JOURNAL OF APPROXIMATION THEORY 12, 385–395 (1974)

Gradient Approximation of Vector Fields*

ROBERT M. GOOR

Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506

Communicated by R. C. Buck

In this paper, we consider an approximation problem which arose in optimal control theory. We seek conditions on a compact subset K of Euclidean n-space such that every continuous vector field on K may be uniformly approximated (on K) by vector fields with strictly positive integrating factors. We prove that such approximation is possible for all K in a particular subclass of the compact sets with topological dimension less than or equal to 1.

INTRODUCTION

The following problem arose in connection with a problem of optimal control in the Lagrange form, where the results of this paper were used to prove existence theorems (see [2]). Let K be a compact subset of a Euclidean space E^n, and let $C(K)^n$ be the Banach space of continuous vector fields $b(x) = (b_1, \ldots, b_n)$ (with n real components) defined on K with

$$\| b \| = \sup_{x \in K} | b(x)| = \sup_{x \in K} [b_1(x)^2 + \cdots + b_n(x)^2]^{1/2}.$$

We let $\tilde{F}(K)$ be the set of vector fields g, defined in an open neighborhood of K, which are of the form $g(x) = c(x) \nabla G(x)$ for some continuous, strictly positive function $c(x)$ and some continuously differentiable function $G(x)$. Here $\nabla G(x)$ is the gradient of G at x. Thus, the set $\tilde{F}(K)$ consists of vector fields defined in a neighborhood of K which have strictly positive integrating factors. The set of vector functions $\tilde{F}(K)$ induces a subset of $C(K)^n$ by restriction, and we denote by $F(K)$ the norm closure of this set in $C(K)^n$. We note that any element g of $\tilde{F}(K)$ is gradient-like, for, if $g(x) = c(x) \nabla G(x)$, then the inner product, $g(x) \cdot \nabla G(x)$, is strictly positive for all x such that $g(x) \neq 0$. Thus, we say that the elements of $F(K)$ are weakly gradient-like. We now pose the problem.

Given a vector field b in $C(K)^n$, under what conditions on b (and K) is b

* This research represents part of the author's Ph.D. thesis, written at the University of Michigan in 1972, under the guidance of Professor L. Cesari.

weakly gradient-like? As it stands, this problem has been too general to admit satisfactory solution. Rather, a restricted problem has proved more susceptible to attack. Namely, for what compact sets K does $F(K) = C(K)^n$? This is the approach we take in this paper.

In Section 1, we define a subset $G(K)$ of $F(K)$ which is useful in the approximation problem under consideration, and we discuss the relations between $G(K)$, $F(K)$ and $C(K)^n$. Section 2 is devoted to the case $n = 2$, i.e., we seek conditions on compact subsets K of the plane so that $F(K) = C(K)^2$. In Section 3, we review some definitions from topological dimension theory, and we use these concepts in our study of the case $n > 2$.

1

We define the subset $\tilde{G}(K)$ of $\tilde{F}(K)$ to be the set of all continuous vector functions g, defined in an open neighborhood of K, which are of the form $g(x) = \nabla G(x)$ for some continuously differentiable function $G(x)$. The set $\tilde{G}(K)$ defines a subset of $F(K)$ by restriction to K, and we define $G(K)$ to be the norm closure of this set. Thus, $G(K) \subseteq F(K) \subseteq C(K)^n$ for any compact set K in E^n. The advantage in working with the set $G(K)$ is that it is a linear subspace of $C(K)^n$ (since $\tilde{G}(K)$ is closed under multiplication by scalars and under addition), while, in general, $F(K)$ is not closed under addition. In fact, we have the following.

THEOREM. *If K is such that $F(K)$ is a linear subspace of $C(K)^n$, then $F(K) = C(K)^n$.*

Proof. Let $\nu = (\nu_1, ..., \nu_n)$ be a Borel measure on K with n real components so that if $g = (g_1, ..., g_n)$ is an element of $F(K)$, then $\int_K g \cdot d\nu = \sum_{i=1}^n \int_K g_i \cdot d\nu_i = 0$. That is, $d\nu$ annihilates the subspace $F(K)$. By definition, for any function $c(x)$, continuous and strictly positive in a neighborhood of K, and any function G, continuously differentiable in a neighborhood of K, $c\nabla G \in F(K)$. We consider the function G to be fixed (but arbitrary). Because $F(K)$ is closed, for any continuous, *nonnegative* function $c(x)$ defined on K, $c(x) \nabla G(x)$ lies in $F(K)$. Let $f(x)$ be any continuous, real-valued function on K. Let $f^+(x) = \max\{0, f(x)\}$, and let $f^-(x) = \min\{0, f(x)\}$. Then $f^+(x) \geqslant 0$ and $f^-(x) \leqslant 0$ for all x in K. Furthermore, $f(x) = f^+(x) + f^-(x)$ so that

$$f(x) \nabla G(x) = f^+(x) \nabla G(x) + f^-(x) \nabla G(x),$$

for all x in K. Therefore, if $H(x) = -G(x)$, then for all x in K,

$$f(x) \nabla G(x) = f^+(x) \nabla G(x) + [-f^-(x)] \nabla H(x).$$

Since each element of the sum is in $F(K)$, and $F(K)$ is assumed to be closed under addition, $f(x)\nabla G(x)$ is in $F(K)$. Therefore,

$$\int_K f(x)\nabla G(x)\cdot d\nu(x)=0,$$

for all continuous, real-valued functions f defined on K. We let $d\mu$ be the real measure $\nabla G\cdot d\nu=\sum_{i=1}^{n}(\partial G/\partial x_i)\,d\nu_i$. Then, we have shown that $\int_K f(x)\,d\mu(x)=0$ for all continuous, real-valued functions f on K. It follows that $d\mu=\nabla G\cdot d\nu$ is the zero measure for all continuously differentiable real-valued functions G. Let $b=(b_1,..., b_n)$ be an arbitrary fixed vector in E^n and define $G(x)=b\cdot x=b_1x_1+\cdots+b_nx_n$ for all x in E^n. Then, $\nabla G(x)=b$ for all x, and $b\cdot d\nu$ is the zero measure. That is,

$$0=\int_{K'} b\cdot d\nu=b\cdot\nu(K'),$$

for all Borel measurable subsets K' of K. Since b is arbitrary, $\nu(K')=0$ for all Borel measurable subsets K' of K. Therefore, ν is the zero measure. Since the only Borel measure on K which annihilates $F(K)$ is the zero measure, and $F(K)$ is a linear subspace of $C(K)^n$, $F(K)$ must equal $C(K)^n$, and the theorem is proved.

This result is not as useful as it is interesting, since demonstrating that $F(K)$ is linear seems to be as difficult as verifying that $F(K)=C(K)^n$ by more direct means.

We conclude this section with an example of a compact set K_1 in E^2 for which $G(K_1)\neq F(K_1)\neq C(K_1)^2$, and an example of a compact set K_2 in E^n for which $G(K_2)=F(K_2)=C(K_2)^n$.

EXAMPLE 1.1. We denote by (x, y) the points in E^2 and we define K to be the set of all (x, y) in E^2 such that $x^2+y^2=1$. For (x, y) in K_1, we define $g(x, y)=(-y, x)$ so that $g\in C(K)^2$. We will show that g is *not* an element of $G(K_1)$ *or* $F(K_1)$. We define the bounded linear functional T on elements $h=(h_1, h_2)$ of $C(K_1)^2$ by setting

$$T(h)=\int_0^{2\pi}[-h_1(\cos t,\sin t)\sin t+h_2(\cos t,\sin t)\cos t]\,dt,$$

so that $T(h)$ is the counter-clockwise path integral of h along K_1. Thus, $T(g)=2\pi$. But T annihilates the elements of $G(K_1)$, so $g\notin G(K_1)$. Now, suppose g is an element of $F(K_1)$. Then, for each $n=1, 2,...,$ there is a continuous, strictly positive function $c_n(x, y)$ and a continuously differentiable function $G_n(x, y)$, both defined in a neighborhood of K_1, such that

$$|g(x, y)-c_n(x, y)\nabla G_n(x, y)|<1/n,$$

for all (x, y) in K_1. Since $|g(x, y)| = 1$ for all (x, y) in K_1, for $n \geqslant 2$, $\nabla G_n(x, y) \neq (0, 0)$ for all (x, y) in K_1. In particular, we look at $n = 2$. Let $\gamma = \max c_2(x, y)$, the maximum taken for (x, y) in K_1. Then, by the Schwarz inequality,

$$| c_2(x, y) \nabla G_2(x, y) \cdot g(x, y) - | g(x, y)|^2 | < \tfrac{1}{2},$$

for all (x, y) in K_1. Since $| g(x, y)|^2 = 1$, we have that

$$c_2(x, y) \nabla G_2(x, y) \cdot g(x, y) > \tfrac{1}{2},$$

for all (x, y). Therefore,

$$\nabla G_2(x, y) \cdot g(x, y) > (1/2c_2(x, y)) \geqslant 1/2\gamma,$$

for all (x, y) in K_1. We now note that

$$T(\nabla G_2) = \int_0^{2\pi} \nabla G_2(\cos t, \sin t) \cdot g(\cos t, \sin t)\, dt \geqslant \pi/\gamma > 0,$$

which contradicts the fact that T annihilates elements of $G(K_1)$. Therefore, the open ball of radius $\frac{1}{2}$ in $C(K_1)^2$, centered at g, contains no elements of $F(K_1)$, so $F(K_1) \neq C(K_1)^2$. From the theorem proved above, it is clear that $G(K_1) \neq F(K_1)$, since $F(K_1)$ cannot be a linear subspace of $C(K_1)^2$. Thus, $G(K_1) \neq F(K_1) \neq C(K_1)^2$.

EXAMPLE 1.2. Let K_2 be the compact subset of E^n defined by

$$K_2 = \{(x_1, x_2, ..., x_n) \mid 0 \leqslant x_1 \leqslant 1, x_2 = \cdots = x_n = 0\}.$$

We will show that $G(K_2) = C(K_2)^n$, and hence that $F(K_2) = C(K_2)^n$. Let $g(x_1, ..., x_n) = (g_1(x_1, ..., x_n), ..., g_n(x_1, ..., x_n))$ be an element of $C(K_2)^n$ such that each function $g_2, ..., g_n$ is continuously differentiable. Such functions are dense in $C(K_2)^n$. Actually, each function g_i depends only on the variable x_1, so we write the function g as $g(x_1) = (g_1(x_1), ..., g_n(x_1))$. Since g_1 is continuous, it is Riemann integrable and we may define $G_1(x_1) = \int_0^{x_1} g_1(t)\, dt$. We define $H(x_1, ..., x_n)$ by setting

$$H(x_1, ..., x_n) = G_1(x_1) + \sum_{i=2}^{n} x_i g_i(x_1).$$

Then, H has continuous partial derivatives and

$$H_{x_i}(x_1, 0, ..., 0) = g_i(x_1),$$

for $i = 1, 2, ..., n$ and for $0 \leqslant x_1 \leqslant 1$. Thus, each such element g of $C(K_2)^n$ is an element of $G(K_2)$. Since the functions with continuous derivatives are dense in $C(K_2)^n$, $G(K_2) = C(K_2)^n$.

The difference between the sets K_1 and K_2 illustrates the underlying theme of this paper. We will consider (one dimensional) compact sets K in E^n which, in an appropriate sense, have the topological properties of K_2 and avoid those of K_1.

2

In this section, we study the case $n = 2$. That is, we investigate conditions on a compact set K in E^2 such that $G(K) = C(K)^2$. For the purposes of this section only, we regard E^2 as the complex plane: $E^2 = \{x + iy \mid x \text{ and } y \text{ lie in } E^1\}$, where $i^2 = -1$. For K a compact subset of E^2, we define $\mathbf{C}(K)$ to be the space of continuous, complex-valued functions defined on K. $\mathbf{C}(K)$ is a Banach space under the norm defined for $f = f_1 + if_2$ by

$$\|f\|_\infty = \sup_{z \in K} |f(z)| = \sup_{z \in K} [f_1(z)^2 + f_2(z)^2]^{1/2}.$$

The subspace $\mathbf{P}(K)$ of $\mathbf{C}(K)$ is defined to be the norm closure in $\mathbf{C}(K)$ of the polynomials on K. That is, $\mathbf{P}(K)$ is generated by functions $p(z)$ which are polynomials in the complex variable z. We will need the following theorem, which is proved in ([1], p. 48).

THEOREM 2.1 (Mergelyan's Theorem). *Let K be a compact subset of E^2 whose complement is connected. If $f \in \mathbf{C}(K)$ and f is analytic on the interior of K, then $f \in \mathbf{P}(K)$.*

It is clear that $C(K)^2$ and $\mathbf{C}(K)$ are isometrically isomorphic. We will use this fact in conjunction with Theorem 2.1 to prove the main theorem of this section, Theorem 2.2.

THEOREM 2.2. *If K is a compact subset of E^2 which is nowhere dense and has connected complement in E^2, then $G(K) = C(K)^2$.*

Proof. For $f = (f_1, f_2)$ in $C(K)^2$, we define $U(f) = f_1 - if_2$ so that $U(f) \in \mathbf{C}(K)$. It is clear that U is an isometry from $C(K)^2$ onto $\mathbf{C}(K)$. Let $p(z)$ be a polynomial in the complex variable $z = x + iy$, defined in the whole complex plane, and hence in a neighborhood of K. Then, there are real polynomials p_1 and p_2 in the two real variables x and y such that $p(x + iy) = p_1(x, y) + ip_2(x, y)$. By the Cauchy–Riemann equations,

$$\partial p_1/\partial y = -\partial p_2/\partial x,$$

for all real x, y. Thus, by Green's theorem, the differential $p_1\,dx - p_2\,dy$ is exact on the entire plane. It follows that there is a continuously differentiable function $P = P(x, y)$ such that $\partial P/\partial x = p_1$ and $\partial P/\partial y = -p_2$. We have therefore shown that $U^{-1}(p) \in G(K)$ for every complex polynomial p. Since

U^{-1} is continuous and maps a dense subset of $\mathbf{P}(K)$ into $G(K)$, we conclude that $U^{-1}(\mathbf{P}(K)) \subseteq G(K)$. We now use our assumptions about K. Since K is nowhere dense, K has empty interior, and every function in $\mathbf{C}(K)$ is analytic on the (empty) interior of K. By Theorem 2.1, then, since K has connected complement, if $f \in \mathbf{C}(K)$, we have that $f \in \mathbf{P}(K)$, i.e., $\mathbf{P}(K) = \mathbf{C}(K)$. Therefore, under our hypotheses on K, $U^{-1}(\mathbf{C}(K)) \subseteq G(K)$. But, U is an isometry of $C(K)^2$ onto $\mathbf{C}(K)$, so $U^{-1}(\mathbf{C}(K)) = C(K)^2$. Thus, $C(K)^2 \subseteq G(K) \subseteq C(K)^2$, and the theorem follows.

In the next section, we attempt to imitate these results when K is a compact subset of E^n, $n > 2$. However, neither of the hypotheses stated in Theorem 2.2 on the subset K of E^2 generalizes verbatim to the case $n > 2$. If K is a compact subset of E^n which is nowhere dense in E^n, then it has no interior in E^n, but for $n > 2$, K could have "dimension" larger than or equal to 2. For example, if $K = \{(x, y, z) \in E^3 \mid x^2 + y^2 \leqslant 1, z = 0\}$, then K is nowhere dense in E^3, but, as in Example 1.1, it may be shown that $G(K) \neq F(K) \neq C(K)^3$. Similarly, the "connected complement" condition is not sufficient for $n > 2$. For example, the set $K = \{(x, y, z) \in E^3 \mid x^2 + y^2 = 1, z = 0\}$ has connected complement in E^3, but, as in Example 1.1, $G(K) \neq F(K) \neq C(K)^3$. Thus, in Section 3, we consider the appropriate topological generalizations for $n > 2$.

3

We will need some of the concepts of topological dimension theory, and for these, we draw on [3]. Let K be a compact subset of E^n and let $\{U_1, ..., U_k\}$ be a covering of K by open sets. The order of the covering $\{U_1, ..., U_k\}$ is defined to be the largest integer N such that there are $N + 1$ members of the covering whose intersection is nonempty. If $\{V_1, ..., V_m\}$ is a covering of K by open sets, we say that $\{V_1, ..., V_m\}$ is a refinement of $\{U_1, ..., U_k\}$ if for each V_j, $j = 1, 2, ..., m$, there is an i, $1 \leqslant i \leqslant k$, such that $V_j \subseteq U_i$. (See [3, pp. 52–53].) We may define the dimension of K, $\dim K$, as follows. We will say that $\dim K \leqslant N$ if every covering of K by finitely many open sets has a refinement of order less than or equal to N. In particular, the empty set is the only set with dimension -1. Further, if $\dim K = 0$, every covering of K by finitely many open sets has a refinement whose elements are pairwise disjoint. We thus have the following theorem.

THEOREM 3.1. *If K is a nonempty compact subset of E^n such that* $\dim K = 0$, *then* $G(K) = C(K)^n$.

Proof. Let g be an arbitrary element of $C(K)^n$ and let $\epsilon > 0$ be arbitrary. For each x_0 in K, there is an open neighborhood $U = U(x_0)$ of x_0 such that

$$| g(x) - g(x_0)| < \epsilon/2,$$

for all x in $K \cap U(x_0)$. Since $\{U(x)\}_{x \in K}$ is a covering of K by open sets, the family $\{U(x)\}_{x \in K}$ has a Lebesgue number δ. Thus, if V is an open subset of E^n of diameter less than δ and if $V \cap K$ is nonempty, then $V \subseteq U(x)$ for some x in K. Since K is compact, K is totally bounded, so there is a covering of K by finitely many open sets $\{V_1, ..., V_m\}$ such that each V_i, $1 \leqslant i \leqslant m$, has diameter less than δ. Since $\dim K = 0$, there is a refinement $\{U_1, ..., U_N\}$ of $\{V_1, ..., V_m\}$ such that $U_i \cap U_j = \phi$ if $i \neq j$. By the definition of the family $\{U_1, ..., U_N\}$ the diameter of U_i is less than δ for each i. For $i = 1, 2, ..., N$, let x_i be any element of $K \cap U_i$. Then, $U_i \subseteq U(x')$ for some x' in K, so that, for any x in $K \cap U_i$,

$$| g(x) - g(x_i)| \leqslant | g(x) - g(x')| + | g(x') - g(x_i)| < \epsilon.$$

We let $b_i = g(x_i) \in E^n$ for $i = 1, 2, ..., N$, and we define a real-valued function G on $U = \bigcup_{i=1}^{N} U_i$ as follows. Let $G(x) = b_i \cdot x$ for all x in U_i. Since the sets $\{U_1, ..., U_N\}$ are pairwise disjoint, G is well-defined. Also, $\nabla G(x) = b_i$ for all x in U_i. It follows that G is continuously differentiable on U and, for x in $K \cap U_i$,

$$| g(x) - \nabla G(x)| = | g(x) - b_i | < \epsilon.$$

Therefore, $| g(x) - \nabla G(x)| < \epsilon$ for all x in K. Since $\epsilon > 0$ was arbitrary, $g \in G(K)$. Since $g \in C(K)^n$ was arbitrary, $G(K) = C(K)^n$ and Theorem 3.1 is proved.

Our main result of the section concerns a subclass of the compact sets K satisfying $\dim K \leqslant 1$.

We note that in Section 2, we have already proved a theorem which can be used to derive a statement of this type. For, by [3, p. 41, Theorem IV 1], $\dim E^n = n$ for all n, and, if $K \subseteq E^n$, then, by [3, p. 44, Theorem IV 3], $\dim K = n$ if and only if K contains a nonempty subset which is open in E^n. In particular, for $n = 2$, $K \subseteq E^2$ is nowhere dense in E^2 if and only if $\dim K \leqslant 1$. Thus, we may restate Theorem 2.2 to say: if K is a compact subset of E^2 such that $\dim K \leqslant 1$ and K has connected complement in E^2, then $G(K) = C(K)^2$. We will prove a similar, though weaker statement for $K \subseteq E^n$, $n > 2$.

We now state some definitions. We will assume that $B_1, ..., B_k$ are open balls in the Euclidean space E_n with centers $x_1, ..., x_k$, respectively. If $B_i \cap B_j \neq \phi$ for $i \neq j$, we denote by L_{ij} the undirected line segment joining x_i and x_j. Thus, $L_{ij} = L_{ji}$. If $B_i \cap B_j$ is empty, we let $L_{ij} = L_{ji} = \phi$, and we set $L_{ii} = \{x_i\}$ for each i. We define a point set $L(B_1, ..., B_k)$ by setting

$$L(B_1, ..., B_k) = \bigcup_{i,j=1}^{k} L_{ij}.$$

We note that $L(B_1, \ldots, B_k)$ is a finite union of piecewise linear curves and that $L(B_1, \ldots, B_k)$ has the same number of connected components as $\bigcup_{i=1}^k B_i$.

We will say that the collection $\{B_1, \ldots, B_k\}$ forms a simple chain if $L(B_1, \ldots, B_k)$ contains no simple closed curve. Let K be a compact subset of E^n. We will say that K has the simple chain covering property if for every $\epsilon > 0$, there is a covering of K by open balls $B_1, \ldots, B_k$ in E^n, each of radius less than ϵ, such that the collection $\{B_1, \ldots, B_k\}$ forms a simple chain. We note that if K has the simple chain covering property, then $\dim K \leqslant 1$, since every covering by open balls $\{B_1, \ldots, B_k\}$ which form a simple chain, has order less than or equal to 1. Furthermore, if a set K has the simple chain covering property, then clearly K contains no simple closed curve. We may prove the following theorem.

THEOREM 3.2. *If K is a compact subset of E^n having the simple chain covering property, then $F(K) = C(K)^n$.*

Proof. Let $g = (g^1, g^2, \ldots, g^n)$ be an element of $C(K)^n$ and let $\epsilon > 0$ be given. For each x_0 in K, there is an open n-ball $B(x_0)$ centered at x_0 such that $| g(x_0) - g(x)| < \epsilon/4$ for x in $K \cap B(x_0)$. The open balls $B(x)$, for x in K, cover the set K. Since K is compact, the covering $\{B(x)\}_{x \in K}$ has a Lebesgue number δ. Let $B_1, \ldots, B_k$ be a covering of K by open balls in E^n of radius less than $\delta/2$, such that $\{B_1, \ldots, B_k\}$ forms a simple chain. Thus, for each B_i, there is an element y_i of K such that $B_i \subseteq B(y_i)$. Let $b_i = g(y_i)$ and write $b_i = (b_i^1, b_i^2, \ldots, b_i^n)$. We let r_i be the radius of B_i and we write $c_i = (c_i^1, c_i^2, \ldots, c_i^n)$ for the center of B_i, $i = 1, 2, \ldots, k$. We now note that since K is compact, there is a positive distance δ' between K and $\partial(\bigcup_{i=1}^k B_i)$. In particular, since $\{B_1, \ldots, B_k\}$ forms a simple chain, $\bigcup_{i \neq j} (\partial B_i \cap \partial B_j) \subseteq \partial(\bigcup_{i=1}^k B_i)$, and if W is a $\delta'/2$-neighborhood of $\bigcup_{i \neq j} (\partial B_i \cap \partial B_j)$, then $K \cap cl W = \phi$. We define B to be the set $\bigcup_{i=1}^k B_i \backslash cl W$, so that B is an open neighborhood of K in E^n. We will define a vector function $f = (f^1, f^2, \ldots, f^n)$ on B such that $| g(x) - f(x)| < \epsilon$ for all x in K, and such that $f \in F(K)$. That is, we will find a continuous, strictly positive function $h(x)$ on B and a real-valued, continuously differentiable function $G(x)$ on B such that $f(x) = h(x) \nabla G(x)$ for all x in B. The vector function f will be defined inductively. In fact, since it is to be constant in each of the sets $B_j \backslash (\bigcup_{i \neq j} B_i)$, the induction step will be obvious when the construction is verified in the sets $B_i \cap B_j$, for $i \neq j$. Since the family $\{B_1, \ldots, B_k\}$ forms a simple chain and hence has order less than or equal to 1, we need never consider the intersections of 3 or more distinct elements of the family.

For x in $B_j \backslash (\bigcup_{i \neq j} B_i \cup cl W)$, we define $f(x) = b_j = g(y_j)$. Here, without loss of generality, we assume that if $B_i \cap B_j \neq \phi$, $i \neq j$, then b_j is *not* perpendicular to the line connecting the centers c_i and c_j. (If this not the case, we may perturb b_j by adding a vector of length less than $\epsilon/4$ so that the above is satisfied. The remaining statements then proceed without difficulty if we consider that $| b_j - g(y_j)| < \epsilon/4$ for each j.) Thus, for each x in $B_j \backslash (\bigcup_{i \neq j} B_i \cup cl W)$,

$$|f(x) - g(x)| \leqslant | b_j - g(x)| \leqslant | b_j - g(y_j)| + | g(y_j) - g(x)| < \epsilon/2.$$

Now, suppose that $f(x)$ is extended to all of B in such a way that for x lying in $B_i \cap B_j$, for $i \neq j$, then $f(x)$ is a convex combination of b_i and b_j. We then have the following. For x in $B_i \cap B_j \cap K$, there is a λ, $0 \leqslant \lambda \leqslant 1$, such that $f(x) = \lambda b_i + (1 - \lambda) b_j$. Thus,

$$|f(x) - g(x)| \leqslant \lambda | b_i - g(y_i)| + \lambda | g(y_i) - g(x)| \\ + (1 - \lambda)| b_j - g(y_j)| + (1 - \lambda)| g(y_j) - g(x)|,$$

which is less than ϵ since each absolute value is bounded by $\epsilon/4$. We will therefore show how to define f in the sets $B_i \cap B_j$, $i \neq j$, so that f is a convex combination of b_i and b_j.

We assume, therefore, that $B_i \cap B_j$ is non-empty, for some $i \neq j$. Without loss of generality, we may assume that the centers c_i and c_j of B_i and B_j both lie along the x_1-axis, and, in fact, we may assume that $c_i = 0$ while $c_j = (a, 0, 0,..., 0)$ for some $a > 0$. Thus, $\partial B_i \cap \partial B_j$ lies in a hyperplane perpendicular to the x_1-axis, say $\{(x_1 ,..., x_n) \mid x_1 = a'\}$, where $0 < a' < a$. Let a_1, a_2 be positive numbers such that $a_1 < a' < a_2$ and such that: the intersection of the hyperplane $x_1 = a_1$ and the set ∂B_j lies in the open set W; the intersection of the hyperplane $x_1 = a_2$ and the set ∂B_i lies in the open set W.

We note that neither of the vectors b_i and b_j are perpendicular to the x_1-axis (by assumption). If $b_i = b_j$, we may set $h(x) = 1$ and $f(x) = b_i = b_j$ for all x in $B_i \cup B_j$. We therefore assume that $b_i \neq b_j$.

Let $\bar{b}_i$ and $\bar{b}_j$ be the vectors in E^{n+1} formed from b_i, b_j by setting $\bar{b}_i = (b_i , b)$, $\bar{b}_j = (b_j , b)$, where b is chosen to be any positive number large enough so that $\bar{b}_i \cdot \bar{b}_j = (b)^2 + \sum_{k=1}^{n} b_i^k b_j^k > 0$. Since $b_i \neq b_j$, $\bar{b}_i$ and $\bar{b}_j$ are linearly independent. We define row vectors v_1, v_2,..., v_{n+1} in E^{n+1} as follows. We let v_3, v_4,..., v_{n+1} be $n - 1$ row vectors in E^{n+1} such that $\bar{b}_i$, $\bar{b}_j$, v_3,..., v_{n+1} are linearly independent and such that v_k is perpendicular to $\bar{b}_i$ and $\bar{b}_j$ for all k, $3 \leqslant k \leqslant n + 1$. We let $v_1 = (1, 0, 0,..., 0)$ and we note that since $v_1 \cdot \bar{b}_i \neq 0$ and $v_1 \cdot \bar{b}_j \neq 0$, the family v_1, $\bar{b}_j$, v_3,..., v_{n+1} is linearly independent. We define $v_2 = \bar{b}_j$. Thus, $\{v_1 , v_2 ,..., v_{n+1}\}$ forms a linearly independent set of row vectors. Let M be the $(n + 1) \times (n + 1)$ matrix whose kth row is v_k.

For d in E^1, let $L = \{\bar{x} = (x_1 ,..., x_{n+1}) \in E^{n+1} \mid x_1 = d\}$. Then, we have defined M so that $M(L_0)$ is a subset of L_0. (When we write $M(\bar{x})$ for some vector $\bar{x}$ in E^{n+1}, we have in mind that $\bar{x}$ is a column vector.) Since the rows of M are linearly independent, M is invertible and $M(L_0) = L_0$. We write $L_d = (d, 0,..., 0) + L_0$. Then, $M(L_d) = M(d, 0, 0,..., 0) + L_0$ so that L_d is parallel to $M(L_d)$. Note, however, that $M(d, 0,..., 0) = (d, d_2, ..., d_{n+1})$ for some scalars $d_2, ..., d_{n+1}$ so that $M(d, 0,..., 0)$ is an element of $(d, 0,..., 0) + L_0$. Thus, $M(L_d) \cap L_d$ is non-empty. It follows that $M(L_d) = L_d$ for all d in E^1.

We define

$$p = M(\bar{b}_i) = (\bar{b}_i \cdot v_1, \bar{b}_i \cdot v_2, \bar{b}_i \cdot v_3, ..., \bar{b}_i \cdot v_{n+1}) = (p_1, p_2, 0, 0,..., 0)$$

where we have constructed M so that p_2 is positive. Similarly, we define $q = M(\bar{b}_j) = (q_1, q_2, 0, 0,..., 0)$, where q_2 is positive. Since $\bar{b}_i$, $\bar{b}_j$ are independent, so are p and q. We will define a vector-valued function $\bar{f}$ on E^{n+1} such that: $\bar{f}(\bar{x}) \in E^{n+1}$ for all $\bar{x}$; $\bar{f}(x_1, ..., x_{n+1}) = p$ for $x_1 \leqslant a_1$; $\bar{f}(x_1, ..., x_{n+1}) = q$ for $x_1 \geqslant a_2$ and $\bar{f}(x_1, ..., x_{n+1})$ is a convex combination of p and q for $a_1 < x_1 < a_2$. For $a_1 \leqslant x_1 \leqslant a_2$, we define

$$\lambda(x_1) = (a_2 - x_1)/\{(p_2/q_2)(x_1 - a_1) + (a_2 - x_1)\}.$$

Then, $\lambda(a_1) = 1$, $\lambda(a_2) = 0$ and $0 \leqslant \lambda(x_1) \leqslant 1$ for all x_1 such that $a_1 \leqslant x_1 \leqslant a_2$. For $a_1 \leqslant x_1 \leqslant a_2$, we define $\bar{f}(x_1, ..., x_{n+1})$ by setting

$$\bar{f}(x_1, ..., x_{n+1}) = \lambda(x_1)\, p + (1 - \lambda(x_1))\, q.$$

We define $\bar{f}$ by continuity and constancy outside the interval $a_1 \leqslant x_1 \leqslant a_2$. We now define a real-valued function $\bar{h}(x_1, ..., x_{n+1})$ by setting

$$\bar{h}(x_1, ..., x_{n+1}) = (a_2 - a_1)/\{(p_2/q_2)(x_1 - a_1) + (a_2 - x_1)\},$$

for $a_1 \leqslant x_1 \leqslant a_2$. We set $\bar{h}(x_1, ..., x_{n+1}) = 1$ for $x_1 \leqslant a_1$ and $\bar{h}(x_1, ..., x_{n+1}) = q_2/p_2$ for $x_1 \geqslant a_2$. Thus, $\bar{h}$ is continuous. Since $\bar{h}$ is monotone in the interval $a_1 \leqslant x_1 \leqslant a_2$, it is easy to see that $\bar{h}$ takes on values between 1 and q_2/p_2. Therefore, $\bar{h}(x_1, ..., x_{n+1}) > 0$ for all $x_1, ..., x_{n+1}$.

We consider the quotient $(\bar{g}_1, \bar{g}_2, ..., \bar{g}_{n+1}) = \bar{f}/\bar{h}$. It is an easy computation to verify that

$$\partial \bar{g}_k / \partial x_m = \partial \bar{g}_m / \partial x_k,$$

at all points $(x_1, ..., x_{n+1})$ in E^{n+1}. Therefore, there is a real-valued, continuously differentiable function $\bar{G}$ such that $\nabla \bar{G}(\bar{x}) = \bar{f}(\bar{x})/\bar{h}(\bar{x})$, or, $\bar{f}(\bar{x}) = \bar{h}(\bar{x})\, \nabla \bar{G}(\bar{x})$ for all $\bar{x}$ in E^{n+1}.

For $\bar{x}$ in E^{n+1}, we define $G_0(\bar{x}) = \bar{G}(M^{-1}(\bar{x}))$ and $h_0(\bar{x}) = \bar{h}(M^{-1}(\bar{x}))$. Thus, $\nabla G_0(\bar{x}) = M^{-1}[\nabla \bar{G}(M^{-1}(\bar{x}))]$. Since $M(L_d) = L_d$ implies $M^{-1}(L_d) = L_d$ for all real d, both $h_0(x_1, ..., x_{n+1})$ and $\nabla G_0(x_1, ..., x_{n+1})$ are constant for $x_1 \leqslant a_1$ and for $x_1 \geqslant a_2$ (as are $\bar{h}$ and $\nabla \bar{G}$). We will write $x = (x_1, ..., x_n)$ and $\bar{x} = (x, x_{n+1}) = (x_1, ..., x_{n+1})$. For $x_1 \leqslant a_1$,

$$h_0(\bar{x})\, \nabla G_0(\bar{x}) = M^{-1}(p) = \bar{b}_i,$$

and for $x_1 \geqslant a_2$,

$$h_0(\bar{x}) \nabla G_0(\bar{x}) = M^{-1}(q) = \bar{b}_j .$$

Since, for $a_1 \leqslant x_1 \leqslant a_2$, $\bar{h}(\bar{x}) \nabla \bar{G}(\bar{x})$ is a convex combination of p and q, in the same interval, $h_0(\bar{x}) \nabla G_0(\bar{x})$ is a convex combination of $\bar{b}_i$ and $\bar{b}_j$. Therefore, the $(n + 1)$st component of $h_0 \nabla G_0$ is constant and equal to b. We may define a function G of n variables, then, by setting $G(x) = G_0(x, 0)$, and, it is clear that if $h(x) = h_0(x, 0)$, then $h(x) \nabla G(x)$ is a convex combination of b_i and b_j for $a_1 \leqslant x_1 \leqslant a_2$. Also, $h(x) \nabla G(x) = b_i$ for $x_1 \leqslant a_1$ and $h(x) \nabla G(x) = b_j$ for $x_1 \geqslant a_2$. Finally, both h and ∇G are constant for $x_1 \leqslant a_1$, and for $x_1 \geqslant a_2$.

For x in $(B_i \cap B_j) \backslash cl W$, we define $f(x) = h(x) \nabla G(x)$. From our earlier demonstration, we have that $|f(x) - g(x)| < \epsilon$ for all x in $K \cap (B_i \cap B_j)$. The extensions of f, h and G to all of B require merely an appropriate scaling of h. Since $\{B_1, ..., B_k\}$ forms a simple chain, all three functions will be well-defined on extension and we have that

$$|f(x) - g(x)| = |h(x) \nabla G(x) - g(x)| < \epsilon \qquad \text{for all } x \text{ in } K.$$

Therefore, since $\epsilon > 0$ was arbitrary, g is an element of $F(K)$. Since $g \in C(K)^n$ was arbitrary, $F(K) = C(K)^n$, and the theorem is proved.

References

1. T. Gamelin, "Uniform Algebras," Prentice-Hall, Englewood Cliffs, NJ, 1969.
2. R. M. Goor, Existence theorems in the calculus of variations and approximation, to appear.
3. W. Hurewicz and H. Walman, "Dimension Theory," Princeton University Press, Princeton, NJ, 1948.