PRELIMINARY VERSION

Fitting a C^m-Smooth Function to Data

by

*Charles Fefferman**

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

Email: cf@math.princeton.edu

*This work was partially supported by NSF grant #DMS-0245242

§0. The Problem

Let $m, n \geq 1$, and suppose we are given N points $(x_1, t_1), \ldots, (x_N, t_N) \in \mathbb{R}^n \times \mathbb{R}$. Let $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$ denote the infimum of $\| F \|_{C^m(\mathbb{R}^n)}$ over all functions $F \in C^m(\mathbb{R}^n)$ whose graphs pass through the given points. To avoid trivial cases, we assume that $x_1, \ldots, x_N$ are all distinct. We want to know how many operations are needed to compute the order of magnitude of $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$. By "order of magnitude", we mean the following: Two numbers $X, Y \geq 0$ determined by $(x_1, t_1), \ldots, (x_N, t_N)$ and m, n are said to have "the same order of magnitude", provided we have $cX \leq Y \leq CX$, with 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 and Y have the same order of magnitude.

Our algorithms work with exact real numbers. We ignore roundoff and overflow errors, although our discussion could be modified to take account of these issues. By an "operation", we mean one of the following:

- An addition, subtraction, multiplication, or division of real numbers.
- A "comparison" of two real numbers x and y, i.e., the decision as to whether $x < y$, $x = y$, or $x > y$.
- The act of reading a real number from memory.
- The act of writing a real number into memory.

We regard m and n as fixed, but N as very large. The main result of this paper is as follows.

Theorem 1: *The algorithm to be explained below computes the order of magnitude of* $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$ *in at most* CN^2 *operations, where* C *depends only on* m *and* n.

Bo'az Klartag has greatly sharpened Theorem 1, obtaining $N \log N$ in place of N^2. Bo'az and I will write an extended version of this paper, including this sharpened result.

A significant feature of our algorithm is that it works for arbitrary collections of N points in $\mathbb{R}^n \times \mathbb{R}$. Under simplifying assumptions on the geometry of the points, it is easy to give fast algorithms to compute the order of magnitude of $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$.

A delicate case arises, eg, when the points $x_1, \ldots, x_N \in \mathbb{R}^2$ all lie close to the curve $V = \{Q = 0\} \subset \mathbb{R}^2$, where Q is a low-degree polynomial. We are looking for a function $F \in C^m(\mathbb{R}^2)$, taking prescribed values at $x_1, \ldots, x_N$, and having essentially the least possible C^m norm. To compute such an F, it would be very helpful to determine the $m^{\underline{\text{th}}}$ order Taylor polynomial of F at each x_i modulo a small error. Call these polynomials $P_i (i = 1, \ldots, N)$. Because $x_1, \ldots, x_N$ lie close to V, it may be hard to determine the P_i, except perhaps modulo Q. On the other hand, suppose that the line segment joining two nearby points x_{10} and x_{27} meets V at a not-so-small angle. For x_i near x_{10} and x_{27}, we expect that $\vec{\nu} \cdot \nabla P_i(x_i) \approx (F(x_{10}) - F(x_{27}))/|x_{10} - x_{27}|$, where $\vec{\nu}$ is the unit vector in the direction of $x_{10} - x_{27}$. Consequently, for some of the x_i, the data $F(x_1) = t_1, \ldots, F(x_N) = t_N$ allow us to distinguish between the two hypotheses $P_i = P$ and $P_i = P + Q$. More generally, anything we learn about some P_j may also teach us something about P_i for $i \neq j$. This remark plays a key rôle in our algorithm. We invite the reader to trace what our algorithm does when $x_1, \ldots, x_N$ are as above. From now on, we confine our attention to the general case.

Instead of demanding that the graph of F pass through $(x_1, t_1), \ldots, (x_N, t_N)$, we could have asked merely that the graph of F lie close to the (x_ν, t_ν). Thus, suppose we are given points $(x_1, t_1), \ldots, (x_N, t_N) \in \mathbb{R}^n \times \mathbb{R}$, and "tolerances" $\sigma_1, \sigma_2, \ldots, \sigma_N \in [0, \infty)$. We define $Norm\{(x_\nu, t_\nu, \sigma_\nu)_{\nu=1,\ldots,N}\}$ as the infimum of all $M > 0$ for which there exists $F \in C^m(\mathbb{R}^n)$, such that

$$\| F \|_{C^m(\mathbb{R}^n)} \leq M \,, \text{ and } |F(x_\nu) - t_\nu| \;\leq\; M\sigma_\nu \text{ for } \nu = 1, \ldots, N \,.$$

(This reduces to $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$ in the special case $\sigma_1 = \sigma_2 = \cdots = \sigma_N = 0$.) An obvious variant of our algorithm allows us to compute the order of magnitude of $Norm\{(x_\nu, t_\nu, \sigma_\nu)\}$ in at most CN^2 operations, as in Theorem 1. The main change needed to accommodate this generalization occurs in Section 8 below, where we would have to apply the main results of [11] in full force, rather than the special case called Theorem 2 in this paper. We omit the details.

This paper is part of a literature on the problem of extending a given function $f : E \to \mathbb{R}$, defined on an arbitrary set $E \subset \mathbb{R}^n$, to a function $F \in C^m(\mathbb{R}^n)$. The question goes back to Whitney [25,26,27], with significant contributions by Glaeser [17], Brudnyi-Shvartsman [4,...,9 and 20,21,22], A. and Y. Brudnyi [3], Zobin [28,29], and Bierstone-Milman-Pawłucki [1,2]; see also my papers [10,...,16]. Here, we take $E = \{x_1, \ldots, x_N\}$ finite, and we pose the question from the viewpoint of theoretical computer science.

We see no reason to believe that our algorithm is best possible, and we look forward to future improvements. For the case $m = 1$, an essentially optimal solution is contained in the work of Har-Peled and Mendel [18]. We thank A. Naor for pointing this out. Section 9 below mentions a few open problems.

I am grateful to Gerree Pecht for LaTeXing this paper to ever-impeccable "Gerree standards."

§1. The Plan

The idea behind our algorithm starts with the following elementary remarks. Suppose $F \in C^m(\mathbb{R}^n)$, with $F(x_\nu) = t_\nu$ for $\nu = 1, \ldots, N$. Let $\bar{M} > 0$ be an upper bound for $\| F \|_{C^m(\mathbb{R}^n)}$, and let C denote a constant depending only on m and n. Let P_ν denote the $(m-1)^{\underline{\text{rst}}}$ degree Taylor polynomial of F at the point x_ν. Then, for each $\nu = 1, \ldots, N$, we have

(1) $|\partial^\alpha P_\nu(x_\nu)| \leq \bar{M}$ for $|\alpha| \leq m-1$; and $P_\nu(x_\nu) = t_\nu$.

Moreover, for any $\nu, \mu = 1, \ldots, N$, Taylor's theorem gives

(2) $|\partial^\alpha (P_\nu - P_\mu)(x_\nu)| \leq C\bar{M}|x_\nu - x_\mu|^{m-|\alpha|}$ for $|\alpha| \leq m-1$.

To exploit these remarks, we introduce the vector space $\mathcal{P}$ of all (real-valued) $m-1^{\underline{\text{rst}}}$ degree polynomials on $\mathbb{R}^n$; and we define a family of (possibly empty) convex subsets $\Gamma(x_\nu, \ell, M) \subset \mathcal{P}$ by the following induction on ℓ:

For $\ell = 1$, $1 \leq \nu \leq N$, $M \in (0, \infty)$, we define

(3) $\Gamma(x_\nu, 1, M) = \{P \in \mathcal{P} : |\partial^\alpha P(x_\nu)| \leq M$ for $|\alpha| \leq m-1$, and $P(x_\nu) = t_\nu\}$.

For $\ell \geq 1$, suppose we have already defined the sets $\Gamma(x_\mu, \ell, M)$. Then, for $1 \leq \nu \leq N$ and $M \in (0, \infty)$, we define

(4) $\Gamma(x_\nu, \ell+1, M) = \{P \in \mathcal{P}$: For each μ, there exists $P' \in \Gamma(x_\mu, \ell, M)$, such that $|\partial^\alpha (P - P')(x_\nu)| \leq M|x_\nu - x_\mu|^{m-|\alpha|}$ for $|\alpha| \leq m-1\}$

Then observations (1),(2), and an obvious induction on ℓ, show that $P_\nu \in \Gamma(x_\nu, \ell, C\bar{M})$ for each ν, ℓ. In particular,

(5) Whenever $M \geq C \cdot Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$, we have $\Gamma(x_\nu, \ell, M) \neq \phi$ for all ℓ, ν. (As usual, ϕ denotes the empty set.)

Conversely, for an $\ell_* \geq 1$, depending only on m and n, the following holds:

(6) Let $M > 0$, and suppose that $\Gamma(x_\nu, \ell_*, M) \neq \phi$ for all ν.
Then $Norm\,\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\} \leq C \cdot M$.

We will prove (6) in Section 8 below, by reducing it to a result from [11]. (See also Brudnyi-Shvartsman [7].) From (5) and (6), we see that

(7) $Norm\,\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$ has the same order of magnitude as
$inf\,\{M > 0 :\, \Gamma(x_\nu, \ell_*, M) \neq \phi$ for each $\nu = 1, \ldots, N\}$.

The idea of our algorithm is to compute the approximate size and shape of the convex sets $\Gamma(x_\nu, \ell, M)$ by following the induction (3), (4); and then to read off the order of magnitude of $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$ from (7). In the next few sections, we explain more precisely what this means, and how to carry it out.

§2. The Data Structures

In this section, we define the basic data structures used to specify the "approximate size and shape" of the convex sets $\Gamma(x_\nu, \ell, M)$ from Section 1. We also define some basic operations on those data structures. Let V be a finite-dimensional (real) vector space. A "blob" in V is a family $\mathcal{K} = (K_M)_{M>0}$ of (possibly empty) convex subsets $K_M \subseteq V$, parametrized by $M \in (0, \infty)$, such that $M < M'$ implies $K_M \subseteq K_{M'}$. The "onset" of a blob $\mathcal{K} = (K_M)_{M>0}$ is defined as the infimum of all the $M > 0$ for which $K_M \neq \phi$. (If all K_M are empty, then onset $\mathcal{K} = +\infty$.)

Suppose $\mathcal{K} = (K_M)_{M>0}$ and $\mathcal{K}' = (K'_M)_{M>0}$ are blobs in V, and let $C \geq 1$ be a constant. We say that $\mathcal{K}$ and $\mathcal{K}'$ are "C-equivalent" if they satisfy $K_M \subseteq K'_{CM}$ and $K'_M \subseteq K_{CM}$ for all $M \in (0, \infty)$.

Note that, if $\mathcal{K}$ and $\mathcal{K}'$ are C_1-equivalent, and if $\mathcal{K}'$ and $\mathcal{K}''$ are C_2-equivalent, then $\mathcal{K}$ and $\mathcal{K}''$ are $C_1 \cdot C_2$-equivalent. Note also that, if $\mathcal{K}$ and $\mathcal{K}'$ are C-equivalent, then

$(1/C) \cdot \text{onset } \mathcal{K} \leq \text{onset } \mathcal{K}' \leq C \cdot \text{ onset } \mathcal{K}$.

For fixed $x_\nu, \ell \geq 1$, the family of sets $(\Gamma(x_\nu, \ell, M))_{M>0}$ from the previous section forms a blob in $\mathcal{P}$, which we call $\Gamma(x_\nu, \ell)$. In the language of blobs, the fundamental result (7) from the previous section becomes

(0) $Norm\{(x_\nu, t_\nu)_{\nu=1,\dots,N}\}$ has the same order of magnitude as
$\max\{\text{onset } \Gamma(x_\nu, \ell_*) : \nu = 1, \dots, N\}$.

Among all the blobs in V, we focus attention on those given by "Approximate Linear Algebra Problems", or "ALPs". To define these, let $\lambda_1, \dots, \lambda_L$ be (real) linear functionals on V, let $b_1, \dots, b_L$ be real numbers, let $\sigma_1, \dots, \sigma_L$ be non-negative real numbers, and let $M_* \in [0, +\infty]$. We call

(1) $\mathcal{A} = [(\lambda_1, \dots, \lambda_L), (b_1, \dots, b_L), (\sigma_1, \dots, \sigma_L), M_*]$
an "ALP" in V. With $\mathcal{A}$ given by (1), we define a blob

(2) $\mathcal{K}(\mathcal{A}) = (K_M(\mathcal{A}))_{M>0}$ in V, by setting

(3) $K_M(\mathcal{A}) = \{v \in V : |\lambda_\ell(v) - b_\ell| \leq M \cdot \sigma_\ell \text{ for } \ell = 1, \dots, L\}$ for $M \geq M_*$, and

(4) $K_M(\mathcal{A}) = \phi$ for $M < M_*$.

(Our definition (3) motivates the use of the phrase "approximate linear algebra problem".) We allow $L = 0$ in (1), in which case (3) says simply that $K_M(\mathcal{A}) = V$ for $M \geq M_*$.

An "ALP" is intermediate in generality between a linear algebra problem and a linear programming problem.

We call $\mathcal{K}(\mathcal{A})$ "the blob (in V) arising from the ALP $\mathcal{A}$". Unlike a general blob, an ALP is specified by finitely many (real) parameters, and may therefore be manipulated by algorithms. We call L and M_* in (1), respectively, the "length" and "threshold" of the ALP $\mathcal{A}$.

To "compute the approximate size and shape" of the $\Gamma(x_\nu, \ell, M)$, we will exhibit an ALP $\mathcal{A}_{\nu,\ell}$ in $\mathcal{P}$, such that the blobs $\mathcal{K}(\mathcal{A}_{\nu,\ell})$ and $\Gamma(x_\nu, \ell)$ are C-equivalent, with C depending only on ℓ, m, n.

In particular, onset $\mathcal{K}(\mathcal{A}_{\nu,\ell_*})$ and onset $\Gamma(x_\nu, \ell_*)$ will therefore have the same order of magnitude, so that (0) implies

(5) $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$ has the same order of magnitude as
$\max\{\text{onset } \mathcal{K}(\mathcal{A}_{\nu,\ell_*}) : \nu = 1, \ldots, N\}$.

To exhibit the $\mathcal{A}_{\nu,\ell}$, we will follow the inductive process (3), (4) from Section 1. We will construct the $\mathcal{A}_{\nu,\ell}$ in the next section. Here, we prepare the way by defining a few elementary operations on blobs and ALPs.

First, suppose $V = V_1 \oplus V_2$ is a direct sum of vector spaces, and let $\mathcal{K} = (K_M)_{M>0}$ be a blob in V_2. Then we obtain trivially a blob $V_1 \oplus \mathcal{K} = (V_1 \oplus K_M)_{M>0}$ in $V_1 \oplus V_2$, by setting $V_1 \oplus K_M = \{(v_1, v_2) \in V_1 \oplus V_2 : v_2 \in K_M\}$. In particular, if $\mathcal{K} = \mathcal{K}(\mathcal{A})$ is the blob in V_2 arising from an ALP $\mathcal{A}$, then also $V_1 \oplus \mathcal{K} = \mathcal{K}(\mathcal{A}^+)$ is the blob in $V_1 \oplus V_2$ arising from an ALP $\mathcal{A}^+$; the ALP $\mathcal{A}^+$ is constructed from $\mathcal{A}$ in a trivial manner. The ALP's $\mathcal{A}^+$ and $\mathcal{A}$ have the same length.

Next, suppose $V = V_1 \oplus V_2$, let $\pi : V \to V_1$ be the natural projection, and let $\mathcal{K} = (K_M)_{M>0}$ be a blob in V. Then we define the blob $\pi\mathcal{K}$ in V_1 by setting

(6) $\pi\mathcal{K} = (\pi K_M)_{M>0}$.

If $\mathcal{K} = \mathcal{K}(\mathcal{A})$ is the blob in V arising from an ALP $\mathcal{A}$, then there exists an ALP $\bar{\mathcal{A}}$ in V_1, such that the blob $\pi\mathcal{K}$ is C-equivalent to $\mathcal{K}(\bar{\mathcal{A}})$, with C depending only on $\dim V$. In Section 6 below, we will show how to compute $\bar{\mathcal{A}}$ from $\mathcal{A}$. Our $\bar{\mathcal{A}}$ will have length at most $\dim V_1$. (One can also exhibit an ALP $\bar{\mathcal{A}}'$ in V_1, such that $\mathcal{K}(\bar{\mathcal{A}}') = \pi\mathcal{K}$. However, the length of $\bar{\mathcal{A}}'$ will be more than $\dim V_1$. We make no use of $\bar{\mathcal{A}}'$.)

Now let V be a vector space, and suppose that $\mathcal{K}^\nu = (K^\nu_M)_{M>0}$ is a blob in V, for each $\nu = 1, 2, \ldots, N$. Then we define the intersection $\mathcal{K}^1 \cap \cdots \cap \mathcal{K}^N$ by setting

$$\mathcal{K}^1 \cap \cdots \cap \mathcal{K}^N \;=\; (K_M^1 \cap \cdots \cap K_M^N)_{M>0}\,.$$

If each $\mathcal{K}^\nu = \mathcal{K}(\mathcal{A}_\nu)$ for an ALP $\mathcal{A}_\nu$, then their intersection has the form

$$\mathcal{K}^1 \cap \cdots \cap \mathcal{K}^N = \mathcal{K}(\hat{\mathcal{A}})\,,$$

for an ALP $\hat{\mathcal{A}}$ determined from $\mathcal{A}_1, \cdots, \mathcal{A}_N$ in an obvious way. In particular, the length of $\hat{\mathcal{A}}$ is the sum of the lengths of the $\mathcal{A}_\nu$.

The above operations on blobs behave well with respect to C-equivalence. In fact, if $\mathcal{K}$ and $\mathcal{K}'$ are C-equivalent blobs in V_2, then (trivially) $V_1 \oplus \mathcal{K}$ and $V_1 \oplus \mathcal{K}'$ are C-equivalent blobs in $V_1 \oplus V_2$. Also, if $\pi : V_1 \oplus V_2 \to V_1$ is the natural projection, and if $\mathcal{K}$ and $\mathcal{K}'$ are C-equivalent blobs in $V_1 \oplus V_2$, then (trivially) $\pi\mathcal{K}$ and $\pi\mathcal{K}'$ are C-equivalent blobs in V_1. Finally, if $\mathcal{K}^\nu$ is C-equivalent to $(\mathcal{K}')^\nu$ for $\nu = 1, \ldots, N$, then (trivially) $\mathcal{K}^1 \cap \cdots \cap \mathcal{K}^N$ is C-equivalent to $(\mathcal{K}')^1 \cap \cdots \cap (\mathcal{K}')^N$.

Let $\mathcal{A}$ be any ALP in a vector space V. Then $\mathcal{K}(\mathcal{A})$ is C-equivalent to $\mathcal{K}(\mathcal{A}^\#)$ for another ALP $\mathcal{A}^\#$, with length $(\mathcal{A}^\#) \leq \dim V$, and with C depending only on $\dim V$. In Section 6 below, we prove this fact and show how to compute $\mathcal{A}^\#$ from $\mathcal{A}$. Moreover, the $\mathcal{A}^\#$ constructed in Section 6 has the additional property that onset $\mathcal{K}(\mathcal{A}^\#)$ = threshold $(\mathcal{A}^\#)$. This allows us to replace a long ALP $\mathcal{A}$ by a short one $\mathcal{A}^\#$, and also to compute the order of magnitude of the onset of $\mathcal{K}(\mathcal{A})$.

§3. The ALPs

In this section, we use the algorithms sketched and promised in the previous section, to compute, for each $\nu = 1, \ldots, N$ and $\ell \geq 1$ an ALP $\mathcal{A}_{\nu,\ell}$ in $\mathcal{P}$, with the following properties:

$(1)_\ell$ length $(\mathcal{A}_{\nu,\ell}) \leq \dim \mathcal{P} + 1$; and

$(2)_\ell$ The blobs $\Gamma(x_\nu, \ell)$ and $\mathcal{K}(\mathcal{A}_{\nu,\ell})$ are C-equivalent, with C depending only on ℓ, m, n.

Once we have found the $\mathcal{A}_{\nu,\ell}$, then, in view of $(2)_\ell$, we will have "computed the approximate size and shape" of $\Gamma(x_\nu, \ell, M)$.

To compute the $\mathcal{A}_{\nu,\ell}$ and prove $(1)_\ell$ and $(2)_\ell$, we proceed by induction on ℓ, following the induction (3), (4) in Section 1.

For $\ell = 1$, we simply recall that $\Gamma(x_\nu, 1, M)$ is defined as

$$\Gamma(x_\nu, 1, M) \;=\; \{P \in \mathcal{P} : |\partial^\alpha P(x_\nu)| \;\leq\; M \cdot 1 \text{ for } |\alpha| \leq m-1, \text{ and } |P(x_\nu) - t_\nu| \;\leq\; M \cdot 0\}$$

for all $M \in (0, \infty)$. Thus, $\Gamma(x_\nu, 1) = \mathcal{K}(\mathcal{A}_{\nu,1})$ for an obvious ALP $\mathcal{A}_{\nu,1}$ of length $\dim \mathcal{P} + 1$. In particular, $(1)_\ell$ and $(2)_\ell$ hold for $\ell = 1$.

For the inductive step, we fix $\ell \geq 1$, and we suppose we are given ALPs $\mathcal{A}_{\nu,\ell}$ in $\mathcal{P}$ (for $\nu = 1, \ldots, N$) satisfying $(1)_\ell$ and $(2)_\ell$. We show how to compute ALPs $\mathcal{A}_{\nu,\ell+1}$ $(\nu = 1, \ldots, N)$ satisfying $(1)_{\ell+1}$ and $(2)_{\ell+1}$. We write C to denote constants depending only on ℓ, m, n.

Let us break up the inductive step (4) in Section 1 into little, easy steps. Given the $\Gamma(x_\nu, \ell)$ $(\nu = 1, \ldots, N)$, we proceed as follows.

Step 1: For each $\mu = 1, \cdots, N$, we form the blob $\Gamma^+(x_\mu, \ell) = \mathcal{P} \oplus \Gamma(x_\mu, \ell)$ in $\mathcal{P} \oplus \mathcal{P}$.

Step 2: For each $\nu, \mu = 1, \ldots, N$, we form the blob $\Omega_{\nu,\mu}$ in $\mathcal{P} \oplus \mathcal{P}$, given by $\Omega_{\nu,\mu} = (\Omega_{\nu,\mu,M})_{M>0}$, with

$$\Omega_{\nu,\mu,M} = \{(P, P') \in \mathcal{P} \oplus \mathcal{P} : |\partial^\alpha (P - P')(x_\nu)| \;\leq\; M |x_\nu - x_\mu|^{m-|\alpha|} \text{ for } |\alpha| \leq m-1\}$$

for $M \in (0, \infty)$.

Step 3: For each $\nu, \mu = 1, \ldots, N$, we form the blob

$$\Gamma^\#(x_\nu, x_\mu, \ell) \;:=\; \Gamma^+(x_\mu, \ell) \,\cap\, \Omega_{\nu,\mu} \text{ in } \mathcal{P} \oplus \mathcal{P}\,.$$

Thus, $\Gamma^\#(x_\nu, x_\mu, \ell) \;=\; (K^\#_{\nu,\mu,\ell,M})_{M>0}$, with

$$K^\#_{\nu,\mu,\ell,M} \;=\; \{(P, P') \in \mathcal{P} \oplus \mathcal{P} : P' \in \Gamma(x_\mu, \ell, M)\,, \text{ and}$$

$$|\partial^\alpha (P - P')(x_\nu)| \;\leq\; M |x_\nu - x_\mu|^{m-|\alpha|} \text{ for } |\alpha| \;\leq\; m-1\} \text{ for } M \in (0, \infty)\,.$$

Step 4: For each $\nu, \mu = 1, \cdots, N$, we form the blob

$$\bar{\Gamma}(x_\nu, x_\mu, \ell) \;=\; \pi \Gamma^\#(x_\nu, x_\mu, \ell) \text{ in } \mathcal{P}\,, \text{ where } \pi : \mathcal{P} \oplus \mathcal{P} \to \mathcal{P}$$

is the projection $(P, P') \mapsto P$.

Thus, $\bar{\Gamma}(x_\nu, x_\mu, \ell) = (\bar{K}_{\nu,\mu,\ell,M})_{M>0}$, where

$$\bar{K}_{\nu,\mu,\ell,M} \;=\; \{P \in \mathcal{P} :\ \text{There exists } P' \in \Gamma(x_\mu, \ell, M)\,,\ \text{with}$$

$$|\partial^\alpha (P - P')(x_\nu)| \;\leq\; M|x_\nu - x_\mu|^{m-|\alpha|} \text{ for } |\alpha| \leq m-1\} \text{ for } M \in (0,\infty)\,.$$

Step 5: For each $\nu = 1, \ldots, N$, we form the blob

$$\Gamma(x_\nu, \ell+1) \;=\; \bigcap_{\mu=1}^{N} \bar{\Gamma}(x_\nu, x_\mu, \ell) \text{ in } \mathcal{P}\,.$$

Thus, $\Gamma(x_\nu, \ell+1) = (\Gamma(x_\nu, \ell+1, M))_{M>0}$, with

$\Gamma(x_{\nu+1}, \ell{+}1, M) \;=\; \{P \in \mathcal{P} :$ For each $\mu = 1, \ldots, N,$ there exists $P' \in \Gamma(x_\mu, \ell, M)\,,$

such that $|\partial^\alpha (P - P')(x_\nu)| \leq M|x_\nu - x_\mu|^{m-|\alpha|}$ for $|\alpha| \leq m-1\}$ for $M \in (0\,,\,\infty)$.

The result of Step 5 agrees precisely with (4) in Section 1.

Using the algorithms sketched or promised in the preceding section, we will carry out the analogue of Steps 1,...,5 for ALPs. We proceed as follows.

Step $1'$: For each $\mu = 1, \ldots, N$, we form an ALP $\mathcal{A}^+_{\mu,\ell}$ in $\mathcal{P} \oplus \mathcal{P}$, such that $\mathcal{P} \oplus \mathcal{K}(\mathcal{A}_{\mu,\ell}) \;=\; \mathcal{K}(\mathcal{A}^+_{\mu,\ell})$.

Step $2'$: For each $\nu, \mu = 1, \ldots, N$, we form an ALP $\mathcal{B}_{\nu,\mu}$ in $\mathcal{P} \oplus \mathcal{P}$ such that $\mathcal{K}(\mathcal{B}_{\nu,\mu}) \;=\; \Omega_{\nu,\mu}$, with $\Omega_{\nu,\mu}$ as in Step 2.

Step $3'$: For each $\nu, \mu = 1, \ldots, N$, we form an ALP $\mathcal{A}^\#_{\nu,\mu,\ell}$ in $\mathcal{P} \oplus \mathcal{P}$ such that $\mathcal{K}(\mathcal{A}^+_{\mu,\ell}) \cap \mathcal{K}(\mathcal{B}_{\nu,\mu}) = \mathcal{K}(\mathcal{A}^\#_{\nu,\mu,\ell})$.

Step $4'$: For each $\nu, \mu = 1, \ldots, N$, we form an ALP $\bar{\mathcal{A}}_{\nu,\mu,\ell}$ in $\mathcal{P}$, such that $\mathcal{K}(\bar{\mathcal{A}}_{\nu,\mu,\ell})$ is C-equivalent to $\pi\mathcal{K}(\mathcal{A}^\#_{\nu,\mu,\ell})$, with π as in Step 4.

Step $5'$: For each $\nu = 1, \ldots, N$, we form an ALP $\tilde{\mathcal{A}}_{\nu,\ell+1}$ in $\mathcal{P}$, such that

$$\mathcal{K}(\tilde{\mathcal{A}}_{\nu,\ell+1}) \;=\; \bigcap_{\mu=1}^{N} \mathcal{K}(\bar{\mathcal{A}}_{\nu,\mu,\ell}).$$

Finally, we carry out one more step, namely

Step 6′: For each $\nu = 1, \ldots, N$, we form an ALP $\mathcal{A}_{\nu,\ell+1}$ in $\mathcal{P}$, such that length $(\mathcal{A}_{\nu,\ell+1}) \leq \dim \mathcal{P}$, and such that the blobs $\mathcal{K}(\tilde{\mathcal{A}}_{\nu,\ell+1})$ and $\mathcal{K}(\mathcal{A}_{\nu,\ell+1})$ are C-equivalent.

The algorithms sketched or promised in the previous section allow us to carry out the above six steps.

Let us compare the outcome of Steps $1'$,...,$6'$ with that of Steps 1,...,5. Comparing Steps 1 and $1'$, and recalling our inductive hypothesis $(2)_\ell$, we see that the blobs $\Gamma^+(x_\mu, \ell)$ and $\mathcal{K}(\mathcal{A}^+_{\mu,\ell})$ are C-equivalent. Since also the blobs $\Omega_{\nu,\mu}$ from Step 2 and $\mathcal{K}(\mathcal{B}_{\nu,\mu})$ from Step $2'$ are equal, it follows that the blobs $\Gamma^\#(x_\nu, x_\mu, \ell)$ and $\mathcal{K}(\mathcal{A}^\#_{\nu,\mu,\ell})$ from Steps 3 and $3'$ are C-equivalent. Consequently, the blobs $\bar{\Gamma}(x_\nu, x_\mu, \ell)$ and $\mathcal{K}(\bar{\mathcal{A}}_{\nu,\mu,\ell})$ from Steps 4 and $4'$ are C-equivalent. This in turn implies that the blobs $\Gamma(x_\nu, \ell+1)$ and $\mathcal{K}(\tilde{\mathcal{A}}_{\nu,\ell+1})$ from Steps 5 and $5'$ are C-equivalent. Since in Step $6'$, we have length $(\mathcal{A}_{\nu,\ell+1}) \leq \dim \mathcal{P}$, with $\mathcal{K}(\tilde{\mathcal{A}}_{\nu,\ell+1})$ and $\mathcal{K}(\mathcal{A}_{\nu,\ell+1})$ being C-equivalent, it follows that the $\mathcal{A}_{\nu,\ell+1}$ satisfy $(1)_{\ell+1}$ and $(2)_{\ell+1}$.

This completes our induction on ℓ. We have succeeded in computing the $\mathcal{A}_{\nu,\ell}$ modulo the algorithms sketched or promised in Section 2.

Regarding the ALPs arising in Steps $1'$,...,$6'$ above, we note that:

- $\mathcal{A}^+_{\mu,\ell}$ has length at most $\dim \mathcal{P} + 1$, since the same is true of $\mathcal{A}_{\mu,\ell}$.
- $\mathcal{B}_{\nu,\mu}$ has length $\dim \mathcal{P}$.
- $\mathcal{A}^\#_{\nu,\mu,\ell}$ has length equal to length $(\mathcal{A}^+_{\nu,\ell})$ + length $(\mathcal{B}_{\nu,\mu})$, which is at most $2 \cdot \dim \mathcal{P} + 1$.
- $\bar{\mathcal{A}}_{\nu,\mu,\ell}$ has length at most $\dim \mathcal{P}$. (See the discussion of $\pi\mathcal{K}$ in Section 2).
- $\tilde{\mathcal{A}}_{\nu,\ell+1}$ has length $\displaystyle\sum_{\mu=1}^{N} \text{length } (\bar{\mathcal{A}}_{\nu,\mu,\ell}) \leq (\dim \mathcal{P}) \cdot N$.
- $\mathcal{A}_{\nu,\ell+1}$ has length at most $\dim \mathcal{P}$, by definition.

Once we have computed $\mathcal{A}_{\nu,\ell_*}$ for $\nu = 1, \ldots, N$, we can then compute the order of magnitude of the onsets of $\mathcal{K}(\mathcal{A}_{1,\ell_*}), \ldots, \mathcal{K}(\mathcal{A}_{N,\ell_*})$, as promised in Section 2.

From estimate (5) in Section 2, we can then read off the order of magnitude of $Norm\{(x_\nu, t_\nu)_{\nu=1,\dots,N}\}$.

To complete the proof of Theorem 1, it remains to exhibit the algorithms promised in Section 2, to estimate the number of operations required for our algorithms, and to prove (6) from Section 1.

§4. Row Operations on ALPs

In this section, we show how to perform elementary row operations on ALPs, analogous to the elementary processes of linear algebra. This will be used later to exhibit the algorithms promised in Section 2. Our row operations are of three types.

To describe our first row operation, let

$$(1)\quad \mathcal{A} = [(\lambda_1, \dots, \lambda_L),\, (b_1, \dots, b_L),\, (\sigma_1, \dots, \sigma_L), M_*]$$

be an ALP in a vector space V, and let $\pi : \{1, \dots, L\} \to \{1, \dots, L\}$ be a permutation.

Then

$$\mathcal{A}^\pi \;=\; [(\lambda_{\pi 1}, \dots, \lambda_{\pi L}),\, (b_{\pi 1}, \dots, b_{\pi L}),\, (\sigma_{\pi 1}, \dots, \sigma_{\pi L}),\, M_*]$$

is again an ALP in V, and, evidently, $\mathcal{K}(\mathcal{A}) = \mathcal{K}(\mathcal{A}^\pi)$. We say that $\mathcal{A}^\pi$ arises from $\mathcal{A}$ by "permuting rows". (In the next section, we will regard each λ_ℓ as a row vector.) Our second type of row operation arises for an ALP (1) in case there is some $\bar{L} < L$ such that $\lambda_{\bar{L}+1} = \lambda_{\bar{L}+2} = \cdots = \lambda_L = 0$. In that case, for $\bar{L} < \ell \le L$, the estimate $|\lambda_\ell(v) - b_\ell| \le M\sigma_\ell$, appearing in the definition of $\mathcal{K}(\mathcal{A})$, reduces to $|b_\ell| \le M\sigma_\ell$, which is equivalent to $M \ge M_\ell^*$ for an $M_\ell^* \in [0, \infty]$ determined trivially by b_ℓ and σ_ℓ. Consequently, we have $\mathcal{K}(\mathcal{A}) = \mathcal{K}(\bar{\mathcal{A}})$, where

$$\bar{\mathcal{A}} = [(\lambda_1, \dots, \lambda_{\bar{L}}), (b_1, \dots, b_{\bar{L}}),\, (\sigma_1, \dots, \sigma_{\bar{L}}),\, \max\{M_*, M^*_{\bar{L}+1}, \dots, M^*_L\}]\,.$$

We say that $\bar{\mathcal{A}}$ arises from $\mathcal{A}$ by "stripping away zeros".

Our third row operation on ALP (1) arises by adding a multiple of one of the functionals $\lambda_1, \dots, \lambda_L$ to each of the other λ's. More precisely, let $\mathcal{A}$ be the ALP given by (1), and let $1 \le \ell_0 \le L$. Suppose we are given real coefficients $\beta_1, \dots, \beta_L$, with $\beta_{\ell_0} = 0$. We define a new ALP $\hat{\mathcal{A}}$ in V, by setting

(2) $\hat{\mathcal{A}} = [(\hat{\lambda}_1, \dots, \hat{\lambda}_L), (\hat{b}_1, \dots, \hat{b}_L), (\sigma_1, \dots, \sigma_L), M_*]$, where

(3) $\hat{\lambda}_\ell = \lambda_\ell + \beta_\ell \lambda_{\ell_0}$ and $\hat{b}_\ell = b_\ell + \beta_\ell b_{\ell_0}$ for $\ell = 1, \dots, L$.

The blobs $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\hat{\mathcal{A}})$ are then related by the following simple result.

Proposition: *Assume that $|\beta_\ell| \cdot \sigma_{\ell_0} \le \sigma_\ell$ for $\ell = 1, \dots, L$. Then the blobs $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\hat{\mathcal{A}})$ are 2-equivalent.*

Proof: Fix $M \ge M_*$, and let $v \in K_M(\mathcal{A})$ (as in (3), (4) from Section 2). Then, for $\ell = 1, \dots, L$, we have $|\lambda_\ell(v) - b_\ell| \le M\sigma_\ell$, and consequently

$$|\hat{\lambda}_\ell(v) - \hat{b}_\ell| \;=\; |[\lambda_\ell(v) - b_\ell] + \beta_\ell[\lambda_{\ell_0}(v) - b_{\ell_0})]| \le\; M\sigma_\ell + |\beta_\ell| \cdot M\sigma_{\ell_0}$$

$$\le\; 2M\sigma_\ell\,, \text{since } |\beta_\ell|\sigma_{\ell_0} \le \sigma_\ell\,.$$

This shows that

(4) $K_M(\mathcal{A}) \subseteq K_{2M}(\hat{\mathcal{A}})$

for all $M \ge M_*$. One the other hand, (4) is obvious for $M < M_*$, since $K_M(\mathcal{A})$ is empty in that case. Thus, (4) holds for all $M > 0$.

Moreover, since $\beta_{\ell_0} = 0$, (3) implies

$$\lambda_\ell = \hat{\lambda}_\ell - \beta_\ell \hat{\lambda}_{\ell_0} \text{ and } b_\ell = \hat{b}_\ell - \beta_\ell\, \hat{b}_{\ell_0} \text{ for } \ell = 1, \dots, L\,.$$

Hence, we may repeat the proof of (4), with the rôles of $\mathcal{A}$ and $\hat{\mathcal{A}}$ interchanged, to conclude that

(5) $K_M(\hat{\mathcal{A}}) \subseteq K_{2M}(\mathcal{A})$

for all $M > 0$. Inclusions (4) and (5) tell us that the blobs $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\hat{\mathcal{A}})$ are 2-equivalent.

The proof of the proposition is complete. □

When $\mathcal{A}$ and $\hat{\mathcal{A}}$ are related as in (1), (2), (3), with $\beta_{\ell_0} = 0$, then we say that $\hat{\mathcal{A}}$ arises from $\mathcal{A}$ by "row addition." If also $|\beta_\ell|\sigma_{\ell_0} \leq \sigma_\ell$ for all $\ell = 1, \ldots, L$ so that the above Proposition applies, then we say that $\hat{\mathcal{A}}$ arises from $\mathcal{A}$ by "stable row addition."

§5. Echelon Form

In this section, we use the elementary row operations from Section 4 to place a given ALP $\mathcal{A}$ into "echelon form", somewhat like the standard echelon form in linear algebra. In the next section, we use our echelon form to exhibit the algorithms promised in Section 2.

We take our vector space V to be $\mathbb{R}^D$. Let

$$(1)\ \ \mathcal{A} = [(\lambda_1, \ldots, \lambda_L), (b_1, \ldots, b_L), (\sigma_1, \ldots, \sigma_L), M_*]$$

be an ALP in V. Each functional λ_ℓ may be identified with a row vector, $\lambda_\ell = (\lambda_{\ell 1}, \ldots, \lambda_{\ell D}) \in \mathbb{R}^D$. Thus the ALP $\mathcal{A}$ may be rewritten in the form

$$(2)\ \ \mathcal{A} = [(\lambda_{\ell j})_{\substack{1\leq \ell \leq L \\ 1 \leq j \leq D}}, (b_\ell)_{1\leq \ell \leq L}, (\sigma_\ell)_{1 \leq \ell \leq L},\ M_*]\,.$$

For $0 \leq I \leq L$, we say that an ALP $\mathcal{A}$ as in (2) is in "echelon form through row I", with "pivots" $p_1, \ldots, p_I$, if the following conditions are satisfied.

$(\mathrm{EF0})_I$: The p_i are integers, and $1 \leq p_1 < p_2 < \cdots < p_I \leq D\,.$

$(\mathrm{EF1})_I$: $\lambda_{ip_i} \neq 0$ for $i = 1, \ldots, I$.

$(\mathrm{EF2})_I$: $\lambda_{ij} = 0$ for $1 \leq j < p_i$, $i = 1, \ldots, I$.

$(\mathrm{EF3})_I$: $\lambda_{ij} = 0$ for $1 \leq j \leq p_I$, $i > I$. (If $I = 0$, then this holds vacuously.)

If $I = 0$, then there are no pivots, and $(\mathrm{EF0})_I, \cdots, (\mathrm{EF3})_I$ hold vacuously. Hence, any ALP is in echelon form through row zero. On the other hand, an ALP can never be in echelon form through row I with $I > D$, as one sees at once from $(\mathrm{EF0})_I$. An ALP $\mathcal{A}$ which is in echelon form through row L, with $L=$ length $(\mathcal{A})$ as in (1), (2), will be said to be in "echelon form". Note that an ALP in $\mathbb{R}^D$ in echelon form has length at most D.

To place a given ALP into echelon form by row operations, we repeatedly apply the following result.

Lemma 1: *Let $\mathcal{A}$ be an ALP as in (2). Suppose $\mathcal{A}$ is in echelon form through row I. Then one of the following alternatives holds.*

Alternative 1: $\lambda_{\ell j} = 0$ *for all* $\ell > I$,$1 \le j \le D$.

Alternative 2: There exists an ALP $\bar{\mathcal{A}}$ in echelon form through row $I+1$, such that the blobs $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\bar{\mathcal{A}})$ are 2-equivalent. Moreover, we can compute $\bar{\mathcal{A}}$ from $\mathcal{A}$, by an algorithm that uses at most CDL operations, where C is a universal constant. Finally, $\bar{\mathcal{A}}$ and $\mathcal{A}$ have the same length.

Proof: Let $p_1, \dots, p_I$ be the pivots for $\mathcal{A}$. Suppose Alternative 1 doesn't hold. We take p_{I+1} to be the least j for which there exists $\ell > I$ with $p_{\ell j} \neq 0$. Thus,

(3) $\lambda_{\ell, p_{I+1}} \neq 0$ for some $\ell > I$,
and

(4) $\lambda_{\ell, j} = 0$ for $j < p_{I+1}, \ell > I$.
Also,

(5) $p_I < p_{I+1} \le D$,
as we see by comparing (3) with $(\mathrm{EF3})_I$.

Among all $\ell > I$ with $\lambda_{\ell, p_{I+1}} \neq 0$, we pick ℓ_0 to minimize $\sigma_\ell / |\lambda_{\ell, p_{I+1}}|$. By permuting rows in $\mathcal{A}$, we may assume without loss of generality that $\ell_0 = I+1$. Thus,

(6) $\lambda_{I+1, p_{I+1}} \neq 0$,
and

(7) $|\lambda_{\ell, p_{I+1}} / \lambda_{I+1, p_{I+1}}| \cdot \sigma_{I+1} \le \sigma_\ell$ for all $\ell > I$.

In fact, (7) holds trivially when $\lambda_{\ell, p_{I+1}} = 0$, and it follows from the minimizing property of $\ell_0 = I+1$ when $\lambda_{\ell, p_{I+1}} \neq 0$.

We now perform "addition of rows" on the ALP $\mathcal{A}$, as in Section 4, using coefficients

(8) $\beta_\ell = -\lambda_{\ell,p_{I+1}} / \lambda_{I+1,p_{I+1}}$ for all $\ell > I+1$,

(9) $\beta_\ell = 0$ for $\ell \le I+1$.

Note that $\beta_{\ell_0} = \beta_{I+1} = 0$, and that $|\beta_\ell| \cdot \sigma_{\ell_0} \le \sigma_\ell$ for all ℓ, thanks to (7), (8), (9). Hence, the Proposition in Section 4 applies. Thus, from $\mathcal{A}$, we obtain by "stable addition of rows", an ALP

(10) $\hat{\mathcal{A}} = [(\hat{\lambda}_{\ell j})_{\substack{1\le\ell\le L \\ 1\le j\le D}},\, (\hat{b_\ell})_{1\le\ell\le L},\, (\sigma_\ell)_{1\le\ell\le L},\, M_*]$,

such that

(11) The blobs $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\hat{\mathcal{A}})$ are 2-equivalent,
and

(12) $\hat{\lambda}_{\ell j} = \lambda_{\ell j} + \beta_\ell\, \lambda_{\ell_0 j}$ for all ℓ, j.

From (8), (9), (12), we see that

(13) $\hat{\lambda}_{\ell j} = \lambda_{\ell j}$ for $\ell \le I+1$,
and

(14) $\hat{\lambda}_{\ell j} = \lambda_{\ell j} - (\lambda_{\ell,p_{I+1}} / \lambda_{I+1,p_{I+1}}) \cdot \lambda_{I+1,j}$ for $\ell > I+1$.

In particular, (4) and (13), (14) give

(15) $\hat{\lambda}_{\ell j} = 0$ for $j < p_{I+1}$, $\ell \ge I+1$.

Another application of (14) gives $\hat{\lambda}_{\ell,p_{I+1}} = 0$ for $\ell > I+1$. Together with (15), this yields

(16) $\hat{\lambda}_{\ell j} = 0$ for $j \le p_{I+1}$, $\ell > I+1$.

It is now easy to check that

(17) $\hat{\mathcal{A}}$ is in echelon form through row $I+1$, with pivots $p_1, \ldots, p_{I+1}$.

In fact:

$(\mathrm{EF0})_{I+1}$ for $\hat{\mathcal{A}}$ follows from (5) and $(\mathrm{EF0})_I$ for $\mathcal{A}$.

$(\mathrm{EF1})_{I+1}$ for $\hat{\mathcal{A}}$ follows from (6), (13) and $(\mathrm{EF1})_I$ for $\mathcal{A}$.

$(\mathrm{EF2})_{I+1}$ for $\hat{\mathcal{A}}$ follows from (13), (15) and $(\mathrm{EF2})_I$ for $\mathcal{A}$.

$(\mathrm{EF3})_{I+1}$ for $\hat{\mathcal{A}}$ is precisely (16).

Thus, (17) holds. In view of (11) and (17), we find ourselves in Alternative 2. Moreover, the above argument produced $\hat{\mathcal{A}}$ from $\mathcal{A}$ by an algorithm, using at most CDL operations, as the reader may easily check. Here, C denotes a universal constant. Finally, note that $\hat{\mathcal{A}}$ and $\mathcal{A}$ above have the same length.

The proof of Lemma 1 is complete. □

Repeatedly applying Lemma 1, we can easily derive the main result of this section.

Lemma 2: *Let $\mathcal{A}$ be an ALP in $\mathbb{R}^D$. Then there exists an ALP $\mathcal{A}^\#$ in echelon form in $\mathbb{R}^D$, such that the blobs $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\mathcal{A}^\#)$ are 2^D-equivalent. Moreover, we can compute $\mathcal{A}^\#$ from $\mathcal{A}$ in at most CD^2L operations, where C is a universal constant.*

Proof: Starting at $\mathcal{A}^0 = \mathcal{A}$, which is in echelon form through row zero, we repeatedly apply Lemma 1, until we find ourselves in Alternative 1 in the statement of that lemma. (If we never reach Alternative 1, then we continue forever.) Thus, we obtain a (finite or infinite) sequence of ALPs $\mathcal{A} = \mathcal{A}^0, \mathcal{A}^1, \mathcal{A}^2, \cdots$, with $\mathcal{A}^I$ in echelon form through row I, and such that the blobs $\mathcal{K}(\mathcal{A}^I)$ and $\mathcal{K}(\mathcal{A}^{I+1})$ are 2-equivalent. An ALP in $\mathbb{R}^D$ can never be in echelon form through row $I > D$, and therefore, our sequence terminates at some $\mathcal{A}^J$, with $J \le D$. Thus, $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\mathcal{A}^J)$ are 2^D-equivalent, $\mathcal{A}^J$ is in echelon form through row J, and $\mathcal{A}^J$ satisfies Alternative 1, i.e.,

$$\mathcal{A}^J = [(\bar{\lambda}_{\ell j})_{\substack{1\le\ell\le L \\ 1\le j\le D}}, (\bar{b}_\ell)_{1\le\ell\le L}, (\bar{\sigma}_\ell)_{1\le\ell\le L}, \bar{M}_*],$$

with $\bar{\lambda}_{\ell j} = 0$ for $J < \ell \leq L$, $1 \leq j \leq D$.

Stripping away zeros from $\mathcal{A}^J$, we obtain an ALP $\mathcal{A}^\#$ in echelon form, with $\mathcal{K}(\mathcal{A}^J) = \mathcal{K}(\mathcal{A}^\#)$. Thus, $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\mathcal{A}^\#)$ are 2^D-equivalent. Moreover, the above argument produces $\mathcal{A}^\#$ from $\mathcal{A}$ using at most CD^2L operations, since we apply Lemma 1 at most D times. Here, C denotes a universal constant. The proof of Lemma 2 is complete. □

Corollary: *In Lemma 2, the ALP $\mathcal{A}^\#$ has length $\leq \min(D, \textit{length } \mathcal{A})$.*

Proof: We have length $\mathcal{A}^\# \leq D$, since $\mathcal{A}^\#$ is in echelon form. Also, Lemma 1 shows that the ALPs $\mathcal{A}^0, \mathcal{A}^1, \cdots, \mathcal{A}^J$ in the proof of Lemma 2 all have the same length. Since length $\mathcal{A}^\# \leq$ length $\mathcal{A}^J$, it follows that length $\mathcal{A}^\# \leq$ length $\mathcal{A}$. □

§6. Applications of Echelon Form

In this section, we apply our work on echelon form, to exhibit the following two algorithms for ALPs. These algorithms were promised in Section 2.

Algorithm 1: *Given an ALP $\mathcal{A}$ in $\mathbb{R}^D$, we exhibit an ALP $\mathcal{A}^\#$ of length $\leq D$ in $\mathbb{R}^D$, such that the blobs $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\mathcal{A}^\#)$ are 2^D-equivalent, and onset $\mathcal{K}(\mathcal{A}^\#)$= threshold $(\mathcal{A}^\#)$.*

Algorithm 2: *Let $\pi : \mathbb{R}^D \to \mathbb{R}^{\bar{D}}$ be the projection $(x_1, \ldots, x_D) \mapsto (x_{D-\bar{D}+1}, \ldots, x_D)$. Given an ALP $\mathcal{A}$ in $\mathbb{R}^D$, we exhibit an ALP $\bar{\mathcal{A}}$ of length $\leq \bar{D}$ in $\mathbb{R}^{\bar{D}}$, such that the blobs $\pi\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\bar{\mathcal{A}})$ are 2^D-equivalent.*

To carry out Algorithm 1, we simply take $\mathcal{A}^\#$ as in Lemma 2 in the previous section. We know that $\mathcal{A}^\#$ has length $\leq D$, and that $\mathcal{K}(\mathcal{A})$ is 2^D-equivalent to $\mathcal{K}(\mathcal{A}^\#)$. It remains to check that onset $\mathcal{K}(\mathcal{A}^\#)$ = threshold $(\mathcal{A}^\#)$. To see this, we recall the definitions.

With

(1) $\mathcal{A}^\# = [(\lambda_1, \ldots, \lambda_L), (b_1, \ldots, b_L), (\sigma_1, \ldots, \sigma_L), M_*]$,
we have

(2) threshold $\mathcal{A}^\# = M_*$,
and

(3) onset $\mathcal{K}(\mathcal{A}^\#) = inf\{M > 0 : K_M(\mathcal{A}^\#) \neq \phi\}$, where

(4) $K_M(\mathcal{A}^{\#}) = \{v \in \mathbb{R}^D : |\lambda_\ell(v) - b_\ell| \leq M\sigma_\ell \text{ for } \ell = 1, \ldots, L\}$ for $M \geq M_*$,
and

(5) $K_M(\mathcal{A}^{\#}) = \phi$ for $M < M_*$.

Consequently, if we can show that

(6) $K_M(\mathcal{A}^{\#}) \neq \phi$ for $M \geq M_*$,

then it follows that onset $\mathcal{K}(\mathcal{A}^{\#})$= threshold $\mathcal{A}^{\#}$. Thus, if we can prove (6), then we have carried out Algorithm 1.

We recall that $\mathcal{A}^{\#}$ is in echelon form. Let $1 \leq p_1 < p_2 < \cdots < p_L \leq D$ be the pivots for $\mathcal{A}^{\#}$. Thus, the functionals λ_ℓ in (1) are given by

(7) $\lambda_\ell : (x_1, \ldots, x_D) \mapsto \lambda_{\ell 1} x_1 + \lambda_{\ell 2} x_2 + \cdots + \lambda_{\ell D} x_D$,
with

(8) $\lambda_{\ell, p_\ell} \neq 0$ and $\lambda_{\ell, j} = 0$ for $j < p_\ell$.

In view of the form (7), (8) of the λ_ℓ, we can successively choose the coordinates $x_D, x_{D-1}, \ldots, x_1$ in such a way that the vector $\bar{v} = (x_1, x_2, \ldots, x_D) \in \mathbb{R}^D$ satisfies $\lambda_\ell(\bar{v}) = b_\ell$ for all $\ell = 1, \ldots, L$. The vector $\bar{v}$ belongs to $K_M(\mathcal{A})$ for all $M \geq M_*$, thanks to (4). This proves (6), and completes our discussion of Algorithm 1.

We turn our attention to Algorithm 2. Again, we produce an ALP $\mathcal{A}^{\#}$ in $\mathbb{R}^D$, as in Lemma 2 of the preceding section. Thus, (1),. . .,(8) hold, as before. Since $\mathcal{K}(\mathcal{A})$ and $\mathcal{K}(\mathcal{A}^{\#})$ are 2^D-equivalent, it follows that $\pi\mathcal{K}(\mathcal{A})$ and $\pi\mathcal{K}(\mathcal{A}^{\#})$ are 2^D-equivalent. Hence, to carry out Algorithm 2, it is enough to produce an ALP $\bar{A}$ of length $\leq \bar{D}$ in $\mathbb{R}^{\bar{D}}$, such that $\pi\mathcal{K}(\mathcal{A}^{\#}) = \mathcal{K}(\bar{A})$.

We recall from the definitions in Section 2 that

(9) $\pi\mathcal{K}(\mathcal{A}^{\#}) = (\pi K_M(\mathcal{A}^{\#}))_{M>0}$.

To understand the $\pi K_M(\mathcal{A}^\#)$, we set

(10) $\Lambda_{\ell o} = \{\ell : p_\ell \le D - \bar{D}\}$ and $\Lambda_{hi} = \{\ell : p_\ell \ge D - \bar{D} + 1\}$.

For $\ell \in \Lambda_{hi}$, (7) and (8) show that $\lambda_\ell(x_1, \dots, x_D)$ may be written in the form $\bar{\lambda}_\ell(x_{D-\bar{D}+1}, \dots, x_D)$ for a functional $\bar{\lambda}_\ell$ acting on $\mathbb{R}^{\bar{D}}$.

We set

(11) $\bar{K}_M = \{(x_{D-\bar{D}+1}, \dots, x_D) \in \mathbb{R}^{\bar{D}} : |\bar{\lambda}_\ell(x_{D-\bar{D}+1}, \dots, x_D) - b_\ell| \le M\sigma_\ell \text{ for } \ell \in \Lambda_{hi}\}$.

Comparing (4) with (10), we see at once that

(12) $\pi K_M(\mathcal{A}^\#) \subseteq \bar{K}_M$ for $M \ge M_*$.

On the other hand, suppose $M \ge M_*$ and $\bar{v} = (x_{D-\bar{D}+1}, \dots, x_D) \in \bar{K}_M$. In view of (7), (8), (10), we may successively choose $x_{D-\bar{D}}, x_{D-\bar{D}-1}, \dots, x_1$ so that $v = (x_1, \dots, x_n)$ satisfies $\lambda_\ell(v) = b_\ell$ for all $\ell \in \Lambda_{\ell o}$. Thus, $v \in K_M(\mathcal{A}^\#)$ by (4), and $\pi v = \bar{v}$ by definition of v. This proves that

(13) $\pi K_M(\mathcal{A}^\#) \supseteq \bar{K}_M$ for $M \ge M_*$.

From (5), (12), (13), we see that

(14) $\pi K_M(\mathcal{A}^\#) = \bar{K}_M$ for $M \ge M_*$,

and

(15) $\pi K_M(\mathcal{A}^\#) = \phi$ for $M < M_*$.

In view of (11), (14), (15), it is trivial to produce an ALP $\bar{\mathcal{A}}$ in $\mathbb{R}^{\bar{D}}$, such that $\pi\mathcal{K}(\mathcal{A}^\#) = \mathcal{K}(\bar{\mathcal{A}})$. Moreover, the length of $\bar{\mathcal{A}}$ is equal to the number of elements in Λ_{hi}. Since $1 < p_1 < p_2 < \cdots < p_L \le D$, we see from (10) that Λ_{hi} contains at most $\bar{D}$ elements.

Thus, we have produced an ALP $\bar{\mathcal{A}}$ of length $\le \bar{D}$ in $\mathbb{R}^{\bar{D}}$, such that $\pi\mathcal{K}(\mathcal{A}^\#) = \mathcal{K}(\bar{\mathcal{A}})$. This is the task to which we had reduced Algorithm 2.

Note that the number of operations needed to carry out Algorithms 1 and 2 is at most $CD^2{\cdot}\text{length}\ (\mathcal{A})$, where C is a universal constant.

§7. The Main Algorithm

The preceding section provides the algorithms promised in Section 2. Hence, we can compute the ALPs $\mathcal{A}_{\nu,\ell_*}(\nu = 1,\dots,N)$, as described in Section 3. As promised in Section 2 and explained in Section 6, we can compute a number Y_ν having the same order of magnitude as onset $(\mathcal{A}_{\nu,\ell_*})$, for each ν. Finally, we return the answer

(1) $ANS = \max\{Y_\nu : \nu = 1,\dots,N\}$.

This completes the description of our main algorithm.

Once we prove the key estimate (6) in Section 1, we will know that the answer returned by our algorithm is of the same order of magnitude as the desired quantity $Norm\{(x_\nu, t_\nu)_{\nu=1,\dots,N}\}$. We will prove this estimate in the next section.

Let us see how many operations are needed to carry out our main algorithm. We start with a few preliminary remarks. We write C to denote constants depending only on m and n. Since $\dim\mathcal{P}$ depends only on m and n, the reduction of an ALP $\mathcal{A}$ in $\mathcal{P}$ or $\mathcal{P}\oplus\mathcal{P}$ to echelon form (as in Section 5) requires at most $C{\cdot}\text{length}\ (\mathcal{A})$ operations. If also length $(\mathcal{A}) \le 2\dim\mathcal{P} + 1$, then at most C operations are required.

In view of the above remarks, we can easily estimate the number of operations used by the algorithms in Section 3. First of all,

(2) Computing the $\mathcal{A}_{\nu,1}$ $(\nu = 1,\dots,N)$ takes at most C operations for each ν, for a total of CN operations.

Moreover, for each $\ell = 1,2,\dots,\ell_* - 1$, we have the following.

(3) Step $1'$ takes at most C operations for each μ, for a total of at most CN operations.

(4) Steps $2'$, $3'$, $4'$ take at most C operations for each ν,μ, for a total of at most CN^2 operations.

(5) Steps $5'$ and $6'$ take at most CN operations for each ν, for a total of at most CN^2 operations.

(To see (3), (4), (5), we use the remarks on the lengths of the ALPs $\mathcal{A}^+_{\mu,\ell}$, $\mathcal{B}_{\nu,\mu}, \ldots, \mathcal{A}_{\nu,\ell+1}$ given at the end of Section 3.)

Combining (3), (4), (5), we see that it takes at most CN^2 operations to pass from the ALPs $(\mathcal{A}_{\nu,\ell})_{\nu=1,\ldots,N}$ to the ALPs $(\mathcal{A}_{\nu,\ell+1})_{\nu=1,\ldots,N}$ for a given ℓ. Recalling that ℓ_* depends only on m and n, we conclude that it takes at most CN^2 operations to compute all the ALPs $\mathcal{A}_{\nu,\ell_*}$ $(\nu = 1, \ldots, N)$.

Once we know these, we can compute a single Y_ν in (1) using at most C operations. Hence, we may pass from the $\mathcal{A}_{\nu,\ell_*}(\nu = 1, \ldots, N)$ to the answer (1) in at most CN operations.

Altogether, then, our main algorithm requires at most CN^2 operations, as asserted in Theorem 1. The bulk of the work goes into producing the ALPs $\mathcal{A}_{\nu,\ell_*}$.

We leave it to the reader to check that the storage required by our main algorithm is at most CN, which is optimal, since the statement of the problem already requires storage CN. ("Storage" means here the maximum number of real numbers that can be held in memory.)

§8. The Proof

In this section, we complete the proof of Theorem 1 by establishing the key estimate (6) from Section 1. We will reduce matters to the following theorem, which is a special case of the main results in [11]. (It had been conjectured earlier by Brudnyi-Shvartsman and proven by them for $m = 2$ [7].) We write $\#(S)$ to denote the number of elements in a set S. (If S is infinite, then $\#(S) = +\infty$.) Recall that $\mathcal{P}$ denotes the vector space of $(m-1)^{\underline{\text{rst}}}$ degree polynomials on $\mathbb{R}^n$.

Theorem 2: *Given $m, n \geq 1$, there exists $k^\#$, depending only on m and n, for which the following holds.*

Let $E \subset \mathbb{R}^n$ be finite, let $f : E \to \mathbb{R}$, and let $M \in (0, \infty)$. Assume that, given any $S \subseteq E$ with $\#(S) \leq k^\#$, there exists a map $y \mapsto P^y$, from S into $\mathcal{P}$, such that:

(a) $|\partial^\alpha P^y(y)| \leq M$ *for* $|\alpha| \leq m-1, y \in S$;

(b) $|\partial^\alpha(P^y - P^{y'})(y)| \leq M|y-y'|^{m-|\alpha|}$ *for* $|\alpha| \leq m-1, y, y' \in S$; *and*

(c) $P^y(y) = f(y)$ *for all* $y \in S$.

Then there exists a C^m *function* F *on* $\mathbb{R}^n$, *such that* $\| F \|_{C^m(\mathbb{R}^n)} \leq CM$, *and* $F = f$ *on* E. *Here,* C *depends only on* m *and* n.

We will also use an elementary "clustering lemma" from [12].

Lemma 1: *Let* $\ell \geq 1$, *and let* $S \subset \mathbb{R}^n$, *with* $\#(S) = \ell + 1$. *Then we can partition* S *into subsets* $S_0, S_1, \ldots, S_{\nu_{\max}}$, *such that*

(a) $\#(S_\nu) \leq \ell$ *for each* ν, *and*

(b) distance $(S_\mu, S_\nu) > c \cdot$ *diameter* (S) *for* $\mu \neq \nu$, *with* c *depending only on* ℓ.

Returning to the setting of Section 1, we set $E = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$, and we define $f : E \to \mathbb{R}$ by $f(x_\nu) = t_\nu$ $(\nu = 1, \ldots, N)$. By induction on $\ell \geq 1$, we will establish the following result.

Lemma 2: *Suppose* $y_0 \in E$, $\ell \geq 1$, $M > 0$, *and* $P \in \Gamma(y_0, \ell, M)$. *Then, for any* $S \subseteq E$, *with* $y_0 \in S$ *and* $\#(S) \leq \ell$, *there exists a map* $y \mapsto P^y$ *from* S *into* $\mathcal{P}$, *such that*:

(A) $P^{y_0} = P$;

(B) $|\partial^\alpha P^y(y)| \leq CM$ *for* $|\alpha| \leq m-1$, $y \in S$;

(C) $|\partial^\alpha(P^y - P^{y'})(y)| \leq CM|y-y'|^{m-|\alpha|}$ *for* $|\alpha| \leq m-1$, $y, y' \in S$; *and*

(D) $P^y(y) = f(y)$ *for all* $y \in S$.

Here, C *depends only on* m, n, ℓ.

Proof: For $\ell = 1$, we have $S = \{y_0\}$. We take $P^{y_0} = P$. Thus, (A) holds by definition, (B) and (D) hold since $P \in \Gamma(y_0, 1, M)$, and (C) holds since $y = y' = y_0$ for $y, y' \in S$. This proves Lemma 2 for $\ell = 1$.

For the inductive step, suppose Lemma 2 holds for a given ℓ. We will prove Lemma 2 with $(\ell+1)$ in place of ℓ.

Thus, suppose $P \in \Gamma(y_0, \ell+1, M)$, and let $S \subseteq E$, with $y_0 \in S$ and $\#(S) \leq \ell+1$. We must produce a map $y \mapsto P^y$ satisfying (A),…, (D). If $\#(S) \leq \ell$, then the desired map exists, thanks to our induction hypothesis. Hence, we may suppose that $\#(S) = \ell+1$. Let $\delta=$ diameter $(S) > 0$. We write c, C, C', etc., to denote constants depending only on m, n, and ℓ.

By Lemma 1, we may partition S into non-empty subsets S_ν $(0 \leq \nu \leq \nu_{\max})$, with the following properties.

(1) $\#(S_\nu) \leq \ell$ for each ν.

(2) $y_0 \in S_0$.

(3) distance $(S_\nu, S_{\nu'}) \geq c\delta$ if $\nu \neq \nu'$.

For each ν $(1 \leq \nu \leq \nu_{\max})$, pick $y_\nu \in S_\nu$. Thus, $y_\nu \in S_\nu$ for $0 \leq \nu \leq \nu_{\max}$.

Since $P \in \Gamma(y_0, \ell+1, M)$, we know that, for each ν, there exists

(4) $P_\nu \in \Gamma(y_\nu, \ell, M)$,

with

(5) $|\partial^\alpha (P_\nu - P)(y_0)| \leq M|y_\nu - y_0|^{m-|\alpha|}$ for $|\alpha| \leq m-1$.

In particular,

(6) $P_0 = P$.

We fix $P_0, P_1, \ldots, P_{\nu_{\max}}$ as in (4), (5), (6).

For each ν, we may apply our induction hypothesis to the point y_ν, the set S_ν, and the polynomial P_ν, thanks to (1) and (4). Hence, there is a map $y \mapsto P^y$ from S_ν into $\mathcal{P}$, satisfying:

(7) $P^{y_\nu} = P_\nu$;

(8) $|\partial^\alpha P^y(y)| \leq CM$ for $|\alpha| \leq m-1$, $y \in S_\nu$;

(9) $|\partial^\alpha (P^y - P^{y'})(y)| \leq CM|y - y'|^{m-|\alpha|}$ for $|\alpha| \leq m-1$, $y, y' \in S_\nu$; and

(10) $P^y(y) = f(y)$ for all $y \in S_\nu$.

Combining the above maps on the S_ν into a single map $y \mapsto P^y$ from S into $\mathcal{P}$, we obtain the following results from (6),. . .,(10).

- $P^{y_0} = P$
- $|\partial^\alpha P^y(y)| \leq CM$ for $|\alpha| \leq m-1$, $y \in S$;
- $P^y(y) = f(y)$ for all $y \in S$.

Thus, our map $y \mapsto P^y$, from S into $\mathcal{P}$, satisfies properties (A), (B), (D) from the statement of Lemma 2. Also, (9) shows that property (C) holds, provided y and y' belong to the same S_ν.

Hence, to complete the proof of Lemma 2, it is enough to prove (C) in the case $y \in S_\nu$, $y' \in S_{\nu'}$, $\nu \neq \nu'$. In view of (3) (with $\delta = \text{diam } S$), this means that

(11) $|\partial^\alpha (P^y - P^{y'})(y)| \leq CM\delta^{m-|\alpha|}$ for $|\alpha| \leq m-1$, $y \in S_\nu$, $y' \in S_{\nu'}$, $\nu \neq \nu'$.

Thus, the proof of Lemma 2 is reduced to (11).

Let $y \in S_\nu$, $y' \in S_{\nu'}$, with $\nu \neq \nu'$. From (7) and (9), we have

(12) $|\partial^\alpha (P^y - P_\nu)(y)| \leq CM|y - y_\nu|^{m-|\alpha|} \leq CM\,\delta^{m-|\alpha|}$ for $|\alpha| \leq m-1$, and

(13) $|\partial^\alpha (P^{y'} - P_{\nu'})(y')| \leq CM|y' - y_{\nu'}|^{m-|\alpha|} \leq CM\,\delta^{m-|\alpha|}$ for $|\alpha| \leq m-1$.

Also, (5) shows that

(14) $|\partial^\alpha (P_\nu - P_{\nu'})(y_0)| \leq M|y_\nu - y_0|^{m-|\alpha|} + M|y_{\nu'} - y_0|^{m-|\alpha|} \leq CM\,\delta^{m-|\alpha|}$ for $|\alpha| \leq m-1$.

Since $y_0, y', y \in S$, we have $|y' - y|$, $|y_0 - y| \leq \text{diam }(S) = \delta$. Consequently, (13) and (14) imply

(15) $|\partial^\alpha(P^{y'} - P_{\nu'})(y)| \leq CM\,\delta^{m-|\alpha|}$ for $|\alpha| \leq m-1$,

and

(16) $|\partial^\alpha(P_\nu - P_{\nu'})(y)| \leq CM\,\delta^{m-|\alpha|}$ for $|\alpha| \leq m-1$.

From (12), (15), (16), we obtain the desired estimate (11).

The proof of Lemma 2 is complete. □

We now prove estimate (6) from Section 1. We take $\ell_* = k^\#$ as in Theorem 2. Thus, ℓ_* depends only on m and n. As in Lemma 2, we define $E = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$, and define $f : E \to \mathbb{R}$ by setting $f(x_\nu) = t_\nu$ for $\nu = 1, \ldots, N$. Let $M > 0$ satisfy the hypothesis of (6), namely

(17) $\Gamma(x_\nu, \ell_*, M) \neq \phi$ for each ν.

We will show that the hypotheses of Theorem 2 hold, for the set E, the function f, and the constant CM, for a large enough C depending only on m and n. To see this, let $S \subset E$, with $\#(S) \leq k^\#$. If S is empty, there is nothing to prove. If S is non-empty, then we pick $y_0 \in S$ and then pick $P \in \Gamma(y_0, \ell_*, M)$. (We can find such a P, thanks to (17).) Applying Lemma 2, with $\ell = \ell_* = k^\#$, we obtain a map $y \mapsto P^y$ from S into $\mathcal{P}$, satisfying

(18) $|\partial^\alpha P^y(y)| \leq CM$ for $|\alpha| \leq m-1$, $y \in S$;

(19) $|\partial^\alpha(P^y - P^{y'})(y)| \leq CM|y-y'|^{m-|\alpha|}$ for $|\alpha| \leq m-1$, $y, y' \in S$; and

(20) $P^y(y) = f(y)$ for all $y \in S$.

In (18), (19), the constant C depends only on m, n and ℓ_*. Since ℓ_* depends only on m and n, it follows that C depends only on m and n. The existence of a map $y \mapsto P^y$ satisfying (18), (19), (20) is precisely the hypothesis of Theorem 2.

Applying Theorem 2, we obtain a function $F \in C^m(\mathbb{R}^n)$, such that:

(21) $\| F \|_{C^m(\mathbb{R}^n)} \leq CM$, with C depending only on m and n; and

(22) $F = f$ on E, i.e., $F(x_\nu) = t_\nu$ for $\nu = 1, \ldots, N$.

From (21), (22) and the definition of $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$, we conclude that $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\} \le CM$, with C depending only on m and n.

This is precisely the conclusion of (6) from Section 1.

The proof of (6) from Section 1 is complete, and with it, the proof of Theorem 1. □

§9. Open Problems

We close the paper by posing a few open problems connected with Theorem 1. First of all, we would like to compute a particular function $F \in C^m(\mathbb{R}^n)$, whose graph passes through N given points $(x_1, t_1), \ldots, (x_N, t_N)$, and whose C^m-norm has the least possible order of magnitude. Say, we enter the data $\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$ into an (idealized) computer. The computer works for a while, performing L_0 operations. It then signals that it is ready to accept further input. Whenever we enter a point $x \in \mathbb{R}^n$, the computer responds by producing an output $F(x)$, using L_1 operations to make the computation. It is guaranteed that $F(x_\nu) = t_\nu$ for $\nu = 1, \ldots, N$, and that $\| F \|_{C^m(\mathbb{R}^n)}$ has the same order of magnitude as $Norm\{(x_\nu, t_\nu)_{\nu=1,\ldots,N}\}$.

We would like to give an algorithm to carry this out, with control over L_0 and L_1.

It is known [8,10] that F as above can be chosen to depend linearly on $t_1, \ldots, t_N$, for fixed $x_1, \ldots, x_N, m, n$. How small can we take L_0 and L_1 above, and maintain this linear dependence?

There are obvious analogues of these questions, with $Norm\{(x_\nu, t_\nu)\}$ replaced by the more general $Norm\{(x_\nu, t_\nu, \sigma_\nu)_{\nu=1,\ldots,N}\}$, mentioned in the introduction. For that matter, we could bring in "Whitney convex sets" in place of the σ_ν. (See [12,14] .)

By combining the ideas in this paper with those of [10,...,16], one can probably give algorithms to compute F as above. However, we have no reason to believe that our methods lead to optimal L_0,L_1.

It is natural to view the computation of $Norm\{(x_\nu, t_\nu, \sigma_\nu)_{\nu=1,\ldots,N}\}$ as "fitting a C^m-smooth function to data". From this viewpoint, it is also natural to allow outliers. We have analyzed the ℓ^∞-norm of the error vector $[(F(x_\nu) - t_\nu)/\sigma_\nu]_{\nu=1,\dots,N}$, for $F \in C^m(\mathbb{R}^n)$, when $\sigma_1, \dots, \sigma_N > 0$. Standard least squares analyzes the ℓ^2-norm of the error vector, for F in a Hilbert space of functions. It would be very interesting, and probably very hard, to say something intelligent about the ℓ^p-norm of the error vector for other p, or for other function spaces (eg, Sobolev spaces $W^{m,p}$). We have no idea how to start on this problem.

§10. 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. A. Brudyni and Y. Brudyni, *Metric spaces with linear extensions preserving Lipschitz condition*, (to appear).

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 subsets of $\mathbb{R}^n$, in Linear and Complex Analysis*, Lect. Notes in Math., Springer-Verlag (1994), pp. 279-281.

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.

7. Y. Brudnyi and P. Shvartsman, *Generalizations of Whitney's extension theorem*, IMRN **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, (to appear).

11. C. Fefferman, *A sharp form of Whitney's extension theorem*, Annals of Math., (to appear).

12. C. Fefferman, *Whitney's extension problem for C^m*, Annals of Math., (to appear).

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

14. C. Fefferman, *A generalized sharp Whitney theorem for jets*, Revista Mathematica Iberoamericana, (to appear).

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

16. C. Fefferman, *C^m Extension by Linear Operators*, (to appear).

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

18. S. Har-Peled and M. Mendel, *Fast construction of nets in low-dimensional metrics, and their applications*, arXiv:cs.DS/0409057 v 1, (2004).

19. B. Malgrange, Ideals of Differentiable Functions, Oxford U. Press, 1966.

20. 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.

21. 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.

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

23. E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton U. Press, 1970.

24. R. Webster, Convexity, Oxford Science Publications, 1994.

25. H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Transactions A.M.S. **36** (1934), 63-89.

26. H. Whitney, *Differentiable functions defined in closed sets I*, Transactions A.M.S. **36** (1934), 369-3897.

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

28. N. Zobin, *Whitney's problem on extendability of functions and an intrinsic metric*, Advances in Math., vol. 133, **1**, (1998), 96-132.

29. N. Zobin, *Extension of smooth functions from finitely connected planar domains*, Journal of Geom. Analysis, vol. 9, **3**, (1999), 489-509.

June 16, 2005:gpp