c *is a small constant depending only on* m *and* n.*) Let* $S_Z^* := \{x_1, \dots, x_Z\}$, *the first* Z *points of our enumeration. Then* $\| f \|_{C^m(E \setminus S_Z^*, \sigma)} \leq C \| f \|_{C^m(E \setminus S, \sigma)}$, *with* C *depending only on* m *and* n.

(B) *For any* $Z \leq N$ *we can compute* $x_1, \dots, x_Z$ *using at most* $CZN \log N$ *work and* CN *storage, where* C *depends only on* m *and* n.

Theorem 6(A) tells us that the set S_Z^* of putative "outliers" is "more-or-less-optimal", in the sense that we cannot make $\| f \|_{C^m(E \setminus S, \sigma)}$ much smaller than $\| f \|_{C^m(E \setminus S_Z^*, \sigma)}$ by using far fewer points than the Z points of S_Z^*.

We have no reason to believe that the work $CZN \log N$ in Theorem 6(B) is best possible.

The proof of Theorem 6 rests on the following refined version [22] of the Brudnyi-Shvartsman finiteness principle in the setting of $C^m(\mathbb{R}^n)$.

Theorem 7. *Given* m, n, E *and* σ *(with* $\#(E) = N$*), there exist* $S_1, S_2, \dots, S_L \subset E$ *with the following properties.*

- $\#(S_\ell) \leq C$ *for each* ℓ, *with* C *depending only on* m *and* n.
- $L \leq CN$, *with* C *depending only on* m *and* n.
- *For any* $f : E \to \mathbb{R}$, *we have*

$$\| f \|_{C^m(E,\sigma)} \sim \max_{\ell=1,\dots,L} \| f \|_{C^m(S_\ell,\sigma)} .$$

Moreover, we can compute $S_1, S_2, \ldots, S_L$ *from* m, n, E, σ *using at most* $CN \log N$ *work and at most* CN *storage, where* C *depends only on* m *and* n*.*

In principle, Theorems 4...7 give excellent answers to our effective, finite problems. Unfortunately, some of the constants C in the statements of these results are likely to be absurdly large. This would rule out practical applications.

Therefore, we pose the following sharpened form of our interpolation problem:

<u>Question 8</u>: Fix $m, n \geq 1$. Let $f : E \to \mathbb{R}$, with $E \subset \mathbb{R}^n$ finite. Let $\epsilon > 0$ be given. Compute an interpolant $F \in C^m(\mathbb{R}^n)$, such that $\| F \|_{C^m(\mathbb{R}^n)} \leq (1+\epsilon) \cdot \| \tilde{F} \|_{C^m(\mathbb{R}^n)}$ for any other interpolant $\tilde{F}$.

Recall that an interpolant is simply a function $F \in C^m(\mathbb{R}^n)$ such that $F = f$ on E.

In order to pose Question 8, we have to return to our definition of the norm on $C^m(\mathbb{R}^n)$. Recall that we have defined

$$(*1) \quad \| F \|_{C^m(\mathbb{R}^n)} = \sup_{x \in \mathbb{R}^n} \max_{|\alpha| \leq m} |\partial^\alpha F(x)|.$$

We could equally well have defined, say,

$$(*2) \quad \| F \|_{C^m(\mathbb{R}^n)} = \sup_{x \in \mathbb{R}^n} \left(\sum_{|\alpha| \leq m} \frac{1}{\alpha!} |\partial^\alpha F(x)|^2 \right)^{1/2}$$

or

$$(*3) \quad \| F \|_{C^m(\mathbb{R}^n)} = \sum_{|\alpha| \leq m} \sup_{x \in \mathbb{R}^n} |\partial^\alpha F(x)|.$$

Each of these definitions of the C^m-norm leads to a different version of Question 8, whereas Questions 0...7 are unaffected by our choice of the C^m norm.

Motivated by $(*1)$ and $(*2)$, we assume that, for each $x \in \mathbb{R}^n$, we are given a norm $|\cdot|_x$ on the vector space $\mathcal{P}$ of m^{th} degree polynomials on $\mathbb{R}^n$. We then define

$$(*) \quad \| F \|_{C^m(\mathbb{R}^n)} = \sup_{x \in \mathbb{R}^n} |J_x(F)|_x.$$

This includes the formulas (∗1) and (∗2), as we see by taking

$$|P|_x = \max_{|\alpha| \leq m} |\partial^\alpha P(x)|$$

and

$$|P|_x = \left(\sum_{|\alpha| \leq m} \frac{1}{\alpha!} |\partial^\alpha P(x)|^2 \right)^{1/2} ,$$

respectively.

(Note that the norm (∗3) is not of the form (∗); we will not study Question 8 for the norm (∗3).) We require that the norms $|P|_x$ satisfy two technical assumptions called the "bounded distortion property" and "approximate translation-invariance"; see [17,18].

Once we have specified the norms $|\cdot|_x$ ($x \in \mathbb{R}^n$), our Question 8 makes sense; it refers to the C^m-norm (∗).

From now on, we assume we are given a family of norms $|\cdot|_x$ as above; and our C^m norm is always defined by (∗).

Any algorithm to answer Question 8 must make use of the norms $|\cdot|_x$. We assume that our computer has access to an Oracle. Given a polynomial $P \in \mathcal{P}$ and a point $x \in \mathbb{R}^n$, the Oracle returns the value of $|P|_x$, and charges us one unit of "work" to do so. (This assumption can be weakened; see [18].)

The following problem is related to Question 8, just as Whitney's Extension Theorem is related to the classical Whitney extension problem. (See Questions 0 and 1.)

<u>Question 9:</u> Fix $m, n \geq 1$. Let $E \subset \mathbb{R}^n$ be finite. For each $x \in E$, let $P^x \in \mathcal{P}$ be given. Let $\epsilon > 0$. Compute a function $F \in C^m(\mathbb{R}^n)$ with the following properties.

- $J_x(F) = P^x$ for each $x \in E$.
- Let $\tilde{F} \in C^m(\mathbb{R}^n)$ be any other function satisfying $J_x(\tilde{F}) = P^x$ for each $x \in E$. Then

$$\| F \|_{C^m(\mathbb{R}^n)} \leq (1 + \epsilon) \cdot \| \tilde{F} \|_{C^m(\mathbb{R}^n)} .$$

We have a sharp answer to Question 9.

Theorem 8. *Given* m, n, ϵ *and* $(P^x)_{x \in E}$ *as in Question 9, with* $\#(E) = N$*, we can compute* $F \in C^m(\mathbb{R}^n)$ *satisfying the conditions in Question 9, with one-time work* $C(\epsilon) N \log N$*, query work* $C(\epsilon) \log N$*, and storage* $C(\epsilon)N$*. Here,* $C(\epsilon)$ *depends only on* m, n, ϵ *and on our choice of the* C^m*-norm.* See [18].

Here, we are not primarily concerned with the dependence of $C(\epsilon)$ on ϵ, since in hoped-for practical applications we can take, say, $\epsilon = 3$.

The proof of Theorem 8 is based on a finiteness principle analogous to the Brudnyi-Shvartsman finiteness principle. Unfortunately, the corresponding finiteness principle in the context of Question 8 is false. (See Fefferman-Klartag [21].)

Regarding Question 8, we at least have the following result, as a consequence of the proof of Theorem 8. (See [18].)

Theorem 9. *Given* m, n, E, f, ϵ*, with* $\#(E) = N$*, we can compute an interpolant* F *as in Question 8, using one-time work* $C(\epsilon)N^5(\log N)^2$*, query work* $C(\epsilon) \log N$*, and storage* $C(\epsilon)N^2$*.*

We can probably do a bit better by tweaking the argument. However, to arrive at an efficient algorithm for Question 8, we will probably need to find new mathematics to fill the gap left by the failure of the relevant finiteness principle.

I've just begun to make progress on this difficult problem.

I think I see how to give an optimal algorithm for Question 8 in the case of $C^2(\mathbb{R}^2)$. As in Theorem 8, the one-time work is $C(\epsilon)N \log N$, the query work is $C(\epsilon) \log N$, and the storage needed is $C(\epsilon)N$. I hope this result survives when I try to write it up carefully.

The general case (i.e., Question 8 for $C^m(\mathbb{R}^n)$) will require substantial additional ideas. Let us see what progress can be made in the years ahead.

References

1. E. Bierstone, P. Milman, W. Pawłucki, *Differentiable functions defined on closed sets. A problem of Whitney*, Inventiones Math. **151, No. 2** (2003), 329-352.

2. E. Bierstone, P. Milman, W. Pawłucki, *Higher-order tangents and Fefferman's paper on Whitney's extension problem*, Annals of Math., (to appear).

3. S. Bromberg, *An extension in the class* C^1, Bol. Soc. Mat. Mex. II, **Ser. 27**, (1982), 35-44.

4. Y. Brudnyi, *On an extension theorem*, Funk. Anal. i Prilzhen. **4** (1970), 97-98; English transl. in Func. Anal. Appl. **4** (1970), 252-253.

5. Y. Brudnyi and P. Shvartsman, *The traces of differentiable functions to closed subsets of* $\mathbb{R}^n$, in Function Spaces (Poznán, 1989), Teubner-Texte Math., **120**, Teubner, Stuttgart (1991), 206-210.

6. Y. Brudnyi and P. Shvartsman, *A linear extension operator for a space of smooth functions defined on closed subsets of* R^n, Dokl. Akad. Nauk SSSR **280** (1985), 268-270. English transl. in Soviet Math. Dokl. **31, No. 1** (1985), 48–51.

7. Y. Brudnyi and P. Shvartsman, *Generalizations of Whitney's extension theorem*, Int. Math. Research Notices **3** (1994), 129-139.

8. Y. Brudnyi and P. Shvartsman, *The Whitney problem of existence of a linear extension operator*, J. Geometric Analysis **7, No. 4** (1997), 515-574.

9. Y. Brudnyi and P. Shvartsman, *Whitney's extension problem for multivariate* $C^{1,w}$ *functions*, Trans. Amer. Math. Soc. **353, No. 6** (2001), 2487-2512.

10. C. Fefferman, *Interpolation and extrapolation of smooth functions by linear operators*, Revista Matematica Iberoamericana, **21 No. 1** (2005), 313-348.

11. C. Fefferman, *A sharp form of Whitney's extension theorem*, Annals of Math., **161**, no. 1 (2005), 509-577.

12. C. Fefferman, *Whitney's extension problem for* C^m, Annals of Math., **164**, no. 1 (2006), 313-359.

13. C. Fefferman, *Whitney's extension problem in certain function spaces*, (preprint).

14. C. Fefferman, *A generalized sharp Whitney theorem for jets*, Revista Matematica Iberoamericana, **21, No. 2** (2005), 577-688.

15. C. Fefferman, *Extension of* $C^{m,\omega}$ *smooth functions by linear operators*, Revista Matematica Iberoamericana, (to appear).

16. C. Fefferman, C^m *extension by linear operators*, Annals of Math., Vol. 166, No. 3, (2007).

17. C. Fefferman, *The* C^m *norm of a function with prescribed jets I*, (to appear).

18. C. Fefferman, *The* C^m *norm of a function with prescribed jets II*, (to appear).

19. C. Fefferman and B. Klartag, *Fitting a* C^m*-smooth function to data I*, Annals of Math., (accepted (2006)).

20. C. Fefferman and B. Klartag, *Fitting a* C^m*-smooth function to data II*, Revista Matemática Iberoamericana, (accepted (2007)).

21. C. Fefferman and B. Klartag, *An example related to Whitney extension with almost minimal* C^m *norm*, (to appear).

22. C. Fefferman, *Fitting a* C^m*-smooth function to data III*, Annals of Math., (accepted (2007)).

23. G. Glaeser, *Etudes de quelques algebres tayloriennes*, J d'Analyse **6** (1958), 1-124.

24. B. Klartag and N. Zobin, C^1 *extensions of functions and stabilization of Glaeser refinements*, Revista Math. Iberoamericana, vol. 23 no. 2, (2007), 635-669..

25. B. Malgrange, Ideals of Differentiable Functions, Oxford U. Press, (1966).

26. J. Merrien, *Prolongateurs de foncions differentiables d'une variale reelle*, J. Math. Pures. Appl. (9), **45** (1966), 291-309.

27. J. von Neumann, First draft of a report on the EDVAC, Contract No. W-670-ORD-492. Reprinted in *IEEE Annals of the History of Computing*, vol. 15, Issue 4 (1993), 27-75.

28. P. Shvartsman, *Lipschitz selections of multivalued mappings and traces of the Zygmund class of functions to an arbitrary compact*, Dokl. Acad. Nauk SSSR **276** (1984), 559-562; English transl. in Soviet Math. Dokl. **29** (1984), 565-568.

29. P. Shvartsman, *On traces of functions of Zygmund classes*, Sibirskyi Mathem. J. **28 N5** (1987), 203-215; English transl. in Siberian Math. J. **28** (1987), 853-863.

30. P. Shvartsman, *Lipschitz selections of set-valued functions and Helly's theorem*, J. Geometric Analysis **12, No. 2** (2002), 289-324.

31. P. Shvartsman, *Sobolev* W^1_p*-spaces on closed sets and domains in* $\mathbb{R}^n$, (to appear).

32. E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton U. Press, (1970).

33. H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. **36** (1934), 63-89.

34. H. Whitney, *Differentiable functions defined in closed sets I*, Trans. Amer. Math. Soc. **36** (1934), 369-389.

35. H. Whitney, *Functions differentiable on the boundaries of regions*, Annals of Math. **35** (1934), 482-485.

36. N. Zobin, *Whitney's problem on extendability of functions and an intrinsic metric*, Advances in Math., **133, No. 1**, (1998), 96-132.

37. N. Zobin, *Extension of smooth functions from finitely connected planar domains*, Journal of Geom. Analysis, **9, No. 3**, (1999), 489-509.