Averaging Sets:
A Generalization of Mean Values and Spherical Designs

P. D. SEYMOUR* and THOMAS ZASLAVSKY†

The Ohio State University, Columbus, Ohio 43210

1. INTRODUCTION, MAIN THEOREM, AND APPLICATIONS

The mean value theorem in its integral form states that, for any continuous function \( f: [0, 1] \rightarrow \mathbb{R} \), there is a point \( x \in (0, 1) \) at which \( f \) assumes its average value, that is,

\[
f(x) = \int_0^1 f(t) \, dt.
\]

Given several functions \( f_1, \ldots, f_m: [0, 1] \rightarrow \mathbb{R} \), one cannot of course expect there to be a single point \( x \) at which every one assumes its mean value, but there is an easy generalization in terms of weighted averages: there exist at most \( m \) points \( x_i \) and positive weights \( a_i \) summing to 1 so that

\[
\sum a_i f_j(x_i) = \int_0^1 f_j(t) \, dt \quad \text{for} \quad j = 1, 2, \ldots, m.
\]

If one wants all the weights to be equal the problem becomes much harder. We prove here that it is possible to find a finite averaging set \( \{x_1, x_2, \ldots, x_n\} \subseteq (0, 1) \), whose unweighted average value equals the integral mean value of \( f_j \) for all the given functions \( f_1, \ldots, f_m \). This was apparently not previously proved even for polynomial functions.

A spherical design of strength \( t \) [4] is a finite subset \( X \) of the unit sphere \( S^d \subseteq \mathbb{R}^{d+1} \) such that, for any polynomial \( f \) of degree at most \( t \), the total value \( \sum \{f(x): x \in X\} \) is invariant under the action of the orthogonal group. One can show [4, p. 372] that \( X \) is a spherical \( t \)-design if and only if it is an averaging set on the sphere for polynomials of degree up to \( t \), that is,

\[
\frac{1}{|X|} \sum_{x \in X} f(x) = \frac{1}{\mu(S^d)} \int_{S^d} f \, d\mu
\]

* Research supported by NSF Grant MCS-8103440. Present address: Bell Communications Research, 600 Mountain Ave., Murray Hill, N.J. 07974.
† Research supported by SGPNR Grant 8110Z354.
for all such polynomials \(f\). (Here \(\mu\) is the usual spherical measure.) The existence of spherical designs has been known for all \(t\) only in the easy case \(d = 1\). (For a partial classification of these "circular" designs see [6].) Small examples are also known for some other parameters \(d\) and \(t\). We prove here that spherical designs exist for all values of \(d\) and \(t\) and that there exist such designs of every sufficiently large size. (How large depends on the parameters and is not decided here.)

Our main theorem is quite general. Let \(\Omega\) be a path-connected topological space provided with a measure \(\mu\) that is finite and positive with full support (that is, \(\mu(S) \geq 0\) for every measurable set and \(\mu(U) > 0\) for every nonvoid open set). It is not necessary that \(\mu\) be countably additive. Let

\[f_1, \ldots, f_m : \Omega \to \mathbb{R}^p\]

be continuous, integrable functions. For instance, they may be a spanning set in a finite-dimensional linear space \(F\) of continuous, integrable functions \(\Omega \to \mathbb{R}^p\); an important example is \(F\) the space of polynomials of degree at most \(t\) in \(x_1, x_2, \ldots, x_d\), if \(\Omega \subseteq \mathbb{R}^d\). An averaging set for \(f_1, \ldots, f_m\) (or, for \(F\)) is a finite set \(X \subseteq \Omega\) having the property

\[
\frac{1}{|X|} \sum_{x \in X} f_j(x) = \frac{1}{\mu(\Omega)} \int_{\Omega} f_j d\mu \quad \text{for } j = 1, 2, \ldots, m.
\]

**Main Theorem.** Given \(\Omega, \mu,\) and \(f_1, \ldots, f_m\) as described, there exist averaging sets \(X\). The size of \(X\) may be any number, with a finite number of exceptions. And \(X\) may be chosen so that the vectors \((f_1(x), \ldots, f_m(x))\) for \(x \in X\) are all distinct.

The most important special case is that in which \(\Omega \subseteq \mathbb{R}^d\) with some measure \(\mu\) and \(F\) the linear space of all polynomials of degree at most \(t\) in \(x_1, x_2, \ldots, x_d\). Then we call \(X\) a (moment) \(t\)-design on \(\Omega\) with respect to \(\mu\). (We may omit mention of \(\mu\) if it is Lebesgue measure.) Spherical designs are one example. Another is that where \(\Omega\) is a region (by which we mean a connected, nonvoid subset of \(\mathbb{R}^d\) that lies in the closure of its own interior) and the measure \(\mu\) is determined by \(d\mu = wdx\), where \(w : \Omega \to \mathbb{R}\) is a positive weighting function whose moments on \(\Omega\) up to order \(t\) are finite.

**Corollary 1.** For each pair of integers \(d\) and \(t > 0\), and for all sufficiently large \(n\), there exist \(d\)-dimensional spherical \(t\)-designs having size \(n\).

**Corollary 2.** For each integers \(d\) and \(t > 0\), region \(\Omega \subseteq \mathbb{R}^d\), and positive weight function \(w : \Omega \to \mathbb{R}\) whose moments up to order \(t\) on \(\Omega\) are finite, and for all sufficiently large \(n\), there exist moment \(t\)-designs of size \(n\).
on $\Omega$ with respect to the measure determined by $wdx$. In particular $t$-designs exist on any bounded region $\Omega$ with respect to Lebesgue measure.

In fact our proof shows that the $t$-designs in each case can be taken to lie in no proper affine subspace of $\mathbb{R}^d$. It also shows that, if a multiset is satisfactory, it is not necessary to have more distinct points than the dimension of the space of polynomial functions of degree at most $t$ on $S^d$ (in Corollary 1) or $\mathbb{R}^d$ (in Corollary 2). (This number is surely further reducible if $d \geq 2$.)

The existence of moment $t$-designs implies the existence of two kinds of experimental designs, all-bias response surface designs and rotatable designs. Each of these is a finite set or multiset $X \subseteq \mathbb{R}^d$. We regard each point $x \in X$ as a list of values to be assigned to the controllable factors in one run of an experiment. In each run a response variable $y$ is measured; we assume that $y$ is given exactly by an unknown polynomial $g_1$ of degree at most $d_1$. Then an approximating polynomial $g_2$ of degree at most $d_2 \leq d_1$ is fitted to the resulting data, let us say by least squares. We require $X$ to be chosen so that $g_2$ is uniquely determinable. (Hence $X$ lies in no affine subspace of $\mathbb{R}^d$.)

In the case of a response surface design we choose $d_2 < d_1$. The objective is to choose $X$ so as to minimize the expected deviation of $g_2$ from $g_1$ on a region $\Omega$ of interest, of which a suitable measure is the integral mean square deviation $\mu(\Omega)^{-1} \int_\Omega (g_2 - g_1)^2 \, d\mu$, $\mu = \text{Lebesgue measure}$. Then $X$ is called an all-bias response surface design of order $d_2$ and degree $d_1$. (We are ignoring other properties that are desirable in a response surface design [2]. Note that it is not required that $X \subseteq \Omega$.) It is proved in [2, Appendix 1] that with the above fitting method and measure of deviation, and assuming one can neglect error other than the “bias” error due to omitting from $g_2$ the terms of high order, a sufficient condition for $X$ to be an all-bias response surface design is that it be a moment $(d_1 + d_2)$-design on $\Omega$. Thus we have, with the stated assumptions:

**Corollary 3.** For any $d_1 > d_2 > 0$ and any bounded region $\Omega \subseteq \mathbb{R}^d$, there exist all-bias response surface designs of order $d_2$ and degree $d_1$ having every sufficiently large size $n$. One can take the designs to be without repetition and contained in $\Omega$.

Of course in statistical applications $n$ should be small rather than large, so our results are not of practical significance. However, it has apparently not been proved before that designs can be found in all cases.

A natural generalization is to allow a positive weight function $w: \Omega \rightarrow \mathbb{R}$ and calculate the bias error by weighted integration. Our theorem, combined with Mallows' generalization of that of Box and Draper [2, p. 634], implies that such generalized all-bias response surface designs also exist.

In the case of a rotatable design we take $d_2 = d_1$. We want $X$ to be such
that the "variance function" [3, Section 4] of the approximation $g_2$ is invariant under rotation around the origin. Equivalently, by [3, Section 5], the moments of $X$ of each order up to $2d_1$ must agree with those of a spherical distribution. That is, there are a ball $B_R$ centered at the origin (whose radius $R$ may be infinite) and a positive "distribution" function $w: [0, R) \to \mathbb{R}$ such that $\int_0^R r^{d+k-1} w(r) dr$ is finite for $k = 0, 1, \ldots, 2d_1$ and $X$ is a moment $2d_1$-design on $\Omega = B_R$ with respect to the measure determined by $du = w(r) \, dr$. (Here $r = r(x)$ denotes the radial distance of the point $x$.) We call $X$ a rotatable design of order $d_1$ for the distribution function $w$.

**COROLLARY 4.** For any dimension $d$, distribution function $w$, and order $d_1$, there exists a rotatable design of each sufficiently large size $n$.

There is a notion of singularity for rotatable designs, which means the singularity of a certain moment matrix. Since $\Omega$ is a solid body, our proof implies that the design can be made nonsingular.

A stronger corollary is that rotatable designs can be constructed from spherical designs. If $Y \subseteq \mathbb{R}^d$, $rY$ denotes the set of multiples $ry$ for $y \in Y$.

**COROLLARY 5.** Let $d$ be a dimension and let $w: [0, R) \to \mathbb{R}$ be a distribution function. For each sufficiently large $N$ there are radii $r_1, r_2, \ldots, r_N < R$, which can be taken to be distinct, such that for every $(d - 1)$-dimensional spherical design $Y$ of strength $2d_1$, the union $X = r_1 Y \cup r_2 Y \cup \cdots \cup r_N Y$ is a $d$-dimensional rotatable design of order $d_1$, for the distribution $w$.

**Proof.** Let us say that $Y$ has radius 1. A choice of set (or multiset) $Z = \{r_1, r_2, \ldots, r_N\}$ will make $X$ a rotatable design if and only if $Z$ is a moment $2d_1$-design (for even powers only) on the interval $[0, R)$ with respect to $wdx$. Such a set exists by Corollary 2.

To show that the stated condition on $Z$ is necessary and sufficient we calculate moments. For $\alpha = (\alpha_1, \ldots, \alpha_d)$ and $x \in \mathbb{R}^d$, we put $|\alpha| = \alpha_1 + \cdots + \alpha_d$ and $x^n = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$. All $\alpha_i$ will be nonnegative integers. The weighted moment of $B_R$ of type $\alpha$ is given by

$$[\alpha; B_R]_w = \frac{\int_{B_R} x^\alpha \, w(r) \, dx}{\int_{B_R} w(r) \, dx} = [\alpha; (0, R)]_w [\alpha; S^{d-1}],$$

where $[\alpha; S^{d-1}]$ is the ordinary moment of $S^{d-1}$. The moment of $X$ is given by

$$[\alpha; X] = \frac{1}{|X|} \sum_{x \in X} x^\alpha = \left( \frac{1}{N} \sum_{j=1}^N r_j^{|\alpha|} \right) [\alpha; Y] = [\alpha; Z] [\alpha; Y].$$

The stated condition is equivalent to $X$ being a rotatable design because $Y$, being a spherical $2d_1$-design, satisfies $[\alpha; Y] = [\alpha; S^{d-1}]$ for $|\alpha| \leq 2d_1$. 


We give three different proofs of the Main Theorem, one fully general and the others more or less specialized. The first one (Section 4) is for analytic and, after slight adaptation, continuously differentiable functions. This is our original proof, which we developed to handle the existence problem for $t$-designs on the unit interval. We include it here, although it is not very general, both because it is quite different from the other two proofs and because it is the starting point for our proof of continuous variability of the averaging set with respect to the functions (Section 7.1). The second proof (Section 5) is for continuous functions on a space $\Omega$ in which we can find a simple path linking any $mp + 1$ points. (Such spaces include all manifolds and consequently all the cases discussed in the corollaries.) This condition enables us to combine steps and yields a considerably more elegant argument than that in our third proof (Section 6), where we treat the most general case.

The first step in all our proofs is to replace the measure and integral by a convexity condition (Section 3). That is possible because the centroid of $f = (f_1, \ldots, f_m)$ is a weighted average of image points of $f$. After that the proofs diverge. In the differentiable case we vary points repeatedly in $\Omega$, using the inverse and implicit function theorems first to rationalize the weights, then to deform multiple points so as to split them apart without changing the sum of their values. (It is because of the smoothness of the inverse and implicit functions that we can deduce a continuous dependence of $X$ on $f$.) This is essentially our original proof for $(0, 1)$, applied to manifolds by confining our attention to suitable curves. In the continuous case things are more difficult because we do not have an inverse function. Instead we prove surjectivity lemmas (Section 5) that permit us first to get averaging multisets and then to separate the multiple points (Section 6). This proof is fairly complicated since it depends on a kind of local surjectivity and on perturbing trial points. The more special proof of Section 5, where $\Omega$ is restricted enough that we can essentially work in the interval $(0, 1)$ instead of $\Omega$ itself, avoids these complexities by replicating and perturbing the weighted averaging set beforehand, thus providing a sufficient variety of points from which to choose those of $X$.

We believe that $X$ can be found so as to depend on $f$ in a locally continuous way, with minor restrictions. We have not settled the conjecture in general, but in Section 7 we prove a version of it for differentiable functions. (Not every averaging set for $f$ can be varied locally; we suspect that small examples, like those sought in [4], are sporadic and depend on $\Omega$ and $f$ being especially nice.) We can also deduce from our proofs of the Main Theorem that the minimum size of an averaging set for $f$ is locally bounded, at least for spaces having simple curves as described above (and in Section 5). In fact we have a formula for such a bound, although one that is hard to evaluate even in the simple case of $t$-designs on an interval. On the other hand we show that there is no global bound on the minimum size of an
averaging set in terms of $\Omega$, $f$, and the codomain dimension. Consequently there is no globally continuous dependence of $X$ on $f$.

By contrast an averaging multiset always exists with a small number of distinct points, independent of $f$ and $\Omega$. In Section 8 we explore the exact necessary size of an averaging multiset, given the dimension of the codomain space. In the case of planar codomain we have an exact solution; in general we are unable to decide among three consecutive integers.

In Section 9 we conclude with a question: How can our theorem be extended to discontinuous functions? The mean value theorem itself shows that some such extension is possible.

2. Notation and Terminology

For a set $S \subseteq \mathbb{R}^p$, we denote by $\text{lin} S$ and $\text{aff} S$ the linear and affine subspaces generated by $S$ and by $\text{conv} S$ the convex hull of $S$. The relative interior of $S$ is the interior of $S$ as a subset of $\text{aff} S$; we denote it by $\text{relint} S$.

All vectors are column vectors. For simplicity of notation we write them horizontally in the text.

For an integrable function $f: \Omega \to \mathbb{R}^p$, we call the point

$$\frac{1}{\mu(\Omega)} \int_{\Omega} f \, d\mu$$

the centroid of $f$ (with respect to $\mu$). The image of $f$ we denote by $\text{Im} f$. We call $f: \Omega \to \mathbb{R}^p$ degenerate or say it has degenerate image if $\text{aff}(\text{Im} f)$ is less than the whole of $\mathbb{R}^p$.

By a curve we mean a continuous image of $(0, 1)$ or the associated mapping of $(0, 1)$. A path is a continuous image of $[0, 1]$ or the associated mapping. A curve or path is simple if it is an injection. The principal chord of a path $g: [0, 1] \to \mathbb{R}^p$ is the vector $g(1) - g(0)$; we write it $\Delta g$.

A useful convention is that a sum $\sum_{i=1}^m$, where $l > m$, should be interpreted as 0 if $l = m + 1$ and as $-\sum_{i=m+1}^{l-1}$ if $l \geq m + 2$.

3. Reduction to Geometry

We begin by observing that we can reduce to a single function

$$f: \Omega \to \mathbb{R}^p$$

by setting $f = (f_1, \ldots, f_m): \Omega \to \mathbb{R}^m$. From now on we consider only the case (*).

The first lemma shows that the measure and integral are not essential.
**Lemma 3.1.** Let $\Omega$ be a topological space and $\mu$ a finite, positive measure on $\Omega$ with full support. Let $f: \Omega \to \mathbb{R}^p$ be continuous and integrable and let $c$ be the centroid of $f$ with respect to $\mu$. Then

$$c \in \text{relint conv}(\text{Im}f).$$

**Proof.** If $h: \mathbb{R}^p \to \mathbb{R}$ is a nonconstant functional of the form $h(x) = a_0 + \sum_{i=1}^{p} a_i x_i$ whose positive closed half-space contains $\text{Im}f$, then

$$h(c) = \frac{1}{\mu(\Omega)} \int_{\Omega} h \circ f \, d\mu \geq 0.$$ 

So $c \in \text{conv}(\text{Im}f)$. If $\ker h \not\supset \text{Im}f$, then there is an $\varepsilon > 0$ such that $\text{Im}f$ meets $h^{-1}((\varepsilon, \infty))$. Then $U = (h \circ f)^{-1}((\varepsilon, \infty))$ is nonvoid and, by the continuity of $f$, open. So

$$h(c) \geq \varepsilon \mu(U) > 0.$$

We conclude that $c \in \text{relint conv}(\text{Im}f)$.

Throughout the rest of the proofs, instead of taking for $c$ the centroid of $f$ with respect to $\mu$ we let it be any point in the relative interior of $\text{conv}(\text{Im}f)$. Then we can forget about the measure altogether. (This is not really a big change since we can locate the centroid anywhere in $\text{relint conv}(\text{Im}f)$ by choosing a suitable measure.) We prove the following result.

**Basic Theorem.** Let $\Omega$ be a path-connected topological space, $f: \Omega \to \mathbb{R}^p$ a continuous function, and $c \in \text{relint conv}(\text{Im}f)$. Then there is a finite set $X \subseteq \Omega$ for which

$$\frac{1}{|X|} \sum_{x \in X} f(x) = c.$$

One can choose $X$ so that all values $f(x)$ are distinct and the size of $X$ is any sufficiently large positive integer (depending on $f$).

For convenience we call $X$ an *averaging set* for $f$ and $c$. To standardize further we assume

$$c = 0 \quad \text{and} \quad \text{aff}(\text{Im}f) = \mathbb{R}^p,$$

respectively by translation and by cutting down the dimension of the codomain. Thus

$$0 \in \text{int conv}(\text{Im}f).$$
It follows by Caratheodory's Theorem that there are points $y_0, y_1, ..., y_p \in \Omega$ such that

$$0 \in \text{int } \text{conv}\{f(y_0), f(y_1), ..., f(y_p)\}.$$ 

That is the starting point of all our proofs.

4. PROOFS FOR ANALYTIC AND DIFFERENTIABLE FUNCTIONS

We assume first that $\Omega$ is an analytic manifold and $f$ is analytic. Later we will show how to relax this requirement.

For each $i = 1, ..., p$, let $P_i$ be an analytic simple curve through $y_i$, chosen so that, for every choice of $y^*_i$ on $P_i$, $0$ lies in the interior of $\text{conv}\{f(y_0), f(y^*_1), ..., f(y^*_p)\}$. (Note that $y_0$ plays a special role here.) Let $\hat{f}_i$ denote the function $f \circ P_i$, so that $\hat{f}_i$ is the directional derivative of $f$ along $P_i$, and let $y_i = P_i(t_i)$.

We wish to show that $y_i$ and $P_i$ (for $i = 1, ..., p$) can be chosen to make the Jacobian matrix

$$J = \left( \hat{f}'_1(t_1), ..., \hat{f}'_p(t_p) \right)$$

nonsingular. If not, assume the $P_i$ and $y_i$ are chosen to maximize the largest size of a minimal dependent subset of the $\hat{f}'_i(t_i)$, say, $\{\hat{f}'_1(t_1), ..., \hat{f}'_q(t_q)\}$. Then $\hat{f}'_i(t^*_i)$ depends on $\hat{f}'_1(t_2), ..., \hat{f}'_q(t_q)$ for all $t^*_i$ near $t_i$. We conclude that the tangent vectors to $\text{Im } f$ near $y_i$ all satisfy a linear equation. It follows by analyticity of $f$ and connectivity of $\Omega$ that $\text{Im } f$ itself satisfies a linear equation, contrary to the assumption that it affinely spans $\mathbb{R}^p$. Thus $J$ can be made nonsingular, and by shrinking the curves if necessary we can assume $J$ remains nonsingular for all choices of $y^*_i$ on $P_i$.

For each choice of $t^* = (t^*_1, ..., t^*_p)$ in $(0, 1)^p$ there is a unique real vector $a^* = (a^*_1, ..., a^*_p) = A(t^*)$ such that

$$F(t^*, a^*) = \sum_{i=1}^{p} a^*_i [\hat{f}'_i(t^*_i) - \hat{f}'_0(t_0)] + \hat{f}'_0(t_0) = 0.$$ 

Moreover $a^*_1, ..., a^*_p$ and $a^*_0 = 1 - \sum_{i=1}^{p} a^*_i$ are all positive. Let $a = A(t)$. By the implicit function theorem, because the Jacobian matrix $D_t F(T, a) = (a_1 \hat{f}'_1(t_1), ..., a_p \hat{f}'_p(t_p))$ is nonsingular, $A$ is invertible in a neighborhood of $t$. Let $\beta$ be a rational point in that neighborhood and $s = (s_1, ..., s_p) = A^{-1}(\beta)$. Multiplying through by a common denominator of $\beta_1, ..., \beta_p$, and setting $x_i = P_i(s_i)$, we have

$$n_0 f(y_0) + n_1 f(x_1) + \cdots + n_p f(x_p) = 0,$$
where the \( n_i \) are positive integers. Thus \( \{ y_0, x_1, \ldots, x_p \} \) is an averaging multiset for \( f \) having \( p + 1 \) distinct points.

To make this into an averaging set we split the \( x_i \) that appear multiply. More generally, for use later in the \( C^1 \) case, suppose we have a finite combination

\[
\sum_{i=1}^{p} n_i f_i(s_i) + \sum_{j=0}^{M} m_j f(z_j) = 0, \quad (*)
\]

where \( M \geq -1 \), the \( n_i \) and \( m_j \) are positive integers, all the values \( f_i(s_i) \) and \( f(z_j) \) are distinct, and

\[
J = (f_1(s_1^*), \ldots, f_p(s_p^*))
\]

is nonsingular. (In the present, analytic case we begin with \( M = 0 \), \( z_0 = y_0 \), \( m_0 = n_0 \).) We split the multiple values by performing two slightly different operations.

If \( f(z_j) \) has multiplicity \( m_j \geq 2 \), we set

\[
F(s^*) = \sum_{i=1}^{p} f_i(s_i^*) \quad \text{and} \quad a = f(z_j) + F(s).
\]

The Jacobian matrix of \( F \) at \( s \) is \( J \), which is nonsingular. Thus choosing \( z_i^* \) so \( f(z_i^*) \) is near but different from \( f(z_i) \), by the inverse function theorem there exists \( s^* \) for which

\[
F(s^*) = a - f(z_j^*)
\]

and replacing one copy each of \( f(z_j), f_1(s_i^*), \ldots, f_p(s_p^*) \), respectively, by \( f(z_j^*) \), \( f_1(s_i^*), \ldots, f_p(s_p^*) \) introduces no new coincidences of values. Performing this replacement then reduces the amount of multiplicity. We repeat this process until all \( m_j = 1 \).

Now the only multiple values remaining, if any, are the \( f_i(s_i) \). Say \( n_i \geq 2 \). Let \( F \) be as before and

\[
a = f_1(s_1) + F(s).
\]

Choosing \( r_1 \) so \( f_1(r_1) \) is near but not equal to \( f_1(s_1^*) \), there exists \( s^* \) such that \( F(s^*) = a - f_1(r_1) \) and there are no new coincident values if we replace one copy of \( f_i(s_i) \) by \( f_i(s_i^*) \) for \( i \geq 2 \) and two copies of \( f_1(s_1) \) by one each of \( f_1(r_1) \) and \( f_1(s_1^*) \). If we can assure \( f_1(r_1) \neq f_1(s_1^*) \), then we have reduced the amount of multiplicity. But if it were impossible to choose \( r_1 \) to separate \( f_1(r_1) \) from \( f_1(s_1^*) \), we would have

\[
G(s^*) = 2f_1(s_1^*) + f_2(s_2^*) + \cdots + f_p(s_p^*) = \text{constant}
\]
for all $s^*_i$ in a small portion of $(0, 1)$ around $s_1$. But $G$ has nonsingular Jacobian matrix at $s$ and consequently is locally one-to-one. This contradiction shows that we can be sure $\hat{f}'_1(r_1)$ and $\hat{f}'_1(s^*_1)$ are distinct. Hence it is possible to eliminate all multiple values. This completes the proof of the Basic Theorem for analytic functions.

We adapt this proof to continuously differentiable functions by separating the two roles of $y_1, \ldots, y_p$, as vertices of a simplex containing the origin and as points at which tangent vectors to Im$f$ form a basis of $\mathbb{R}^p$.

Let $\Omega$ be a differentiable manifold and $f$ be of class $C^1$. Choose $x_1, \ldots, x_p \in \Omega$ and simple curves $P_i$ of class $C^1$ containing the $x_i$, so that the $p$ directional derivatives of $f$ along $P_i$ at $x_i$ are linearly independent. To put this another way, let $\hat{f}_i = f \circ P_i$ and $x_i = f_i(s_i)$ for $s_i \in (0, 1)$; we are choosing $x_i$ and $P_i$ so the Jacobian matrix

$$J = (\hat{f}'_1(s_1), \ldots, \hat{f}'_p(s_p))$$

is nonsingular. This is possible because Im$f$ satisfies no linear equation. (It may be that some of the $y_j$ happen to coincide with some $x_i$; in fact one can always make one $y_j$ and $x_i$ be the same.) By reducing the size of the $P_i$ we can guarantee that the Jacobian remains nonsingular for all choices of $x_i^*$ on $P_i$.

Now we take a large positive number $q$, so large that

$$-\frac{1}{q} \sum \hat{f}'_i(s_i) \in \text{int conv}\{f(y_0), f(y_1), \ldots, f(y_p)\}.$$

Thus

$$\sum \hat{f}'_i(s_i) = -\sum_{j=0}^{p} (q\alpha_j) f(y_j),$$

where $\sum \alpha_j = 1$ and all $\alpha_j > 0$. Since the Jacobian of the left side is nonsingular, we can change the $qa_j$ to nearby rational $\beta_j$ for which $\sum \hat{f}'_i(s_i^*) = -\sum \beta_j f(y_j)$ has a solution. Then clearing denominators we have an averaging multiset consisting of the $2p + 1$ points $y_0, y_1, \ldots, y_{2p}$, $P_i(s_i^*), \ldots, P_p(s_p^*)$. Let us drop the asterisks, so the averaging multiset consists of $y_0, y_1, \ldots, y_p$, $x_1 = P_1(s_1), \ldots, x_p = P_p(s_p)$.

To split the points we consider the general situation (*), beginning now with $M = p$, $z_j = y_j$ for $j = 0, 1, \ldots, p$, and $n_i$ and $m_j$ the multiplicities of the $x_i$ and $y_j$ in the averaging multiset. The argument is henceforth the same as in the analytic case.

To prove that all sufficiently large sizes $n$ are attainable, note that when rationalizing the coefficients $\alpha_i$ (in the analytic case) or $q\alpha_j$ (in the $C^1$ case),
we can choose the $\beta$ values to have any large enough common denominator $N$. In the analytic case, $n = N$. In the differentiable case we have $n \approx pN + q$; by adjusting the $\beta_i$ slightly we can get all large values of $n$.

5. Special Proof in the Continuous Case

In this section we prove the Basic Theorem for an arbitrary continuous function $f$ defined on a domain that is slightly restricted. We assume that for any $p + 1$ distinct points in $\Omega$ there is a simple curve containing them. Manifolds, for instance, satisfy this condition. The proof here does not establish the existence of an averaging set $X$ with distinct values $f(x)$ for $x \in X$.

If we prove the theorem for domain $(0, 1)$, then we have it for any $\Omega$ satisfying the assumption. For we choose a simple curve $P$ containing $y_0, y_1, \ldots, y_p$, and apply the theorem to $\Omega^* = (0, 1), f^* = f \circ P$. Then we have an averaging set for $f$ contained in $P$. Henceforth we consider only the case $\Omega = (0, 1)$ and we number the $y_i$ so that

$$0 < y_0 < y_1 < \cdots < y_p < 1.$$

We need a substitute for the inverse function theorem, which of course is inapplicable to nondifferentiable functions. This substitute is the first lemma; the second is an application that will be particularly useful in the general proof of Section 6 and that illustrates relatively simply the way we use the first lemma in this section.

**Lemma 5.1.** Let $p: \mathbb{R}^p \to \mathbb{R}^p$ be a continuous function that differs by a bounded amount from a nonsingular linear transformation $L$, that is, $\sup \|p(x) - L(x)\| < \infty$. Then $p$ is surjective.

**Proof.** We can assume $L$ is the identity, hence $\|p(x) - x\| \leq M$ for some quantity $M$. The mapping

$$\psi: x \to x/(1 + \|x\|)$$

is a homeomorphism of $\mathbb{R}^p$ with the open unit ball $B$. The composition

$$\psi \circ p \circ \psi^{-1}: B \to B$$

is continuous and extends continuously to the identity map on the boundary of $B$. The extension is thus a continuous self-map of the closed unit ball leaving the boundary pointwise fixed. It is well known that such a map is surjective (because $B$ does not retract onto its boundary [5, (15.6)]). Since $\psi$ is a bijection, the lemma follows.
LEMA 5.2. Let $g_1, \ldots, g_p$ be paths in $\mathbb{R}^p$ whose principal chords are linearly independent. Then the image of the function $G: [0, 1]^p \to \mathbb{R}^p$ defined by

$$G(x_1, x_2, \ldots, x_p) = g_1(x_1) + g_2(x_2) + \cdots + g_p(x_p)$$

has nonempty interior.

Proof. We extend $g_i$ continuously to $\mathbb{R}$ by

$$g_i(v + \lambda) = vAg_i + g_i(\lambda),$$

where $v \in \mathbb{Z}$ and $\lambda \in [0, 1)$. Geometrically this means we are piecing together translates of $g_i([0, 1])$, arranged so the $v$th copy begins where the $(v-1)$st copy ends. The function $\varphi: \mathbb{R}^p \to \mathbb{R}^p$ defined by $\varphi(x_1, \ldots, x_p) = \sum g_i(x_i)$ differs by at most

$$M = 2 \sum_{i=1}^p \max_{[0,1]} \| g_i(t) \|$$

from the linear transformation $L(x_1, \ldots, x_p) = \sum x_i Ag_i$. Thus by Lemma 5.1, $\varphi$ is surjective.

Since $\varphi - L$ is bounded by $M$, the box $J(0, 0, \ldots, 0)$, where

$$J(v_1, v_2, \ldots, v_p) \equiv \left\{ \sum (v_i + \lambda_i) Ag_i : \lambda_i \in [0, 1] \right\},$$

lies in the union of finitely many images $\varphi(J(v_1, v_2, \ldots, v_p))$. Therefore at least one of these images must have nonempty interior. It follows that $J(0, 0, \ldots, 0) = \text{Im } G$ has nonempty interior.

Proof of the theorem. Let $U_i$ be an interval around $y_i$, for $i = 0, 1, \ldots, p$, such that $0 \in \text{int conv}\{f(z_0), f(z_1), \ldots, f(z_p)\}$ for all choices of $z_i \in U_i$. Let $y_i^0, y_i^1, y_i^2, \ldots$ be a sequence in $U_i$ descending to $y_i$ for which all $f(y_i^j)$ are distinct. We can choose the sequence to make $\| f(y_i^j) - f(y_i) \|$ as fast as we like, say, faster than $2^{-j}$.

Using these sequences we construct functions $\varphi_i, \tilde{\varphi}_i: \mathbb{R} \to \mathbb{R}^p$ for $i = 1, \ldots, p$. If $x \in \mathbb{R}$, we write $v$ and $\lambda = x - v$ for the integral and fractional parts of $x$. We set

$$\tilde{\varphi}_i(x) = v[f(y_i^j) - f(y_{i-1})] + [f((1 - \lambda) y_{i-1} + \lambda y_i) - f(y_{i-1})]$$

and we define $\varphi_i(x)$ by the same formula if $x < 0$, by

$$\varphi_i(x) = \sum_{j=0}^{v-1} [f(y_i^j) - f(y_{i-1}^j)] + [f((1 - \lambda) y_{i-1}^v + \lambda y_i^v) - f(y_{i-1}^v)].$$
if \( x \geq 0 \). These functions are continuous. Geometrically \( \phi_t \) is obtained by splicing together copies of \( f([y_{i-1}, y_i]) \) as in the proof of Lemma 5.2. If we replace the \( K \)th copy by \( f([y_{i-1}^\kappa, y_i^\kappa]) \) for \( \kappa \geq 0 \), we have \( \phi_i \).

Now let \( \phi, \hat{\phi} : \mathbb{R}^p \to \mathbb{R}^p \) be defined by

\[
\phi(x) = \sum \phi_i(x_i) \quad \text{and} \quad \hat{\phi}(x) = \sum \hat{\phi}_i(x_i).
\]

Then \( \hat{\phi} \) differs by a bounded amount from the nonsingular linear transformation

\[
L(x) = \sum x_i |f(y_i) - f(y_{i-1})|.
\]

We wish to show that \( \|\phi(x) - \phi(x)\| \) is bounded. Let us write out in full the definitions of \( \phi \) and \( \hat{\phi} \). We get

\[
\hat{\phi}(x) = \sum_{i=1}^p (v_i - v_{i+1}) f(y_i) - v_1 f(y_0)
\]

\[+ \sum_{i=1}^p [f((1 - \lambda_i) y_{i-1} + \lambda_i y_i) - f(y_{i-1})]
\]

and

\[
\phi(x) = \sum_{i=1}^p \sum_{j=1}^{v_i-1} f(y_{i-1}^{j+1}) - \sum_{j=0}^{v_i-1} f(y_0)
\]

\[+ \sum_{i=1}^p [f((1 - \lambda_i) y_{i-1}^{v_i} + \lambda_i y_i^v) - f(y_{i-1}^{v_i+1})]
\]

with the convention \( v_{p+1} = 0 \). Thus

\[
\|\phi(x) - \hat{\phi}(x)\| \leq \sum_{i=1}^p \sum_{j=0}^\infty \|f(y_{i-1}^{j+1}) - f(y_i)\| + \sum_{j=0}^\infty \|f(y_0) - f(y_0)\|
\]

\[+ \sum_{i=1}^p \|f((1 - \lambda_i) y_{i-1}^{v_i} + \lambda_i y_i^v) - f((1 - \lambda_i) y_{i-1} + \lambda_i y_i)\|
\]

\[\leq p(\delta + 2),
\]

where

\[
\delta = \text{diameter of } f([y_0, y_p^0]).
\]

We conclude that \( \|\phi(x) - L(x)\| \) is bounded, whence \( \phi \) is surjective by Lemma 5.1.
Pick a positive integer \( n \). Because \( \varphi \) is surjective, there are values \( x_1, \ldots, x_p \) for which

\[
- \sum_{j=0}^{n-1} f(y_j) = \varphi(x_1, \ldots, x_p)
\]

\[
= \sum_{i=1}^{p} \sum_{j=\nu_i+1}^{\nu_i-1} f(y_j) - \sum_{j=0}^{\nu_1-1} f(y_j)
\]

\[
+ \sum_{i=1}^{p} [f((1 - \lambda_i) y_{i-1}^{\nu_i} + \lambda_i y_i^{\nu_i}) - f(y_{i-1}^{\nu_i})].
\]

Thus with the convention \( v_0 = n \) we have

\[
0 = \sum_{i=0}^{p} \sum_{j=\nu_i+1}^{\nu_i-1} f(y_j) + \sum_{i=1}^{p} [f((1 - \lambda_i) y_{i-1}^{\nu_i} + \lambda_i y_i^{\nu_i}) - f(y_{i-1}^{\nu_i})].
\]

It follows that we have an averaging set \( X \) consisting of the \( n \) points

\[
\begin{align*}
&y_0^{n-1}, y_0^{n-2}, \ldots, y_0^{n+1}, (1 - \lambda_1) y_0^{\nu_1} + \lambda_1 y_1^{\nu_1}, \\
y_1^{\nu_1-1}, y_1^{\nu_1-2}, \ldots, y_1^{\nu_1+1}, (1 - \lambda_2) y_1^{\nu_2} + \lambda_2 y_2^{\nu_2}, \\
y_p^{\nu_p-1}, y_p^{\nu_p-2}, \ldots, y_p^{\nu_p+1}, (1 - \lambda_p) y_p^{\nu_p} + \lambda_p y_p^{\nu_p},
\end{align*}
\]

all distinct (since they are in ascending order as listed), provided that

\[
v_0 = n > v_1 > v_2 > \cdots > v_p > 0 = v_{p+1}.
\]

To prove (*) we in effect compare \( \varphi(x) \) and \( \tilde{\varphi}(x) \). In outline, since \( \varphi \) and \( \tilde{\varphi} \) are not very much different, \( \tilde{\varphi}(x)/n \) is near 0 (for large \( n \)); consequently,

\[
w \equiv \sum_{i=0}^{p} \frac{v_i - v_{i+1}}{n} f(y_i) \approx 0.
\]

But \( \sum (v_i - v_{i+1})/n = 1 \), so we have an affine combination of the \( f(y_i) \) lying inside

\[
A = \text{conv}\{f(y_0), f(y_1), \ldots, f(y_p)\}.
\]

This implies \( (v_i - v_{i+1})/b > 0 \), which is (*). The detailed proof is simplified by a more direct calculation. We have
Therefore \( \|w\| < \frac{p(\delta + 2)}{n} \). Letting

\[
\rho = \text{dist}(0, \text{bdy } A),
\]

we conclude that \( w \in \text{int } A \) if \( n > \frac{p(\delta + 2)}{\rho} \). Thus averaging sets \( X \) exist of all sizes

\[
n > \frac{(p(\delta + 2) + 2)}{\rho}.
\]

(5.1)

6. General Proof of the Continuous Case

In this section we treat continuous functions on an arbitrary path-connected space. We may as well simplify the notation by assuming \( \Omega \subseteq \mathbb{R}^p \) and \( f = \text{the identity} \). We thus begin with

\[
Y = \{ y_0, y_1, \ldots, y_p \} \subseteq \Omega \subseteq \mathbb{R}^p
\]

for which \( 0 \in \text{int conv } Y \).

We will need sets

\[
P, Q_1, \ldots, Q_p \subseteq \Omega
\]

such that \( \sum Q_i \) contains an open ball \( B \), \( P + \cdots + P \) \( (p \text{ times}) \) contains an open ball \( A \), and the \( Q_i \) are pairwise disjoint from each other and from \( P \cup Y \). We get these sets by letting \( P = P_1 \cup \cdots \cup P_p \) and taking \( P_1, \ldots, P_p, Q_1, \ldots, Q_p \) to be paths in \( \Omega \) chosen so that the principal chords \( \Delta P_1, \ldots, \Delta P_p \) are linearly independent, as are \( \Delta Q_1, \ldots, \Delta Q_p \). Then the existence of \( A \) and \( B \) follows from Lemma 5.2.

We have to show that it is possible to obtain such paths \( P_i \) and \( Q_i \). For this purpose we define

\[
w(u) = \sup \{(y - x) \cdot u : x, y \in \Omega\},
\]

the width of \( \Omega \) in the direction \( u \) if \( \|u\| = 1 \), and

\[
w = \inf \{w(u) : u \in \mathbb{R}^p, \|u\| = 1\},
\]

AVERAGING SETS
the minimum width of $\Omega$. Since $\Omega$ does not lie in any hyperplane, all $w(u) > 0$, whence $w > 0$. (It is possible that $w = \infty$.) Let $w'$ be a positive number less than $w/2p$. For the $P_i$ and $Q_i$, we consider paths in $\Omega$ with diameter $\leq w'$. Suppose we have found $2k < 2p$ such paths, $P_1, \ldots, P_k$ and $Q_1, \ldots, Q_k$, whose principal chords span subspaces $S$ and $T$, respectively. Since the total width of the components of

$$R = P_1 \cup \cdots \cup P_k \cup Q_1 \cup \cdots \cup Q_k$$

in a direction $u \in S$ is at most $kw' < w(u)$, there is a path in $\Omega \setminus (Y \cup R)$ whose principal chord has nonzero $u$-component. By subdividing this path we can get a path $P_{k+1}$ of diameter $\leq w'$ whose principal chord is independent of $\Delta P_1, \ldots, \Delta P_k$. Now the total width of the components of $R \cup P_{k+1}$ in a direction $v \in T$ is at most $(2k + 1)w' < w(v)$, hence in $\Omega \setminus (Y \cup R \cup P_{k+1})$ we can find a path $Q_{k+1}$ of diameter $< w'$ for which $\Delta Q_{k+1}$ is independent of $\Delta Q_1, \ldots, \Delta Q_k$. By induction on $k$ this construction yields $2p$ pairwise disjoint paths $P_1, \ldots, P_p, Q_1, \ldots, Q_p$ with principal chords as required.

Let the center of $A$ be $z_1 + \cdots + z_p$, where $z_i \in P$, and let

$$A_{\mu p + k} = \mu A + z_1 + \cdots + z_k$$

for $0 \leq k < p, \mu \geq 1$, where $\mu A = \{\mu x : x \in A\}$, so that $A_{m} \subseteq P + \cdots + P$ ($m$ times). The radius of $A_{m}$ increases roughly linearly with $m$ while the distance between successive centers does not increase. Consequently the overlap of successive balls becomes large and there is an open cone $C$ with vertex $0$ such that all but a finite part of $C$ lies in $\bigcup_{m}^{\infty} A_{m}$. Let $r$ be a positive number large enough that

$$C_r = \{x \in C : \|x\| > r\} \subseteq \bigcup_{m}^{\infty} A_{m}.$$

Let $v_0 = \sum y_i \in C$ be a rational convex combination of $y_0, y_1, \ldots, y_p$. By taking a sufficiently large integral multiple of $v_0$ we get a vector

$$v = \sum_{i=1}^{p} q_i y_i \in -C_r$$

with positive integral coefficients $q_i$, such that

$$v + B \subseteq -C_r. \quad (*)$$

Choose any $z_1' + \cdots + z_p' \in B$, where $z_i' \in Q_i$. From $(*)$ we see that

$$v + \sum_{i=1}^{p} z_i' = -\sum_{j=1}^{m} x_j$$
for some $x_1, \ldots, x_m \in P$. We rewrite this as

$$\sum_{i=0}^{p} q_i y_i + \sum_{j=1}^{m} x_j = -\sum_{i=1}^{p} z_i' \in -B.$$  

We want to split apart the $q_i$ copies of $y_i$ and any repetitions among the $x_j$. To do this we simply take $q_i$ points in $\Omega$ very near $y_i$ and enough points very near each multiple $x_j$ to eliminate the repetition. Let $y_i^*, \ldots, y_q^*$ (where $q = \sum q_i$) and $x_1^*, \ldots, x_m^*$ be the new points. Because they are close to the old points, $\sum y_i^* + \sum x_j^*$ is still inside $-B$. So there are points $z_i^* \in Q_i$ for which

$$\sum_{i=1}^{q} y_i^* + \sum_{j=1}^{m} x_j^* = -(z_1^* + \cdots + z_p^*).$$

By the construction of the $Q_i$ all the $y_i^*, x_j^*$, and $z_i^*$ are distinct (if we picked the $y_i^*$ and $x_j^*$ near enough to the $y_i$ and $x_j$). Therefore we have an averaging set of size

$$n = q + m + p.$$  

We have proved the basic existence result. To show that all large sizes $n$ are attainable we need a refinement of the final step. From (*) and $v \in -C$ it follows that

$$Nv + B \subseteq -C_r \quad \text{for all } N \geq 1.$$  

Let $l$ be such that $v \in A_l$. If $M$ is great enough, any ball of fixed radius in $C_r \cap A_M$ will belong to at least $q + l$ consecutive $A_m$ including $A_M$. Let $N_1$ be so large that

$$-(N_1 v + B) \subseteq A_{M_1 + m} \quad \text{for } m = 1, 2, \ldots, q + l,$$

for some positive integer $M_1$. Suppose

$$-(Nv + B) \subseteq A_{M_m} \quad \text{for } m = 1, 2, \ldots, q + l.$$  

Then

$$-([N+1] v + B) \subseteq -v + A_{M_m} \subseteq A_{M_1 + l + m}$$  

for the same $m$. By induction on $N$ starting at $N_1$ we deduce that

$$-(Nv + B) \subseteq A_{M_1+(N-N_1)l+m}.$$  

Thus we have expressions

$$Nv + \sum_{i=1}^{M_1+(N-N_1)l+m} y_i + \sum_{i=1}^{M_1+(N-N_1)l+m} z_i = 0.$$  


for all \( m = 1, 2, \ldots, q + l \). By the splitting process we described above each such expression leads to an averaging set of size

\[
n = Nq + p + M_1 + (N - N_1)l + m = p + M_1 + N_1q + (N - N_1)(q + l) + m.
\]

Thus all sizes

\[
n > p + M_1 + N_1q
\]

are attainable.

7. CONTINUITY, UNIFORMITY, AND BOUNDS

We would like to have not only an averaging set \( X \) for each function \( f \) but also a continuous dependence of \( X \) of \( f \). In this section we discuss ways in which \( X \) can and cannot vary with \( f \).

Let \( C(\Omega, \mathbb{R}^p) \) be the class of continuous functions from \( \Omega \) into \( \mathbb{R}^p \) and let

\[
\mathcal{C}(\Omega, \mu, \mathbb{R}^p) = \{ f \in C(\Omega, \mathbb{R}^p) : f \text{ is } \mu\text{-integrable} \}.
\]

We provide these function spaces with the "norm" (whose value can be infinite)

\[
\|f\| = \sup_{x \in \Omega} \|f(x)\|
\]

which defines the topology of uniform convergence. (A topology based on compact sets would be equally satisfactory.) An ideal result would be the existence of a number \( n \) and a (uniformly) continuous global mapping

\[
X : \mathcal{C}(\Omega, \mu, \mathbb{R}^p) \to \Omega^n
\]

such that \( X(f) \) is a \( \mu \)-averaging set for each \( f \). There are at least two reasons that cannot happen. For one thing it entails the existence of a size that works for every \( f \); which, as we show by example in Section 7.3, is impossible. Moreover it requires that for each \( f \) every nearby \( g \) has an averaging set \( X(g) \) near \( X(f) \), but if \( f \) has degenerate image, that is, if \( \text{aff}(\text{Im} f) \neq \mathbb{R}^p \), then no matter how we choose \( X(f) \) there is a sequence \( g_m \to f \) for which no averaging sets approach \( X(f) \) (see Section 7.4). Thus it is possible neither to define \( X \) globally, even excluding degenerate functions, nor locally near a degenerate function. So we should confine ourselves to the classes

\[
\mathcal{C}_\star(\Omega, \mathbb{R}^p) = \{ f \in C(\Omega, \mathbb{R}^p) : \text{aff}(\text{Im} f) = \mathbb{R}^p \}
\]
and $\mathcal{F}_*(\Omega, \mu, \mathbb{R}^p)$ of functions with nondegenerate images and we should seek only local continuity of $X$.

**Conjecture 1 (Local Continuity and Uniformity).** Given $\Omega$, $\mu$, and $p$, there are an open covering $\mathcal{U}$ of $\mathcal{F}_*(\Omega, \mu, \mathbb{R}^p)$ and for each $U \in \mathcal{U}$ a natural number $n(U)$ and a continuous mapping $X_U: U \to \Omega^{n(U)}$ such that for any $f \in U$, $X_U(f)$ is an averaging set for $f$ with respect to $\mu$. And $X_U$ can be made uniformly continuous if $\Omega$ is a metric space.

In view of the fact that we can eliminate $\mu$ in favor of an arbitrarily specified centroid (Section 3), we can reformulate Conjecture 1. For $A \subseteq \mathbb{R}^p$, let

$$\mathcal{F}_*(\Omega, \mathbb{R}^p; A) = \{(f, c) \in \mathcal{F}(\Omega, \mathbb{R}^p) \times A : c \in \text{int conv}(\text{Im } f)\}$$

with a product "norm" like $\| (f, c) \| = \max(\| f \|, \| c \|)$. (We are interested in $A = \mathbb{R}^p$ and $A = \{0\}$. Note that the $f$ involved necessarily have nondegenerate image.) Evidently $\mathcal{F}_*(\Omega, \mu, \mathbb{R}^p)$ can be regarded as a topological subspace of $\mathcal{F}_*(\Omega, \mathbb{R}^p; \mathbb{R}^p)$. So the next conjecture is an extension of Conjecture 1; in fact it is essentially equivalent to (uniformly) continuous dependence of $X$ on $f$ and $\mu$.

**Conjecture 2 (Local Continuity and Uniformity).** Given $\Omega$ and $p$, there are an open cover $\mathcal{U}$ of $\mathcal{F}_*(\Omega, \mathbb{R}^p; \mathbb{R}^p)$ and for each $U \in \mathcal{U}$ a natural number $n(U)$ and a continuous mapping $X_U: U \to \Omega^{n(U)}$ such that for any $(f, c) \in U$, $X_U(f, c)$ has average $f$-value equal to $c$. And $X_U$ can be made uniformly continuous if $\Omega$ is a metric space.

It is clearly sufficient to prove Conjecture 2 just for $\mathcal{F}_*(\Omega, \mathbb{R}^p; \{0\})$.

We do not have a proof of local continuity in general but we do have as supporting evidence a proof for differentiable functions—with a sort of uniformity that does not require a metric domain—and the observation that a local bound $n(U)$ on size exists for quite general $\Omega$.

Notice that we did not conjecture that every averaging set for $f$ can be made to vary continuously in a neighborhood of $f$. In Section 7.5 we show this is impossible even for the case $p = 1$, the ordinary mean value theorem. In general we suspect that the existence and nature of the smallest averaging sets may depend strongly on the particular properties of $f$, especially when $\Omega$ and $f$ are highly structured as in the case of spherical designs; and only for relatively large averaging sets can one expect continuous variation with $f$.

7.1. **Local Continuity and Uniformity in the Derivative Norm**

We let $\Omega$ be a differentiable manifold and $\mathcal{C}^1$ be the class of continuously differentiable members of $\mathcal{F}$ for all the various $\mathcal{F}$ we have mentioned, such as $\mathcal{F}(\Omega, \mu, \mathbb{R}^p)$ and $\mathcal{F}_*(\Omega, \mathbb{R}^p; \mathbb{R}^p)$. We put on $\mathcal{C}^1(\Omega, \mathbb{R}^p)$ the topology of the first derivative "norm,"

$$\| f \|_1 = \max(\| f \|, \| Df \|),$$
where $Df$ is the derivative of $f$. For an explanation of these definitions in general we refer to [1]. A remark on uniform continuity: Since $\Omega$ is a manifold one can use local coordinates to define local uniform continuity even if $\Omega$ is not metrized, and this is what we need.

**Theorem (C^1 Local Continuity and Uniformity).** Given $\Omega, \mu$, and $p$, there is an open cover $\mathcal{U}$ of $C^1_b(\Omega, \mu; \mathbb{R}^p)$ with the following properties. For each $U \in \mathcal{U}$ there are a natural number $n(U)$ and a mapping $X_U : U \rightarrow \Omega^{n(U)}$, of class $C^1$ and uniformly continuous, such that for every $f \in U$, $X_U(f)$ is an averaging set for $f$. Moreover the $X_U$ can be chosen so that only $p$ of the $n(U)$ members of $X_U(f)$ vary with $f$.

More generally given $\Omega$ and $p$, there is an open cover $\mathcal{U}$ of $C^1_b(\Omega, \mathbb{R}^p; \mathbb{R}^p)$ having $n(U)$ and $X_U$ as above, such that $X_U(f, c)$ is an averaging set for each $(f, c) \in U$.

**Proof.** It suffices to show that any $(f, 0) \in C^1_b(\Omega, \mathbb{R}^p; \{0\})$ has a neighborhood $U$ on which $X_U$ can be defined. For this purpose we require the implicit function theorem on Banach spaces. For a proof we refer to [1, Theorem 20.1].

We begin with the conclusion of the proof in Section 4. There we obtained an averaging set $X = (x_1, \ldots, x_n)$ for $f$ with all $f(x_i)$ distinct, having the properties that (after suitable renumbering) $x_1, \ldots, x_p$ lie in simple curves $P_1, \ldots, P_p$ of class $C^1$, say, $x_i = P_i(s_i)$ for $i = 1, \ldots, p$, and the Jacobian matrix $J(f, s) = (\dot{f}_i(s_1), \ldots, \dot{f}_i(s_p))$ is nonsingular, where $\dot{f}_i = f \circ P_i$. Let

$$F(g, t) = \sum_{i=1}^p g \circ P_i(t_i) + \sum_{j=1}^n g(x_j)$$

for $g \in C^1(\Omega, \mathbb{R}^p)$ and $t \in (0, 1)^p$. We see that $\partial F/\partial t_i = \dot{g}_i(t_i)$, where $\dot{g}_i = g \circ P_i$, so the derivative $D_t F(g, t)$ of $F$ with respect to $t$ is $J(g, t)$. Thus $F(f, s) = 0$ and $D_t F(f, s)$ is nonsingular. These are the hypotheses of the implicit function theorem—except for one problem: $C^1_b(\Omega, \mathbb{R}^p)$ is not a Banach space if $\Omega$ is noncompact.

In order to get a Banach space we first modify the curves $P_i : (0, 1) \rightarrow \Omega$. There is a closed interval $[a, b] \subseteq (0, 1)$ such that all $s_i \in (a, b)$. We have

$$X = \tilde{X} \equiv \bigcup_{i=1}^p P_i((a, b)) \cup \{x_{p+1}, \ldots, x_n\}.$$

Letting $B$ be the Banach space of $C^1$-bounded functions on $\tilde{\Omega}$, then

$$B \supset \{g | \tilde{\Omega} : g \in C^1_b(\Omega, \mathbb{R}^p)\}.$$
Since $F(g, t)$ depends only on $\hat{g} = g|\tilde{\Omega}$, we can regard $F$ as a function $\tilde{F}$ from $B \times (a, b)^p$ into $\mathbb{R}^p$. As such it is of class $C^1$ (because the evaluation map $e(g, y) = g(y)$ is, by [1, Theorem 10.3]). Thus by the implicit function theorem there are a neighborhood $U_0 \subseteq B$ of $\tilde{f} = f|\tilde{\Omega}$ and a function $T: U_0 \rightarrow (a, b)^p$ of class $C^1$ such that $\tilde{F}(g, T(g)) = 0$ for $g \in U_0$. Let

$$U_1 = \{ g \in C^1(\Omega, \mathbb{R}): g|\tilde{\Omega} \in U_0 \}.$$

For $g \in U_1$ we define

$$X_1(g) = (x_1(g), ..., x_p(g), x_{p+1}, ..., x_n),$$

where $x_i(g) = P_i(T_i(\hat{g}))$. We know that $X_1(g)$ is an averaging multiset; by shrinking the neighborhood $U_1$ to a smaller one $U_2$ we can make all its elements distinct. If we take $X_2 = X_1|U_2$, we have an averaging set varying in a continuously differentiable manner on a neighborhood $U_2$ of $f$.

The next step is to prove that $T$ is uniformly continuous. If $\|DT(g)\|$ is bounded on a neighborhood of $\tilde{f}$, then $T$ is uniformly continuous on that neighborhood. That is a consequence of the simplest form of Taylor's theorem for Banach spaces [1, p. 61]. Since $T$ is a $C^1$ function, we have

$$T(g + h) = T(g) + DT(g) h + \rho(g, h),$$

where $\|\rho(g, h)\|/\|h\|_1 \to 0$ as $(g, h) \to (f, 0)$. If we take $g$ sufficiently near $f$ and $h$ sufficiently near $0$, we get $\|\rho(g, h)\| < M \|h\|_1$, where $M$ is some positive number. Thus

$$\|T(g + h) - T(g)\| \leq (\|DT(g)\| + M) \|h\|_1.$$

So we have local uniform continuity of $T$ near $f$ if we can bound $\|DT(g)\|$ for $g$ near $\tilde{f}$.

To get this bound we write $\tilde{F}$ in terms of the evaluation map,

$$\tilde{F}(g, t) = \sum_{i=1}^{p} e(g, P_i(t_i)) + \sum_{j=p+1}^{n} e(g, x_j).$$

Thus $D_1\tilde{F}(g, t) = \tilde{F}(\cdot, t)$. Consequently, by the fact that $\tilde{F}(g, T(g)) = 0$ and the chain rule, we have

$$0 - D_1\tilde{F}(g, T(g)) = \tilde{F}(\cdot, T(g)) + J(g, T(g)) DT(g).$$

We can solve for $DT(g)$; we get

$$DT(g) = -J(g, T(g))^{-1} \tilde{F}(\cdot, T(g)).$$
So the operator norm of $DT(g)$ is

$$\| DT(g) \| = \sup_{h \neq 0} \frac{\| DT(g) h \|}{\| h \|} \leq \| J(g, T(g))^{-1} \| \cdot \| \tilde{F}(\cdot, T(g)) \| \leq n \| J(g, T(g))^{-1} \|,$$

since the operator norm of $\tilde{F}$ is $\leq n$.

Now we need a positive lower bound for $|\det J(g, T(g))|$ near $f$. We know that $J(g, t)$ is a continuous function of $g$ and $t$ because we are using the $C^1$ norm. Hence there are $\delta_1, \delta_2 > 0$ such that

$$|\det J(g, t)| > \frac{1}{2} |\det J(f, s)| \quad \text{if} \quad \| g - f \|_1 < \delta_1 \text{ and } \| t - s \| < \delta_2.$$  

Since $T$ is continuous, there is a positive $\delta \leq \delta_1$ such that $\| T(g) - s \| < \delta_2$ if $\| g - f \|_1 < \delta$. Thus

$$|\det J(g, T(g))| > \frac{1}{2} |\det J(f, s)| \quad \text{if} \quad \| g - f \|_1 < \delta.$$  

This establishes the bound on $\| J(g, T(g))^{-1} \|$ we needed to deduce uniform continuity of $T$ in a neighborhood of $f$.

If $\Omega$ is metrized, then $P_t[a, b]$ is a continuous function on a compact domain, hence is uniformly continuous. Therefore the composite functions $x_t(g) = P_t \circ T_t(g)$ are uniformly continuous on a neighborhood $U \subseteq U_2$ of $f$. Thus $X_2 = X_2 \cap U$ is uniformly continuous. This concludes the proof.

### 7.2. Bounds on Size

The bound of Formula (5.1) on the necessary size of an averaging set has two interesting consequences. First of all we can improve it. Let $c(f)$ denote the centroid $\mu(\Omega)^{-1} \int f du$.

**Proposition 7.1 (Size Bound).** Suppose $\Omega$ is as in Section 5, for instance, a manifold with or without boundary, and $f: \Omega \to \mathbb{R}^p$ is continuous with nondegenerate image. Choose $Y = \{ y_0, y_1, \ldots, y_p \} \subseteq \Omega$ so that $c(f) \in \text{int conv} f(Y)$. Then there exists an averaging set for $f$ of each size

$$n > p \frac{\text{diam } f(Y)}{\text{dist}(c(f), \text{bdy conv } f(Y))}.$$  

Moreover there is a neighborhood of $f$ in the sup norm in which every function $g$ has an averaging set of each such size $n$.

**Proof.** The second part of the proposition follows from the first by the local continuity near $f$ of

$$H(g) = p \frac{\text{diam } g(Y)}{\text{dist}(c(g), \text{bdy conv } g(Y))}.$$
For the validity of the bound $H(f)$, first let

$$p = \text{dist}(c(f), \text{bdy conv} f(Y)).$$

According to (5.1) every $n$ such that

$$n > \frac{p}{\rho} \text{diam} f(\Omega) + \frac{p + 1}{\rho} 2$$

is the size of an averaging set. We showed this by choosing points $y^i_j$ such that

$$\sum_{i=0}^{\infty} \|f(y^i_j) - f(y_j)\| < 2.$$

We could just as well have made the sum less than an arbitrary positive number $\varepsilon$. And we could have replaced $\text{diam} f(\Omega)$ by

$$\text{diam}\{f(y_j), f(y^i_j) : 0 \leq j \leq p, i \geq 0\},$$

which we can make $< \varepsilon + \text{diam} f(Y)$. Thus every

$$n > \frac{p}{\rho} (\varepsilon + \text{diam} f(Y)) + \frac{p + 1}{\rho} \varepsilon$$

is the size of an averaging set. Letting $\varepsilon$ approach 0, we have the proposition.

From Proposition 7.1 we immediately deduce the second consequence of (5.1).

**COROLLARY 7.1 (SIZE LOCAL BOUND).** Let $\Omega$ be as in Section 5. There exist an open covering $\mathcal{U}$ of $C^*_\infty(\Omega, \mu, \mathbb{R}^p)$ and for each $U \in \mathcal{U}$ a natural number $n(U)$ such that each $f \in U$ has an averaging set of every size $n \geq n(U)$.

It seems much harder, although evidently desirable, to deduce a size local bound in full generality from our proof in Section 6. A crucial step would be to show that one can find balls $A$ and $B$ that work for all $g$ near $f$.

### 7.3. Global Nonuniformity

Here we prove that there cannot in general be any overall bound on the minimum size of an averaging set for functions in $C^*_\infty(\Omega, \mu, \mathbb{R}^p)$, indeed not even for polynomials on $(0, 1)$. Our construction is stated for $\Omega = (0, 1)$ with $\mu = \text{Lebesgue measure}$ and $p = 2$, but clearly it generalizes to all other cases.
EXAMPLE 7.1. Take any positive integer \( n \). Let \( 0 < \delta < 1/(n + 1) \). Choose \( f = (f_1, f_2) \) so that

\[
\int_0^1 f \, dx = (0, 0),
\]

\( f_1(x) < 0 \) for \( x > \delta \),

\( f_2(x) \geq n \) for \( x \leq \delta \),

\( \inf f_2 \geq -1 \).

Any averaging set \( X \) must include a point \( x_i \) at which \( f_1 \) is positive, hence \( x_i < \delta \). But then \( f_2(x_i) \geq n \) and must be balanced by at least \( n \) points at which \( f_1 \) is negative. So \( |X| > n \).

Note that it is possible to perform this construction with polynomials, hence to produce an example with analytic functions on a compact domain, \( \Omega = [0, 1] \). Obviously one can adapt \( f \) to make \( p \) be any desired dimension greater than 2. As for other domains \( \Omega \) besides \( (0, 1) \), any one that has a nonconstant \( f \) can be mapped onto \( [0, 1] \). So we have a result for all \( \Omega \) and \( p \geq 2 \).

PROPOSITION 7.2. If \( \Omega \) supports a nonconstant real-valued function and \( p \geq 2 \), then the minimum size of an averaging set for a nondegenerate continuous function \( f : \Omega \to \mathbb{R}^p \) is unbounded. The same holds for \( C^r \) functions if \( \Omega \) is a \( C^r \) manifold and for analytic functions if \( \Omega \) is an analytic manifold.

In some sense our construction seems to require a large derivative. We are not sure how to make this precise. It does not help to pass over to \( \mathcal{C}^1((0, 1), \mu, \mathbb{R}^p) \) with the \( C^1 \) norm, since in our example \( \|f\|_1 \) can be made arbitrarily small by choice of \( \epsilon \).

7.4. Functions with Degenerate Image

We show here that functions whose images are not full dimensional must be treated separately. The first example is analytic on \( (0, 1) \); the second generalizes readily to all domains \( \Omega \).

EXAMPLE 7.2. We set \( f^{(n)}_\epsilon(x) = (\epsilon f_1(x), \epsilon f_2(x), x) \), where \( f_1 \) and \( f_2 \) are as in Example 7.1. Recall that we can take \( f_1 \) and \( f_2 \) to be polynomials. Since we can choose \( f^{(n)}_\epsilon \) so that \( f^{(n)}_\epsilon \to f_0 = (0, 0, x) \) uniformly as \( \epsilon \to 0 \) and since each \( f^{(n)}_\epsilon \) requires more than \( n \) points in an averaging set, we see that \( f_0 \) is the uniform limit of polynomial functions with unbounded minimum size of averaging sets. Thus there is no local uniformity of size around \( f_0 \) even within the class of polynomials.
EXAMPLE 7.3. We let $\Omega = (0, 1)$ with Lebesgue measure and $p = 2$; we take $f_0(x) = (x, 0)$. Suppose $X_0 = \{x_1, x_2, ..., x_n\}$ is an averaging set for $f_0$. The centroid of $f_0$ of course lies on the line $\mathbb{R} \times \{0\}$ spanned by $\text{Im} f_0$. Now choose an interval $(t_0, t_1) \subseteq (0, 1)$ at positive distance $\delta$ from every point of $X_0$ and choose a $C^\infty$ function $g: (0, 1) \to \mathbb{R}$ such that $g \geq 0$, $g(x) = 0$ for $x \in (t_0, t_1)$, and $\sup g(x) = 1$. Let

$$f_\varepsilon(x) = (x, \varepsilon g(x)).$$

Then $f_\varepsilon \to f_0$ as $\varepsilon \to 0$, but since every averaging set for $f_\varepsilon$ must contain a point in $(t_0, t_1)$, there can be no continuous mapping

$$X: \{f_\varepsilon: 0 \leq \varepsilon \leq 1\} \to (0, 1)^n$$

such that $X(f_\varepsilon)$ is an averaging set or even an averaging multiset for $f_\varepsilon$ and $X(f_0) = X_0$.

Evidently this construction generalizes to any $f, \in \mathcal{E}(\Omega, \mu, \mathbb{R}^p)$ with degenerate image. And it applies even within the class of $C^\infty$ functions, although not analytic ones. So it is essential to exclude all functions with degenerate image from questions of continuity of averaging sets.

7.5. Averaging Sets That Do Not Vary Continuously

We give an example to show that not every averaging set for $f$ can be varied continuously in a neighborhood of $f$, not even when $f$ has nondegenerate image. The mean value theorem itself provides our example: the point at which the mean value occurs may not be continuously variable with $f$.

EXAMPLE 7.4. We take $\Omega = (0, 1)$ and $p = 1$. We want a family of functions $f_\varepsilon: (0, 1) \to \mathbb{R}$ indexed by $\varepsilon \in (-1, 1)$ such that $f_\varepsilon$ is strictly increasing on $(0, \frac{1}{4})$ and on $(\frac{3}{4}, 1)$, it has the constant value $\varepsilon$ on $[\frac{1}{4}, \frac{3}{4}]$, and its mean value is 0, and such that $f_\delta \to f_\varepsilon$ uniformly as $\delta \to \varepsilon$. This is easy to arrange with $C^\infty$ functions.

Now $f_0$ has many averaging sets of size one, indeed any $\{x\}$ where $x \in [\frac{1}{4}, \frac{3}{4}]$. But any other $f_\varepsilon$ has a unique one-point averaging set $\{x_\varepsilon\}$, and

$$x_\varepsilon \to \frac{1}{4}^- \quad \text{if} \quad \varepsilon \to 0^-,$$

$$x_\varepsilon \to \frac{3}{4}^+ \quad \text{if} \quad \varepsilon \to 0^+.$$

So no matter which one-point averaging set we choose for $f_0$, there is no way to vary it continuously on a neighborhood of $f_0$.

There are also many $n$-point averaging sets for $f_0$, with $n \geq 2$, that cannot
be varied continuously on a neighborhood of \( f_0 \), to wit any \( n \)-point subset of \([\frac{1}{2}, \frac{3}{2}]\) that does not include both endpoints.

We give this example because it holds good simultaneously for every \( n \). There are of course for each \( n \geq 1 \) polynomials \( f \) having an \( n \)-point averaging set that does not vary continuously near \( f \).

8. Averaging Multisets

In the course of our construction of averaging sets we found that there is always an averaging multiset with at most \( 2p + 1 \) points, \( p + 1 \) in the case of analytic functions. This was explicit for analytic functions and in the proof in Section 5; in the general case (Section 6) it follows from the fact that \( v \in -C_\mu \), whence \( \sum_0^p q_i y_i + \sum_1^p \mu x_i = 0 \) for some \( x_1, \ldots, x_p \in P \). It is easy to adapt the proof for continuously differentiable functions to reduce this number to \( 2p \). On the other hand we will demonstrate that \( 2p - 1 \) points are sometimes necessary. This leaves open the question of the minimum size of an averaging multiset sufficient for all functions: is it \( 2p - 1 \), \( 2p \), or \( 2p + 1 \)? (And for analytic functions can it be reduced to \( p \)?) The planar case \( p = 2 \) suggests that \( 2p - 1 \) may be the correct answer and that only special functions require even that many.

**Example 8.1.** This example requires \( 2p - 1 \) different points, with \( p \geq 2 \). The image of \( f \) is the subset \( A \) of \( \mathbb{R}^p \) consisting of \( p \) vertical "prongs"

\[
P_0 = \{ (0, 0, \ldots, 0, t) : t \in [0, 1] \},
\]

\[
P_i = \{ (0, \ldots, 0, 1, 0, \ldots, 0, t) : t \in [0, 1] \} \quad \text{for} \quad i = 1, \ldots, p - 1,
\]

and the \( p - 1 \) horizontal "spokes"

\[
S_i = \{ (0, \ldots, 0, t_i, 0, \ldots, 0, 0) : t_i \in [0, 1] \} \quad \text{for} \quad i = 1, \ldots, p - 1.
\]

We map \( \Omega \) (say, \( \Omega = (0, 1) \)) onto \( A \) in such a way that the centroid \( c = (c_1, \ldots, c_{p-1}, 1 - \delta) \), where \( \delta > 0 \) is small and \( c_1, \ldots, c_{p-1} \) are irrationals approximately equal to \( 1/p \). Having obtained this centroid we no longer need \( f \), so from now on we identify \( \Omega \) with \( \text{Im} f \subseteq \mathbb{R}^d \).

Let \( X \) be an averaging multiset for \( f \). For each \( i = 0, 1, \ldots, p - 1 \) there is an affine hyperplane through \( c \) that intersects \( A \) only in the \( i \)th prong, and that at a height \( x_i \) near 1, certainly above \( 1 - (p + 1)\delta \). The existence of these hyperplanes obliges \( X \) to contain a point \( y_i \) near the tip of each prong at height greater than \( 1 - (p + 1)\delta \), since no open halfspace with \( c \) in its boundary can contain \( X \). Because any rational combination of \( y_0, y_1, \ldots, y_{p-1} \) will differ from \( c \) by irrational quantities in each of the first \( p - 1 \) coor-
dinates, $X$ must also contain a point in every spoke to make up the difference. Hence $X$ has at least $2p - 1$ distinct points.

The planar case ($p = 2$). Consider continuous functions from $[0, 1]$ to $\mathbb{R}^2$; that is, paths in the plane. Not only can we prove that $2p - 1 = 3$ distinct points suffice to make an averaging multiset, we can classify the rather exceptional functions for which two points are insufficient. We can do the same for functions on any other domain $\Omega$; we omit the classification for general domains since it is more complicated but involves no new ideas.

**Proposition 8.1.** Let $f: \Omega \to \mathbb{R}^2$ be continuous and $c \in \text{relint conv}(\text{Im} f)$. Then $(f, c)$ has an averaging multiset with at most three distinct points.

To state the classification we need two special kinds of plane sets.

**Type I.** The graph of a function $r = \rho(\theta)$, in polar coordinates, of the following kind. Take $\beta \geq 2\pi$; take a positive function $\rho: [0, \pi] \to \mathbb{R}$ with irrational quotient $\gamma = \rho(\pi) / \rho(0)$ and extend it to $[0, \beta]$ by

$$\rho(\theta) = \gamma \rho(\theta - \pi) \quad \text{for } \theta > \pi.$$

**Type II.** Take $\beta \in (0, \pi)$ and a positive function $\rho_0: [0, \beta] \to \mathbb{R}$. Also take integers $m, n \geq 0$, numbers $\beta_1, \ldots, \beta_n \in [0, \beta]$ and $\alpha_1, \ldots, \alpha_m \in [\pi, \pi + \beta]$, distinct positive numbers $\delta_1 = 1, \delta_1, \ldots, \delta_n$ and distinct $\gamma_0, \gamma_1, \ldots, \gamma_m > 0$ such that all ratios $\gamma_i / \delta_i$ are irrational. Define

$$\rho_i = \delta_i \rho_0 \mid [\beta_i, \beta] \quad \text{for } i = 1, \ldots, n,$$

$$\rho^* : [\pi, \pi + \beta] \to \mathbb{R} \quad \text{by } \rho^*(\theta) = \rho_0(\theta - \pi),$$

$$\rho_0^* = \gamma_0 \rho^*,$$

$$\rho_j^* = \gamma_j \rho^* \mid [\pi, \alpha_j] \quad \text{for } j = 1, \ldots, m.$$  

A set of type II is the union of the graphs of $\rho_0, \rho_1, \ldots, \rho_n, \rho_0^*, \rho_1^*, \ldots, \rho_m^*$ and a compact, connected, locally connected set in the sector $\beta \leq \theta \leq \pi$ (excluding the origin) that meets the boundary rays $R_\beta : \theta = \beta$ and $R_\pi$ in finitely many points including all points $r = \rho(\beta)$ on $R_\beta$ and $r = \rho^*(\pi)$ on $R_\pi$.

We recall that a Peano space, a Hausdorff continuous image of $[0, 1]$, is any compact, connected, locally connected, metrizable space (Hahn–Mazurkiewicz theorem; cf. [7, Section 31]).

**Proposition 8.2.** Let $f$ be a continuous function from a Peano space $\Omega$, say $[0, 1]$ or $S^1$, to $\mathbb{R}^2$; and let $c \in \text{relint conv}(\text{Im} f)$. Then $(f, c)$ has an averaging multiset of one point if $c \in \text{Im} f$, of three distinct points (but no fewer) if there is a combination of translation and rotation carrying $\text{Im} f$ to $a$
set of type I or II and c to the origin, and of two distinct points (but no
fewer) otherwise. Every set of type I or II can occur as an image unless \( \Omega \)
supports only constant functions \( f: \Omega \to \mathbb{R} \).

We omit the proof, which can be easily reconstructed.

9. DISCONTINUOUS FUNCTIONS

We have generalized the integral mean value theorem for continuous
functions. But that theorem applies more generally, to any function that is a
derivative. This leads us to suspect that averaging sets exist more widely
than we have proved. What is the correct general class of functions that have
averaging sets?

For functions from \((0, 1)\) to \(\mathbb{R}^p\), the natural class to examine is of course
that of derivatives of differentiable functions. For them Lemma 3.1 goes
through, so the measure can be eliminated. However, we have not progressed
any further than that.

ACKNOWLEDGMENTS

Several of our colleagues at Ohio State gave us the benefit of their stimulating and helpful
comments. In particular we appreciate Eiichi Bannai's asking us to prove the existence of
spherical designs and \([0, 1]\) moment designs, Zbigniew Fiedorowicz's suggesting the
applicability of differential topology, Norman Levine's help with Peano spaces, Dijen K. Ray-
Chaudhuri's directing us to, and Zabeda Ally's explaining, the statistical applications, and
Steve Rallis's numerous questions.

REFERENCES

1. R. ABRAHAM AND J. ROBBIN, "Transversal Mappings and Flows," Benjamin, New York,
1967.
2. G. E. P. BOX AND N. R. DRAPER, A basis for the selection of a response surface design. J.
3. G. E. P. BOX AND J. S. HUNTER, Multi-factor experimental designs for exploring response
Dedicata 6 (1977), 363–388.