JOURNAL OF APPROXIMATION THEORY 73, 218–236 (1993)

Interpolation by Ridge Functions

DIETRICH BRAESS

Institut für Mathematik, Ruhr-Universität, D-4630 Bochum, Germany

AND

ALLAN PINKUS

Department of Mathematics, Technion, I.I.T., Haifa, Israel

Communicated by E. W. Cheney

Received May 1, 1991; accepted in revised form February 19, 1992

We consider the problem of interpolation by linear combinations of ridge functions. A ridge function is a function of the form $f(\mathbf{a}\cdot\mathbf{x})$ where $f\colon \mathbb{R}\to\mathbb{R}$, $\mathbf{a}\in\mathbb{R}^d\backslash\{\mathbf{0}\}$ is a fixed vector, and $\mathbf{x}\in\mathbb{R}^d$ is the variable. The solvability of the interpolation problem is characterized by geometric properties. © 1993 Academic Press, Inc.

1. INTRODUCTION

Ridge functions are functions of the form

$$f(\mathbf{a}\cdot\mathbf{x}),$$

where $\mathbf{x}=(x_1,\ldots,x_d)\in\mathbb{R}^d$ are the variables, $\mathbf{a}\in\mathbb{R}^d\backslash\{\mathbf{0}\}$ is a fixed vector (direction), $\mathbf{a}\cdot\mathbf{x}=\sum_{i=1}^d a_i x_i$, and $f\colon\mathbb{R}\to\mathbb{R}$. In other words, the class of ridge functions is a simple subset of the set of all d-variable real-valued functions $g\colon\mathbb{R}^d\to\mathbb{R}$, given by composition of an inner product (linear functional on $\mathbb{R}^d$) with a one variable real-valued function. The literature abounds with uses of such functions, or combinations thereof, to approximate d-variable functions. In this paper we consider the problem of interpolation by ridge functions with fixed directions. Our results are very partial and represent a first step in an interesting problem

We call the vector $\mathbf{a}\in\mathbb{R}^d\backslash\{\mathbf{0}\}$ a *direction* because for any $f\colon\mathbb{R}\to\mathbb{R}$, the function $f(\mathbf{a}\cdot\mathbf{x})$ is a constant on

$$\Gamma_{\mathbf{a}}(\alpha):=\{\mathbf{x}:\mathbf{a}\cdot\mathbf{x}=\alpha\}.$$

0021-9045/93 $5.00

For each direction $\mathbf{a}$, let

$$\sum(\mathbf{a}) := \{f(\mathbf{a}\cdot\mathbf{x})\colon \text{all } f\colon \mathbb{R}\to\mathbb{R}\},$$

i.e., in $\sum(\mathbf{a})$ we vary over all f. Given $\mathbf{a}^1, ..., \mathbf{a}^k \in \mathbb{R}^d\setminus\{\mathbf{0}\}$, we consider the following problem:

Characterize those points $\mathbf{x}^i, ..., \mathbf{x}^m \in \mathbb{R}^d$ *(any* m*) such that for every choice of data* $\alpha_1, ..., \alpha_m$ $(\alpha_i \in \mathbb{R},\ i=1, ..., m)$, *there exists a function*

$$g \in \sum(\mathbf{a}^1) + \cdots + \sum(\mathbf{a}^k)$$

satisfying

$$g(\mathbf{x}^i) = \alpha_i, \qquad i = 1, ..., m.$$

That is, there exist $f_1, ..., f_k\colon \mathbb{R}\to\mathbb{R}$ satisfying

$$\sum_{j=1}^{k} f_j(\mathbf{a}^j\cdot\mathbf{x}^i) = \alpha_i, \qquad i=1, ..., m.$$

In this paper we will restrict ourselves to the case where $d=2$ (the plane). We will consider in detail the case $k=3$ and also $k>3$ for a significant subclass.

The first non-trivial case $k=2$ is well understood; see Section 2 and, e.g., Dyn *et al.* [1]. In this case we may apply a linear transformation so that the two directions are parallel to the coordinates. Then the problem is reduced to interpolation by functions of the form

$$f_1(x_1) + f_2(x_2).$$

One solution in this case is the following. Given a set of points $\{\mathbf{x}^i\}_{i\in I}$, the interpolation problem is not solvable if and only if there is a subset of $2q$ points $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^{2q} = \mathbf{x}^0$ such that

$$\begin{aligned} x_1^i &= x_1^{i-1} \quad &&\text{if } i \text{ is even},\\ x_2^i &= x_2^{i-1} \quad &&\text{if } i \text{ is odd}. \end{aligned}$$

A situation in which interpolation is not possible is shown in Fig. 1.1.

It is our aim to establish an analogous characterization when there are more than two directions.

Before considering specific cases, we introduce some general notation and establish a useful reformulation of the problem.

DEFINITION 1.1. Given directions $\{\mathbf{a}^j\}_{j=1}^k \subset \mathbb{R}^d\setminus\{\mathbf{0}\}$, we say that the set of points $\{\mathbf{x}^i\}_{i=1}^m \subset \mathbb{R}^d$ has the *NI-property* (non-interpolation property)

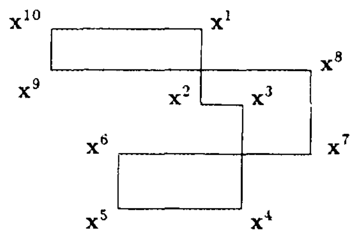

FIGURE 1.1

with respect to $\{\mathbf{a}^j\}_{j=1}^k$ if there exist $\{\alpha_i\}_{i=1}^m \subset \mathbb{R}$ such that we *cannot* find $f_j: \mathbb{R} \to \mathbb{R}$, $j = 1, ..., k$, satisfying

$$\sum_{j=1}^{k} f_j(\mathbf{a}^j \cdot \mathbf{x}^i) = \alpha_i, \qquad i = 1, ..., m.$$

We say that the set $\{\mathbf{x}^i\}_{i=1}^m \subset \mathbb{R}^d$ has the *MNI-property* (minimal non-interpolation property) with respect to the $\{\mathbf{a}^j\}_{j=1}^k$, if $\{\mathbf{x}^i\}_{i=1}^m$ but no proper subset thereof has the *NI*-property.

In other words, set

$$\mathcal{M} = \{g(\mathbf{x}^1), ..., g(\mathbf{x}^m)\} \tag{1.1}$$

where

$$g(\mathbf{x}) = \sum_{j=1}^{k} f_j(\mathbf{a}^j \cdot \mathbf{x}) \tag{1.2}$$

and the f_j range over all arbitrary functions from $\mathbb{R}$ to $\mathbb{R}$. $\mathcal{M}$ is a linear subspace of $\mathbb{R}^m$, and $\{\mathbf{x}^i\}_{i=1}^m$ has the *NI*-property if and only if $\mathcal{M}$ is a proper subspace of $\mathbb{R}^m$. The following result easily follows from the above definitions.

PROPOSITION 1.1. *Given directions $\{\mathbf{a}^j\}_{j=1}^k \subset \mathbb{R}^d \backslash \{\mathbf{0}\}$, the set $\{\mathbf{x}^i\}_{i=1}^m \subset \mathbb{R}^d$ has the NI-property if and only if there exists a vector $\boldsymbol{\beta} = (\beta_1, ..., \beta_m) \in \mathbb{R}^m \backslash \{\mathbf{0}\}$ such that*

$$\sum_{i=1}^{m} \beta_i f_j(\mathbf{a}^j \cdot \mathbf{x}^i) = 0 \tag{1.3}$$

for all $f_j: \mathbb{R} \to \mathbb{R}$ and each $j = 1, ..., k$. The set $\{\mathbf{x}^i\}_{i=1}^m \subset \mathbb{R}^d$ has the MNI-property if and only if the vector $\boldsymbol{\beta}$ in (1.3) is unique up to multiplication by a constant and has no zero component.

Furthermore, if (1.3) *holds for some* $\boldsymbol{\beta} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$, *then there exists a* $\boldsymbol{\beta}' \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$ *satisfying* (1.3), *i.e., all of whose coordinates are integers.*

Remark. The existence of $\boldsymbol{\beta} \neq \mathbf{0}$ satisfying (1.3) is the existence of a nontrivial linear functional supported on the points $\{\mathbf{x}^i\}_{i=1}^m$ annihilating all functions of the form (1.2). The important part of the proposition is the result that the coefficients of the functional may be chosen as integers.

Proof. Let $\mathcal{M}$ be as given in (1.1). Since $\mathcal{M}$ is a linear subspace of $\mathbb{R}^m$, it does *not* span $\mathbb{R}^m$ if and only if there exists a $\boldsymbol{\beta} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ such that

$$\sum_{i=1}^{m} \beta_i g(\mathbf{x}^i) = 0$$

for all g of the form (1.2). Obviously this is equivalent to (1.3). The fact that the *MNI*-property is equivalent to the uniqueness of the $\boldsymbol{\beta}$ up to multiplication by a constant with no zero component, easily follows from similar reasoning.

It remains to prove that we may choose $\boldsymbol{\beta}' = (\beta'_1, ..., \beta'_m) \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ with $\beta'_i \in \mathbb{Z}$, $i = 1, ..., m$, whenever the points $\{\mathbf{x}^i\}_{i=1}^m$ have the *NI*-property. To this end, set

$$\Lambda_j = \{\mathbf{a}^j \cdot \mathbf{x}^i : i = 1, ..., m\} \qquad \text{for} \quad j = 1, ..., k.$$

The set Λ_j contains say r_j distinct values in $\mathbb{R}$ with $1 \leqslant r_j \leqslant m$. Let us denote these values by $\gamma_1^j, ..., \gamma_{r_j}^j$, i.e.,

$$\Lambda_j = \{\gamma_1^j, ..., \gamma_{r_j}^j\}, \qquad j = 1, ..., k,$$

and $\gamma_l^j \neq \gamma_n^j$ for $l \neq n$. Let

$$h_j^l(y) = \begin{cases} 1 & \text{if} \quad y = \gamma_l^j, \\ 0 & \text{if} \quad y = \gamma_n^j, n \neq l, \end{cases} \qquad \text{for} \quad l = 1, ..., r_j.$$

When considering functions only on the points $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^m$, the functions from $\sum(\mathbf{a}^j)$ are spanned by h_j^l, $l = 1, ..., r_j$. Consequently, (1.3) is equivalent to

$$\sum_{i=1}^{m} \beta_i h_j^l(\mathbf{a}^j \cdot \mathbf{x}^i) = 0 \qquad \text{for} \quad l = 1, ..., r_j. \tag{1.4}$$

Thus $\{\mathbf{x}^i\}_{i=1}^m$ has the *NI*-property if and only if there exists a $\boldsymbol{\beta} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ such that (1.4) holds for $l = 1, ..., r_j, j = 1, ..., k$. We have reduced the interpolation problem to the matrix problem

$$\boldsymbol{\beta} C = \mathbf{0}, \tag{1.5}$$

where $\boldsymbol{\beta}$ is a vector in $\mathbb{R}^m$ and C is an m by $r := \sum_{j=1}^{k} r_j$ matrix. The crucial property of C is the fact that all of its entries, i.e., the $h_j^l(\mathbf{a}^j \cdot \mathbf{x}^i)$, are 0's and 1's and no row or column of C is identically zero. Since Eq. (1.5) has a non-trivial solution $\boldsymbol{\beta}$, and all entries of C are integers, it follows from Cramer's rule that there exists a non-trivial solution $\boldsymbol{\beta}'$, all of whose entries are integers. ∎

For later use we will rewrite (1.4). This equation implies that

$$\sum_i{}' \beta_i = 0, \tag{1.6}$$

if the sum runs over all indices i for which $\mathbf{x}^i \in \Gamma_{\mathbf{a}^j}(\gamma_l^j)$.

Remark. We want to emphasize that our problem is one of characterizing sets of points $\{\mathbf{x}^i\}_{i=1}^m$ satisfying (1.6) for some $\boldsymbol{\beta} \neq \mathbf{0}$, for given $\mathbf{a}^1, ..., \mathbf{a}^k$, and we may totally disregard the interpolation problem.

Remark. We associate a (non-directed) graph to a set $\{\mathbf{x}^i\}_{i=1}^m$ with the *NI*-property. The points $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^m$ are the *vertices* of the graph. A pair $(\mathbf{x}^i, \mathbf{x}^n)$ is an *edge* of the graph if $\mathbf{x}^i, \mathbf{x}^n \in \Gamma_{\mathbf{a}^j}(\alpha)$ for some j, $1 \leqslant j \leqslant k$, and some $\alpha \in \mathbb{R}$. If the points have the *MNI*-property, then at least one edge in each of the k directions is adjacent to each vertex. This is a direct consequence of (1.6). If we have a set of points with at least one edge in each of the k directions adjacent to each vertex ($d = 2, k \geqslant 3$), then it does not necessarily follow that these points have the *MNI* (or *NI*-) property.

This paper is organized as follows. In Section 2 we review the results in the cases $k = 1$ and $k = 2$. In Section 3 we consider any k and identify those $\{\mathbf{x}^i\}_{i=1}^m$ for which interpolation is not always possible over a significant subclass. In Section 4 we detail the case $k = 3$, and see how, in fact, the problem differs from that suggested in Section 3.

2. The Cases $k = 1$ and $k = 2$

For $k = 1$ the problem and its solution are simple. Given $\mathbf{a} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$ ($d \geqslant 2$), what are conditions on the points $\{\mathbf{x}^i\}_{i=1}^m$ in $\mathbb{R}^d$ such that for every choice of $\alpha_1, ..., \alpha_m$ there exists a function $f: \mathbb{R} \to \mathbb{R}$ such that

$$f(\mathbf{a} \cdot \mathbf{x}^i) = \alpha_i, \qquad i = 1, ..., m?$$

The answer is that such functions exist if and only if

$$\mathbf{a} \cdot \mathbf{x}^r \neq \mathbf{a} \cdot \mathbf{x}^s \qquad \text{for all} \quad r \neq s.$$

For $k=2$ we restrict ourselves to $d=2$. Here the problem and its solution were essentially given in Dyn *et al.* [1]. The motivation of that paper was different, and in fact the paper deals with a slightly different problem. In [1] the concern is with interpolation at the $\{\mathbf{x}^i\}_{i=1}^m$ using linear combinations of the functions $\{\|\mathbf{x}-\mathbf{x}^i\|_1\}_{i=1}^m$ (where $\|\cdot\|_1$ is the usual l_1 norm on $\mathbb{R}^2$). However, from this problem the authors were naturally led to a consideration of interpolation by functions of the form

$$g(x, y) = f_1(x) + f_2(y).$$

Formally they considered the special case of the two directions $\mathbf{a}^1 = (1, 0)$ and $\mathbf{a}^2 = (0, 1)$. On the other hand, given any two directions, a linear transformation takes them to the vectors $\mathbf{a}^1$ and $\mathbf{a}^2$ above. Thus the result are effectively the same for any two directions.

Based on [1], we therefore list a series of conditions characterizing points $\{\mathbf{x}^i\}_{i=1}^m \subset \mathbb{R}^2$ with the *NI*-property. For completeness, we also include the proofs. We first need a definition which refers to a generalization of the situation depicted in Fig. 1.1.

DEFINITION 2.1. A set of points $\{\mathbf{v}^i\}_{i=1}^p$ is a *closed path* with respect to the distinct directions $\mathbf{a}^1$ and $\mathbf{a}^2$ if $p=2q$, and for some permutation of the $\{\mathbf{v}^i\}_{i=1}^{2q}$ (which we assume to be as given)

$$\mathbf{a}^1 \cdot \mathbf{v}^{i+1} = \mathbf{a}^1 \cdot \mathbf{v}^i, \qquad i = 1, 3, ..., 2q-1,$$

and

$$\mathbf{a}^2 \cdot \mathbf{v}^{i+1} = \mathbf{a}^2 \cdot \mathbf{v}^i, \qquad i = 2, 4, ..., 2q,$$

where we set $\mathbf{v}^{2q+1} = \mathbf{v}^1$.

Geometrically this simply says that the points $\mathbf{v}^1, ..., \mathbf{v}^p$ and $\mathbf{v}^1$ again form a closed path with edges in alternating directions.

THEOREM 2.1. (Dyn *et al.*[1]). *Given two distinct directions* $\mathbf{a}^1$ *and* $\mathbf{a}^2$ *in* $\mathbb{R}^2$, *the following are equivalent*:

(a) *The set of points* $\{\mathbf{x}^i\}_{i=1}^m$ *has the NI-property.*

(b) *There exists a subset* $\{\mathbf{y}^i\}_{i=1}^s$ *of the* $\{\mathbf{x}^i\}_{i=1}^m$ *such that*

$$|\Gamma_{\mathbf{a}^j}(\alpha) \cap \{\mathbf{y}^i\}_{i=1}^s| \neq 1$$

for $j=1, 2$, *and every* $\alpha \in \mathbb{R}$.

(c) *There exists a subset of the* $\{\mathbf{x}^i\}_{i=1}^m$ *which forms a closed path.*

(d) *There exists a subset* $\{\mathbf{z}^i\}_{i=1}^t$ *of the* $\{\mathbf{x}^i\}_{i=1}^m$ *and* $\varepsilon_i \in \{-1, 1\}$, $i = 1, ..., t$, *such that*

$$\sum_{i=1}^{t} \varepsilon_i f_j(\mathbf{a}^j \cdot \mathbf{z}^i) = 0$$

for every $f_j: \mathbb{R} \to \mathbb{R}$ *and* $j = 1, 2$.

Proof. (a)$\Rightarrow$(b). Assume that the points $\{\mathbf{x}^i\}_{i=1}^m$ have the *NI*-property. By Proposition 1.1 there exists a $\boldsymbol{\beta} \in \mathbb{R}^m \backslash \{\mathbf{0}\}$ such that (1.3) holds for all $f_j: \mathbb{R} \to \mathbb{R}$, and $j = 1, 2$. Let $\{\mathbf{y}^l\}_{l=1}^s$ denote the subset of the $\{\mathbf{x}^i\}_{i=1}^m$ for which $\beta_i \neq 0$. That is, after renumbering

$$\sum_{l=1}^{s} \beta_l' f_j(\mathbf{a}^j \cdot \mathbf{y}^l) = 0$$

for all $f_j: \mathbb{R} \to \mathbb{R}$ and $j = 1, 2$, and $\beta_l' \neq 0$, $l = 1, ..., s$. If $\Gamma_{\mathbf{a}^j}(\alpha) \cap \{\mathbf{y}_l\}_{l=1}^s$ is not empty for some α, then it follows from (1.6) that $\sum_l' \beta_l' = 0$, where l runs over the set of indices for which $\mathbf{y}^l \in \Gamma_{\mathbf{a}^j}(\alpha)$. Hence, the set contains at least two points.

(b)$\Rightarrow$(c). Assume that the set $\{\mathbf{y}^i\}_{i=1}^s$ satisfies (b). Set $\mathbf{z}^1 = \mathbf{y}^1$. By assumption, there exists a $\mathbf{y}^{l_2}$, $l_2 \neq 1$, such that

$$\mathbf{a}^1 \cdot \mathbf{y}^{l_2} = \mathbf{a}^1 \cdot \mathbf{z}^1.$$

Set $\mathbf{z}^2 = \mathbf{y}^{l_2}$. By assumption, there exists a $\mathbf{y}^{l_3}$, $l_3 \neq l_2$, such that

$$\mathbf{a}^2 \cdot \mathbf{y}^{l_3} = \mathbf{a}^2 \cdot \mathbf{z}^2.$$

Set $\mathbf{z}^3 = \mathbf{y}^{l_3}$. Continue in this fashion alternating the directions at each step.

Since we can continue this process indefinitely, but there are only s distinct points $\mathbf{y}^1, ..., \mathbf{y}^s$, we must reach a stage where

$$\mathbf{z}^n \in \{\mathbf{z}^1, ..., \mathbf{z}^{n-1}\}.$$

Assume $\mathbf{z}^l = \mathbf{z}^n$ where $l < n$.

If l and n have the same parity, then the set $\{\mathbf{z}^l, ..., \mathbf{z}^{n-1}\}$ is a closed path with repect to $\mathbf{a}^1$ and $\mathbf{a}^2$. If l and n have opposite parity, then the set $\{\mathbf{z}^{l+1}, ..., \mathbf{z}^{n-1}\}$ is a closed path with respect to $\mathbf{a}^1$ and $\mathbf{a}^2$.

(c)$\Rightarrow$(d). Let $\{\mathbf{z}^i\}_{i=1}^{2q}$ form a closed path (with vertices ordered as in the definition of a closed path). Then

$$\sum_{i=1}^{2q} (-1)^i f_j(\mathbf{a}^j \cdot \mathbf{z}^i) = 0$$

for all $f_j: \mathbb{R} \to \mathbb{R}$, and $j = 1, 2$. For example, for $j = 1$ we have $\mathbf{a}^1 \cdot \mathbf{z}^{2i-1} = \mathbf{a}^1 \cdot \mathbf{z}^{2i}$, $i = 1, ..., q$. Thus

$$-f_1(\mathbf{a}^1 \cdot \mathbf{z}^{2i-1}) + f_1(\mathbf{a}^1 \cdot \mathbf{z}^{2i}) = 0, \qquad i = 1, ..., q$$

for any $f_1: \mathbb{R} \to \mathbb{R}$, and

$$\sum_{i=1}^{2q} (-1)^i f_1(\mathbf{a}^1 \cdot \mathbf{z}^i) = 0.$$

A similar argument holds for $j = 2$.

(d) $\Rightarrow$ (a). This is a consequence of Proposition 1.1. ∎

As a consequence we obtain a characterization of minimal sets with the *NI*-property:

A set of points has the MNI-property if and only if the points form a closed path and $|\Gamma_{\mathbf{a}^j}(\alpha)| \leq 2$ *for all* α *and* $j = 1, 2$.

3. Difference Commensurable Points

In what follows we assume that we are given k distinct directions $\mathbf{a}^1, ..., \mathbf{a}^k$ in $\mathbb{R}^2$. For ease of notation, we assume that $\mathbf{a}^i = (\sin\theta_i, -\cos\theta_i)$ where $0 = \theta_1 < \theta_2 < \cdots < \theta_k < \pi$, and $\mathbf{x} = (x, y)$. This means that

$$\mathbf{a}^i \cdot \mathbf{x} = x \sin\theta_i - y \cos\theta_i$$

is a constant along any straight line which intersects the x-axis with positive angle θ_i. Moreover, set $\mathbf{b}^i := (\cos\theta_i, \sin\theta_i)$.

We will first define what we mean by a *brick*. A brick is determined by the directions $\mathbf{a}^1, ..., \mathbf{a}^k$ and sides of length $\sigma_1, ..., \sigma_k$ ($\sigma_i > 0$, $i = 1, ..., k$). It is a set of 2^k points (vertices) in $\mathbb{R}^2$ with the *NI*-property. (In certain non-generic cases, some of these 2^k points may coincide.) It is constructed as follows, up to translation.

Given $i \in \{1, 2, ..., 2^k\}$ consider the representation of $i - 1$ as a binary number

$$i - 1 = \sum_{j=1}^{k} d_j 2^{j-1}, \tag{3.1}$$

where $d_j = d_j(i)$ and set

$$\mathbf{x}^i = \sum_{j=1}^{k} d_j \sigma_j \mathbf{b}^j.$$

In the case $k=2$ we obtain the vertices of a parallelogram

$$\mathbf{x}^1=\mathbf{0}, \qquad \mathbf{x}^2=\sigma_1\mathbf{b}^1, \qquad \mathbf{x}^3=\sigma_2\mathbf{b}^2, \qquad \mathbf{x}^4=\sigma_1\mathbf{b}^1+\sigma_2\mathbf{b}^2.$$

For $k=3$, the eight points form the vertices of a figure which looks like a projection of a parallelopiped. Hence the name "brick."

Now, we alternately associate with each of the 2^k vertices $\{\mathbf{x}^i\}_{i=1}^{2^k}$ a value $\varepsilon_i \in \{-1, 1\}$. Referring to the binary decomposition (3.1) we set

$$\varepsilon_i=(-1)^{n_i} \qquad \text{with} \quad n_i=\sum_{j=1}^{k} d_j(i). \tag{3.2}$$

As is easily checked, the resulting vector $\boldsymbol{\varepsilon}=(\varepsilon_1, ..., \varepsilon_{2^k})$ has the property that

$$\sum_{i=1}^{2^k} \varepsilon_i g(\mathbf{x}^i)=0$$

for every g of the form (1.2). That is, the vertices of the brick have the NI-property.

In light of the situation depicted in Fig. 1.1, bricks seem to be so natural that it is tempting to ask whether all sets of points with the NI-property contain a subset which can be obtained by taking a (finite) sum of bricks. We will first explain what we mean by the above statement.

A brick B is determined by 2^k vertices and a vector $\boldsymbol{\varepsilon} \in \{-1, 1\}^{2^k}$, as given above. Given r bricks $B_1, ..., B_r$ and numbers $\alpha_1, ..., \alpha_r$, by

$$\sum_{j=1}^{r} \alpha_j B_j$$

we mean the set of points $\{\mathbf{y}^i\}_{i=1}^{p}$ with associated values $\{\gamma_i\}_{i=1}^{p}$, where each $\mathbf{y}^i$ is in at least one of the B_j, $j=1, ..., r$, the value γ_i is the sum of $\alpha_j\varepsilon_l^j$ for $j \in \{1, ..., r\}$ and $l \in \{1, ..., 2^k\}$ such that $\mathbf{x}^l$ in B_j is $\mathbf{y}^i$, and $\gamma_i \neq 0$. A vertex of a brick is ignored if the associated "weight" γ_i is zero. Note two important facts. First,

$$\sum_{i=1}^{p} \gamma_i g(\mathbf{y}^i)=0 \tag{3.3}$$

holds for all g as in (1.2) since (3.3) is obtained as a sum of such equations. Second, among the points $\{\mathbf{y}^i\}_{i=1}^{p}$ we do *not* include those points for which $\gamma_i=0$. For example, assume $k=2$ and we are given two bricks $\{\mathbf{x}^i\}_{i=1}^{4}$ and $\{\mathbf{z}^i\}_{i=1}^{4}$ as given in Fig. 3.1 (with directions parallel to the axes).

If $\mathbf{x}^4=\mathbf{z}^3$, then the resulting B_1+B_2 is given in Fig. 3.2a. If $\mathbf{x}^4=\mathbf{z}^1$, then B_1-B_2 is depicted in Fig. 3.2b.

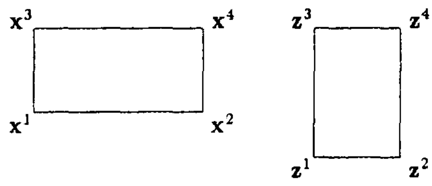

FIGURE 3.1

In this way some of the points may be cancelled, and we always remain with a set of points with the *NI*-property.

Although we did not state it in Theorem 2.1, it is not difficult to ascertain that for $k=2$ every set of points with the *NI*-property contains a subset which is obtained by a sum of bricks (which are parallelograms). The closed paths as specified in Theorem 2.1(c) are obviously sums of parallelograms. For $k=3$, as we shall see in the next section, this is not true in general. There we will find more basic structures than merely the bricks. Nevertheless, for a large class of points $\{\mathbf{x}^i\}_{i=1}^m$ and any $k \geqslant 3$, every set of points with the *NI*-property does contain a subset obtained by taking sums of bricks (of a specific type). We will delineate the additional assumption and prove this result.

DEFINITION 3.1. A set of points $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ is said to be *difference commensurable* (or has the *DC*-property) with respect to a direction $\mathbf{a}$ if there exists a number $\delta > 0$ and *integers* $\{\mu_{ij}\}_{i,j=1}^m$ such that

$$\mathbf{a} \cdot \mathbf{x}^i - \mathbf{a} \cdot \mathbf{x}^j = \mu_{ij}\, \delta$$

for every $i, j = 1, ..., m$.

That is, the points $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ have the *DC*-property with respect to $\mathbf{a}$ with difference δ if all these points lie on the regular grid lines

$$\mathbf{a} \cdot \mathbf{x} = n\delta + v$$

for some v fixed, and $n = 0, \pm 1, \pm 2, ...$

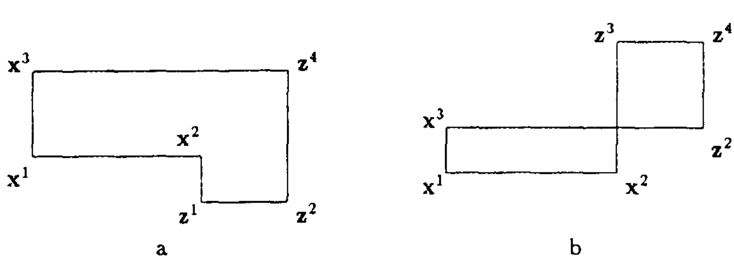

FIGURE 3.2

Given the directions $\mathbf{a}^1, ..., \mathbf{a}^k$ (as above), there are bricks, the vertices of which have the DC-property with respect to $\mathbf{a}^1$ with difference δ. They are given by the previous construction where σ_1 is arbitrary, but for each of the other σ_i's the product $\sigma_i \sin \theta_i$ is an integer multiple of δ. Since we will add bricks anyway, we will use the *elementrary bricks* which satisfy $\sigma_i \sin \theta_i = +\delta$, $i = 2, ..., k$. Since $0 < \theta_2 < \theta_3 < \cdots < \theta_k < \pi$, the positive sign could be chosen. This just implies that for each i,

$$\mathbf{a}^1 \cdot \mathbf{x}^i = m_i \delta \qquad \text{for some } m_i \in \{0, 1, ..., k-1\}. \tag{3.4}$$

In fact, $m_i = \sum_{j=2}^{k} d_j(i)$. Moreover (the equation with maximal m_i),

$$\mathbf{a}^1 \cdot \mathbf{x}^i = (k-1)\,\delta \text{ holds for exactly two vertices.} \tag{3.5}$$

To ease notation we call such bricks *DC-bricks with respect to* $\mathbf{a}^1$ with difference δ. Note that such bricks are uniquely determined up to translation, and the choice of σ_1.

We can now state the main result of this section.

THEOREM 3.1. *Let $\mathbf{a}^1, ..., \mathbf{a}^k$ be k distinct directions in $\mathbb{R}^2$, and assume that the set $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ has the DC-property with respect to $\mathbf{a}^l$, some $l \in \{1, ..., k\}$, with difference δ. Then the points $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ have the NI-property with respect to the directions $\mathbf{a}^1, ..., \mathbf{a}^k$ if and only if a subset of these points may be obtained as a finite sum of DC-bricks with respect to $\mathbf{a}^l$ with difference δ.*

The proof of Theorem 3.1 heavily depends on the following lemma.

LEMMA 3.2. *Assume that the points $\{\mathbf{y}^1, ..., \mathbf{y}^n\}$ have the MNI-property with respect to the distinct directions $\mathbf{a}^1, ..., \mathbf{a}^k$. Further assume that these points have the DC-property with respect to $\mathbf{a}^1$ with difference δ. Then*

$$\max_{i, j \in \{1, ..., n\}} |\mathbf{a}^1 \cdot \mathbf{y}^i - \mathbf{a}^1 \cdot \mathbf{y}^j| \geqslant (k-1)\,\delta.$$

Proof. Consider the convex hull $\mathscr{C}$ of the points $\{\mathbf{y}^i\}_{i=1}^n$. It is a convex polygon. From (1.6) it follows that no line of the form $\mathbf{a}^j \cdot \mathbf{x} = \alpha$ contains exactly one point from $\{\mathbf{y}^i\}_{i=1}^n$ (for $1 \leqslant j \leqslant k$, $\alpha \in \mathbb{R}$). Therefore, exactly two edges of the polygon are parallel to $\mathbf{b}^j$ whenever $1 \leqslant j \leqslant k$. Hence $\mathscr{C}$ is a polygon with exactly $2k$ edges. For each $j \in \{1, ..., k\}$, there are $\alpha_1^j < \alpha_2^j$ such that $\Gamma_{\mathbf{a}^j}(\alpha_1^j)$ and $\Gamma_{\mathbf{a}^j}(\alpha_2^j)$ contain sides of $\mathscr{C}$. Let $\mathscr{C} \cap \Gamma_{\mathbf{a}^j}(\alpha_2^j)$ be the straight line with endpoints $\mathbf{y}_1^j, \mathbf{y}_2^j$, where $\mathbf{a}^1 \cdot \mathbf{y}_2^j > \mathbf{a}^1 \cdot \mathbf{y}_1^j$, $j = 2, ..., k$. Thus $\mathbf{y}_2^j = \mathbf{y}_1^{j+1}$, $j = 2, ..., k-1$. Now

$$\mathbf{a}^1 \cdot \mathbf{y}_2^j - \mathbf{a}^1 \cdot \mathbf{y}_1^j \geqslant \delta$$

since the $\{\mathbf{y}^1, ..., \mathbf{y}^n\}$ have the DC-property with respect to $\mathbf{a}^1$ with difference δ. Summing j over 2, ..., k, we obtain that

$$\mathbf{a}^1 \cdot \mathbf{y}_2^k - \mathbf{a}^1 \cdot \mathbf{y}_1^2 \geqslant (k-1)\,\delta,$$

which proves the lemma. ∎

Proof of Theorem 3.1. One direction is obvious. If a subset of the $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ is obtained as a finite sum of DC-bricks with respect to $\mathbf{a}^l$ with difference δ, then the points $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ have the NI-property with respect to the directions $\mathbf{a}^1, ..., \mathbf{a}^k$.

It remains to prove the converse direction. For convenience, we assume that $l=1$ and set $\mathbf{a} := \mathbf{a}^1$. From Proposition 1.1 we have existence of a vector $\boldsymbol{\beta} \in \mathbb{R}^m \backslash \{\mathbf{0}\}$ such that

$$\sum_{i=1}^{m} \beta_i f_j(\mathbf{a}^j \cdot \mathbf{x}^i) = 0 \tag{3.6}$$

for all $f_j\colon \mathbb{R} \to \mathbb{R}$ and $j = 1, ..., k$.

Let

$$\max\{\mathbf{a} \cdot \mathbf{x}^i - \mathbf{a} \cdot \mathbf{x}^j \colon i, j \in \{1, ..., m\},\ \beta_i, \beta_j \neq 0\} = n\,\delta.$$

From Lemma 3.2, $n \geqslant k-1$. For convenience, assume that

$$\min\{\mathbf{a} \cdot \mathbf{x}^i \colon \beta_i \neq 0\} = 0.$$

Let $i_0, j_0 \in \{1, ..., m\}$ be such that $\mathbf{a} \cdot \mathbf{x}^{i_0} = \mathbf{a} \cdot \mathbf{x}^{j_0} = n\,\delta$, $\beta_{i_0}, \beta_{j_0} \neq 0$, and $\mathbf{x}^{i_0} \neq \mathbf{x}^{j_0}$. It follows from (1.6) that such i_0 and j_0 exist. Let B_{i_0} denote the DC-brick with respect to $\mathbf{a}$ with difference δ, where $\sigma_1 = \|\mathbf{x}^{i_0} - \mathbf{x}^{j_0}\|$, and the points $\mathbf{x}^{i_0}$ and $\mathbf{x}^{j_0}$ are the vertices of the topmost row of B_{i_0}. From (3.4) and (3.5) we conclude that $0 \leqslant \mathbf{a} \cdot \mathbf{y} < n\delta$ holds for any other vertex $\mathbf{y}$ in B_{i_0}. We now add $\pm\beta_{i_0} B_{i_0}$ to the set $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ in the manner previously indicated, i.e., with respect to (3.6). The sign is chosen so that the new coefficient of $\mathbf{x}^{i_0}$ is zero. Since B_{i_0} has the NI-property, we obtain a new set of points $\{\mathbf{x}_1^1, ..., \mathbf{x}_1^{m_1}\}$ with the NI-property. We note further properties:

(a) There exists a $\boldsymbol{\beta}^1 \in \mathbb{R}^{m_1} \backslash \{\mathbf{0}\}$ such that

$$\sum_{i=1}^{m_1} \beta_i^1 f_j(\mathbf{a}^j \cdot \mathbf{x}_1^i) = 0$$

for all $f_j\colon \mathbb{R} \to \mathbb{R}$ and $j = 1, ..., k$.

(b) $\min\{\mathbf{a} \cdot \mathbf{x}_1^i \colon \beta_i^1 \neq 0\} \geqslant 0$.

(c) $\max\{\mathbf{a} \cdot \mathbf{x}_1^i \colon \beta_i^1 = 0\} \leqslant n\,\delta$.

(d) The set of $\mathbf{x}_1^i$ with $\beta_i^1 \neq 0$ and $\mathbf{a} \cdot \mathbf{x}_1^i = n\,\delta$ is a strict subset of the set of $\mathbf{x}^i$ with $\beta_i \neq 0$ and $\mathbf{a} \cdot \mathbf{x}^i = n\delta$, since no new point has been added, while $\mathbf{x}^{i_0}$ has been discarded.

The new set of points may not have these properties in that we may have all the $\beta_i^1 = 0$. If this is so, then we are finished and the theorem is proved. We wish to show that it is this situation, after a finite number of steps, which must occur.

To this end we continue the above process. Since the number of points $\mathbf{x}^i$ for which $\beta_i \neq 0$ and $\mathbf{a} \cdot \mathbf{x}^i = n\,\delta$ is finite, we eventually reach a step l_1, where

$$\max\{\mathbf{a} \cdot \mathbf{x}_{l_1}^i : \beta_i^{l_1} \neq 0\} \leqslant (n-1)\,\delta.$$

At this step we have identified the (given) set with the *NI*-property with maximal level n as a sum of elementary bricks and a set with the *NI*-property with maximal level $n-1$.

By repeating the process we may represent the original set as a sum of elementary bricks and a set with the *NI*-property with maximal level $k-2$. From Lemma 3.2 we conclude that the latter can only be the (trivial) configuration with each $\beta_i = 0$. Hence, we are done.

4. The r-Hexagon and $k=3$

In this section we completely classify all sets of points with the *NI*-property where we are given three distinct directions $\mathbf{a}^1$, $\mathbf{a}^2$, $\mathbf{a}^3$ in $\mathbb{R}^2$. For ease of notation we assume, as in Section 3, that $\mathbf{a}^i = (\sin\theta_i, -\cos\theta_i)$, $i = 1, 2, 3$, where $0 = \theta_1 < \theta_2 < \theta_3 < \pi$. The $\mathbf{a}^1$, $\mathbf{a}^2$, and $\mathbf{a}^3$ are fixed throughout this section. We also define the orthogonal directions $\mathbf{b}^i := (\cos\theta_i, \sin\theta_i)$.

The basic building blocks of sets of points satisfying the *NI*-property are not bricks, but a subset of hexagons. We call a hexagon a *regular hexagon* or an *r-hexagon* if its vertices satisfy the *NI*-property (with respect to $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$). There are r-hexagons, and we can characterize them in various ways. One characterization, up to translation, is the following. Let $A = (0, 0)$, $B = (\sigma, 0)$, $F = (\delta\cos\theta_3, \delta\sin\theta_3)$, where $\sigma, \delta > 0$. (That is, A and B lie on a line orthogonal to $\mathbf{a}^1$, and A and F lie on a line orthogonal to $\mathbf{a}^3$.) Then there is a unique r-hexagon containing the vertices A, B, and F. The remaining three vertices C, D, and E are determined as cutting points of lines. C is the point of intersection of the line through B orthogonal to $\mathbf{a}^2$, and the line through F orthogonal to $\mathbf{a}^1$. E is the point of intersection of the line through B orthogonal to $\mathbf{a}^3$, and the line through

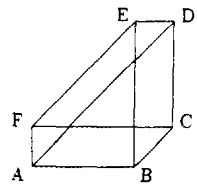

FIGURE 4.1

F orthogonal to $\mathbf{a}^2$. Finally, D is the point of intersection of the line through E orthogonal to $\mathbf{a}^1$, and the line through C orthogonal to $\mathbf{a}^3$ (see Fig. 4.1).

For an algebraic representation observe that the vectors $\mathbf{b}^1$, $\mathbf{b}^2$, and $\mathbf{b}^3$ are linearly dependent, i.e.,

$$t_2\mathbf{b}^2 = t_1\mathbf{b}^1 + t_3\mathbf{b}^3.$$

Since $0 < \theta_2 < \theta_3$, we have $t_1, t_2, t_3 > 0$. Set $\bar{\sigma} := \sigma/t_1$ and $\bar{\delta} := \delta/t_3$. An easy calculation shows that

$$A = \mathbf{0}, \qquad B = \bar{\sigma}t_1\mathbf{b}^1, \qquad C = \bar{\sigma}t_1\mathbf{b}^1 + \bar{\delta}t_2\mathbf{b}^2,$$
$$D = (\bar{\sigma} + \bar{\delta})\, t_2\mathbf{b}^2, \qquad E = \bar{\sigma}t_2\mathbf{b}^2 + \bar{\delta}t_3\mathbf{b}^3, \qquad F = \bar{\delta}t_3\mathbf{b}^3.$$

We associate alternately the weights $+1$ and -1 to the vertices, A, B, C, D, E, F. It follows now that the points form a set with the *NI*-property. If σ and/or δ are negative, a similar construction may be performed, except in the case where $\bar{\sigma} = -\bar{\delta}$. In particular, assume we are given three distinct points $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, and any permutation i, j, k of 1, 2, 3. If $\mathbf{a}^i \cdot \mathbf{x} = \mathbf{a}^i \cdot \mathbf{y}$ and $\mathbf{a}^j \cdot \mathbf{x} = \mathbf{a}^j \cdot \mathbf{z}$, then there is a unique r-hexagon containing the points $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, as vertices, unless $\mathbf{a}^k \cdot \mathbf{y} = \mathbf{a}^k \cdot \mathbf{z}$. In general, any given trapezoid with sides in the direction $\mathbf{b}^1, \mathbf{b}^2$, and $\mathbf{b}^3$, which is not a parallelogram may be completed to an r-hexagon by adding two points.

We will consider *finite sums* of r-hexagons. By this we mean exactly what was meant by a finite sum of bricks. We conjecture that r-hexagons cannot be represented as sums of bricks if the quotient $\bar{\sigma}/\bar{\delta}$ is an irrational number. On the other hand, the converse is true.

LEMMA 4.1. *Let* $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$ *be as above. Then any brick (based on* $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$*) is the sum of at most* 4 *r-hexagons.*

Proof. Let $a, b, c, d, e, d, e, f, g$, and h be the eight vertices of a brick as shown in Fig. 4.2. At the first stage we add four new vertices l, m, n, o to the above. Let m be any point on the line through b and f, i.e., line through b or f in the direction $\mathbf{b}^2$. Let l be the intersection of the line through d in

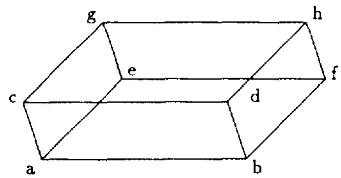

FIGURE 4.2

the direction $\mathbf{b}^2$ and the line through m in the direction $\mathbf{b}^1$. Let n be the intersection of the line through l in the direction $\mathbf{b}^3$ and the line through c in the direction $\mathbf{b}^2$. Let o be the intersection of the line through m in the direction $\mathbf{b}^3$ and the line through n in the direction $\mathbf{b}^1$. It follows that o is also on the line through a in the direction $\mathbf{b}^2$. Thus the points a, b, c, d, l, m, n, and o satisfy the NI-property. In general, these eight vertices are not vertices of a brick. They are the vertices of a brick with a twist, a sort of "Escher brick." Moreover e, f, g, h, l, m, n, and o are also the vertices of an "Escher brick," and our original brick (with its appropriate weights) may be obtained as the difference of these two "Escher bricks" with the appropriate weights.

Consider the "Escher brick" given by the vertices a, b, c, d, l, m, n, and o as obtained above. Depending on the placement of m it may look like Fig. 4.3. Let p be the intersection of the line through nl and the line through ab. Let q be the intersection of the line through om and the line through cd. The vertices a, c, n, o, p, q and the vertices b, d, l, m, p, q each form an r-hexagon. Thus each "Escher brick" is the difference of two r-hexagons with appropriate weights. This proves the lemma.

Remark. Any brick is actually the sum of at most three r-hexagons. But the exact number is immaterial for our purpose and the proof is easier to explain in the case of four r-hexagons. To see that it is the sum of at most three r-hexagons, choose $m = f$ or $m = b$ in the above construction. In this case the brick decomposes into the sum of an "Escher brick" and an r-hexagon.

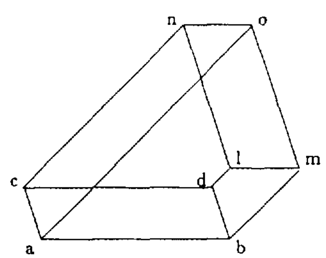

FIGURE 4.3

The main result of this section is:

THEOREM 4.2. *Let $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3$ be three distinct directions in $\mathbb{R}^2$. Then a set of points $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ has the NI-property if and only if a subset of these points may be obtained as a finite sum of r-hexagons.*

Proof. One direction is simple. We therefore assume that we are given a set of points $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ with the *NI*-property. We wish to prove that it contains a (non-trivial) subset which can be obtained as a finite sum of r-hexagons. After reducing the set, if necessary, we may assume that $\{\mathbf{x}^1, ..., \mathbf{x}^m\}$ has the *MNI*-property. By Proposition 1.1 there exists a vector $\boldsymbol{\beta} \in \mathbb{R}^m$, all of whose coefficients are non-zero integers, satisfying

$$\sum_{i=1}^{m} \beta_i f_j(\mathbf{a}^j \cdot \mathbf{x}^i) = 0 \tag{4.1}$$

for all $f_j\colon \mathbb{R} \to \mathbb{R}$, $j = 1, 2, 3$. To avoid ambiguities, we divide the β_i's by their common divisor, and set $M := \sum_{i=1}^{m} |\beta_i|$. We will call M the *size* of the set of points.

Our proof will involve induction on M. The minimal M is $M = 6$ (see the proof of Lemma 3.2), i.e., an r-hexagon. Thus for $M = 6$ the theorem holds.

In the proof of this theorem we will also use the concept of a *cycle*. A cycle is similar to a closed path as given in Definition 2.1. The q distinct points $\mathbf{x}^{i_1}, ..., \mathbf{x}^{i_q}$ (from the $\{\mathbf{x}^i\}_{i=1}^m$) form a cycle if there exist $l_j \in \{1, 2, 3\}$, $j = 1, ..., 2r$, such that

$$\mathbf{a}^{l_j} \cdot \mathbf{x}^{i_j} = \mathbf{a}^{l_j} \cdot \mathbf{x}^{i_{j+1}}, \qquad j = 1, ..., q$$

(here we set $i_{q+1} := i_1$ and $l_{q+1} := l_1$) where

$$\beta_{i_j} \beta_{i_{j+1}} < 0 \qquad \text{and} \qquad l_j \neq l_{j+1}, \qquad j = 1, ..., q. \tag{4.2}$$

Cycles exist since they may be constructed using arguments in the proof of Theorem 2.1.

We associate with a cycle given by the points $\mathbf{x}^{i_1}, ..., \mathbf{x}^{i_{2r}}$ the *vector of weights* $(\operatorname{sgn} \beta_{i_1}, ..., \operatorname{sgn} \beta_{i_{2r}})$. The *length* of a cycle is the number of vertices or sides, i.e., q in the above. We will sometimes consider a cycle from the point of view of the consecutive directions, i.e., $(\mathbf{a}^{l_1}, ..., \mathbf{a}^{l_{2r}})$ in the above example.

The idea of the proof of Theorem 4.1 is the following. If there is a cycle of length 4, we show how to decrease the size M, and thus show the result by induction on M. If there is no cycle of length 4, we show how to alter a cycle so that either M is decreased, or M remains constant, but the length of the cycle is decreased. In this way we eventually decrease M since we will arrive at a cycle of length 4. Of course, all changes made involve adding r-hexagons.

The proof of the theorem is somewhat lengthy and technical. For this reason we divide it by separating out one more lemma.

LEMMA 4.3. *Assume that the* $\{\mathbf{x}^i\}_{i=1}^m$ *as above contain a cycle of length* 4. *Then we can add r-hexagons so as to decrease the size M.*

Proof. If there exists a cycle of length 4 then we label its vertices by $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4$. After choosing the starting point of the cycle we may assume that the directions $\{\mathbf{a}^{q_1}, \ldots, \mathbf{a}^{q_4}\}$ have either the form $\{\mathbf{a}^i, \mathbf{a}^j, \mathbf{a}^i, \mathbf{a}^j\}$ or $\{\mathbf{a}^i, \mathbf{a}^j, \mathbf{a}^k, \mathbf{a}^j\}$, where the $\{i, j, k\}$ are a permutation of $\{1, 2, 3\}$.

Case 1. $\mathbf{a}^i\mathbf{a}^j\mathbf{a}^i\mathbf{a}^j$. The $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4$ are the vertices of a parallelogram. Let $k \in \{1, 2, 3\}\backslash\{i, j\}$. Along the line through $\mathbf{x}^1$ in the direction $\mathbf{a}^k$, there exists an $\mathbf{x}^l$, $l \in \{5, \ldots, m\}$ with $\beta_l\beta_1 < 0$. There is a unique brick containing the vertices $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4$, and $\mathbf{x}^l$. It is obtained by translating the vertices $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4$ in the direction $\mathbf{a}^k$ so that $\mathbf{x}^1$ is mapped onto $\mathbf{x}^l$. This brick contains three new vertices. Multiplying by -1, if necessary, we can assume that the weight at the vertex $\mathbf{x}^i$ is $-\operatorname{sgn}\beta_i$, for $i = 1, 2, 3, 4, l$. We add this brick to our original set. Each β_i, $i \in \{1, 2, 3, 4, l\}$, is reduced by 1 in absolute value, while three (perhaps) new vertices have been added with weights 1 in absolute value. The size of this new set is therefore at most $M-2$. Applying Lemma 4.1 proves this case.

Case 2. $\mathbf{a}^i\mathbf{a}^j\mathbf{a}^k\mathbf{a}^j$. The points $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4$ are the vertices of a trapezoid. From the construction of r-hexagons at the beginning of this section we know that we can obtain an r-hexagon by adding two vertices. From (4.2) we conclude that, after multiplication by -1 if necessary, the weight at the vertex $\mathbf{x}^i$ is $-\operatorname{sgn}\beta_i$, $i = 1, 2, 3, 4$. Adding this r-hexagon to our original set, we simultaneously decrease the size M by 4, while increasing it by at most 2 (because of the additional two vertices of the r-hexagon). Therefore, also in this case the size of the new set is at most $M-2$.

Proof of Theorem 4.2 (Continued). In order to complete the induction argument via the size M, it is necessary that we consider the case where all cycles are of length at least 6. Let C be a cycle of minimal length. In the first two cases we will consider four consecutive directions in C and show how to alter C so that M is decreased, or we obtain a cycle of length less than the length of C. In what follows i, j, k represent any permutation of 1, 2, 3.

Case 1. $\mathbf{a}^i\mathbf{a}^j\mathbf{a}^i\mathbf{a}^k$. Let $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5$ be the five consecutive vertices connected via the above directions, respectively. That is, $\mathbf{a}^i \cdot \mathbf{x}^1 = \mathbf{a}^i \cdot \mathbf{x}^2$,

$\mathbf{a}^j \cdot \mathbf{x}^2 = \mathbf{a}^j \cdot \mathbf{x}^3$, etc.... We claim that there is a (unique) r-hexagon containing the vertices $\mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4$. We have $\mathbf{a}^j \cdot \mathbf{x}^2 = \mathbf{a}^j \cdot \mathbf{x}^3$ and $\mathbf{a}^i \cdot \mathbf{x}^3 = \mathbf{a}^i \cdot \mathbf{x}^4$. Such an r-hexagon exists unless $\mathbf{a}^k \cdot \mathbf{x}^2 = \mathbf{a}^k \cdot \mathbf{x}^4$. However, if $\mathbf{a}^k \cdot \mathbf{x}^2 = \mathbf{a}^k \cdot \mathbf{x}^4$, then since $\mathbf{a}^k \cdot \mathbf{x}^4 = \mathbf{a}^k \cdot \mathbf{x}^5$, it follows that $\mathbf{a}^k \cdot \mathbf{x}^2 = \mathbf{a}^k \cdot \mathbf{x}^5$ and thus there is a cycle exactly like the cycle C except that we replace $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5$ by $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^5$. However, this contradicts the minimality of the length of C.

Now, if $\mathbf{x}^1$ or $\mathbf{x}^5$ is a vertex of this r-hexagon, then adding (or subtracting) this r-hexagon to our original set reduces the size M by at least 2, and we can apply the induction argument.

Assume that neither $\mathbf{x}^1$ nor $\mathbf{x}^5$ is a vertex of this r-hexagon. From the construction of the r-hexagon there exists a vertex $\mathbf{y}$ satisfying $\mathbf{a}^i \cdot \mathbf{x}^2 = \mathbf{a}^i \cdot \mathbf{y}$ and $\mathbf{a}^k \cdot \mathbf{x}^4 = \mathbf{a}^k \cdot \mathbf{y}$. Adding (or subtracting) this r-hexagon does not increase the size M. (We subtract at least 3 and add at most 3 to M.) However, such an addition (subtraction) permits us to replace $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5$ in the cycle C by $\mathbf{x}^1, \mathbf{y}, \mathbf{x}^5$. (Note that $\mathbf{a}^i \cdot \mathbf{x}^1 = \mathbf{a}^i \cdot \mathbf{x}^2 = \mathbf{a}^i \cdot \mathbf{y}$ and $\mathbf{a}^k \cdot \mathbf{y} = \mathbf{a}^k \cdot \mathbf{x}^4 = \mathbf{a}^k \cdot \mathbf{x}^5$.) Thus the cycle of minimal length of this new set is at least 2 less than it was.

Case 2. $\mathbf{a}^i\mathbf{a}^j\mathbf{a}^i\mathbf{a}^j$. Let $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5$ be as in Case 1, i.e., $\mathbf{a}^i \cdot \mathbf{x}^1 = \mathbf{a}^i \cdot \mathbf{x}^2$, $\mathbf{a}^j \cdot \mathbf{x}^2 = \mathbf{a}^j \cdot \mathbf{x}^3$, etc.... We construct a brick containing the vertices $\mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4$ as follows. First, let $\mathbf{y}$ satisfy $\mathbf{a}^i \cdot \mathbf{x}^2 = \mathbf{a}^i \cdot \mathbf{y}$ and $\mathbf{a}^j \cdot \mathbf{x}^4 = \mathbf{a}^j \cdot \mathbf{y}$. The $\mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{y}$ are the vertices of a parallelogram. From (4.1) there exists an $\mathbf{x}^l$, $l \in \{1, ..., m\} \backslash \{3\}$ such that $\beta_3 \beta_l < 0$ and $\mathbf{a}^k \cdot \mathbf{x}^3 = \mathbf{a}^k \cdot \mathbf{x}^l$. From the position of $\mathbf{x}^3$ vis-à-vis $\mathbf{x}^2$ and $\mathbf{x}^4$, we have that $l \notin \{2, 4\}$. Furthermore $l \notin \{1, 5\}$ since $\beta_3 \beta_1 > 0$ and $\beta_3 \beta_5 > 0$. Thus $l \in \{6, ..., m\}$. There exists a brick containing the vertices $\mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{y}$, and $\mathbf{x}^l$. (If $\mathbf{y} = \mathbf{x}^l$, the resulting brick is an r-hexagon. The argument remains the same.) We obtain this brick by translating the parallelogram with vertices $\mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{y}$ in the direction $\mathbf{b}^k$ so that $\mathbf{x}^3$ is projected onto $\mathbf{x}^l$. We add (or subtract) this brick to our original set. Since at least four vertices of the brick were part of the original set, with weights of the appropriate sign, the size M is not increased. If M is decreased, we are finished. Assume M is not decreased. Thus $\mathbf{y}$ is neither $\mathbf{x}^1$ nor $\mathbf{x}^5$. Now $\mathbf{a}^i \cdot \mathbf{x}^1 = \mathbf{a}^i \cdot \mathbf{x}^2 = \mathbf{a}^i \cdot \mathbf{y}$, and $\mathbf{a}^j \cdot \mathbf{y} = \mathbf{a}^j \cdot \mathbf{x}^4 = \mathbf{a}^j \cdot \mathbf{x}^5$. We can therefore replace $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5$ in the cycle C by $\mathbf{x}^1, \mathbf{y}, \mathbf{x}^5$ (the signs alternate as they must). We thus reduce the minimal length by at least 2.

Case 3. Assume that a minimal cycle C does not contain four consecutive directions as given by Cases 1 or 2 (recall that there is no starting point or forward or backward to C). The only remaining case is therefore given by C of the form $..., \mathbf{a}^i, \mathbf{a}^j, \mathbf{a}^k, \mathbf{a}^i, \mathbf{a}^j, \mathbf{a}^k, ...$, i.e., repeats of $\mathbf{a}^i, \mathbf{a}^j, \mathbf{a}^k$.

Let $\mathbf{a}^i \cdot \mathbf{x}^1 = \mathbf{a}^i \cdot \mathbf{x}^2$, $\mathbf{a}^j \cdot \mathbf{x}^2 = \mathbf{a}^j \cdot \mathbf{x}^3$, $\mathbf{a}^k \cdot \mathbf{x}^3 = \mathbf{a}^k \cdot \mathbf{x}^4$, ect.... There exists a brick (which may reduce to an r-hexagon) containing the vertices $\mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5$. The remaining four vertices $\mathbf{y}^1, \mathbf{y}^2, \mathbf{y}^3, \mathbf{y}^4$ are constructed as follows. Let $\mathbf{y}^1$ satisfy $\mathbf{a}^i \cdot \mathbf{x}^3 = \mathbf{a}^i \cdot \mathbf{y}^1$ and $\mathbf{a}^k \cdot \mathbf{x}^5 = \mathbf{a}^k \cdot \mathbf{y}^1$. Let $\mathbf{y}^2$ satisfy $\mathbf{a}^i \cdot \mathbf{x}^2 = \mathbf{a}^i \cdot \mathbf{y}^2$ and $\mathbf{a}^j \cdot \mathbf{y}^1 = \mathbf{a}^j \cdot \mathbf{y}^2$. Let $\mathbf{y}^3$ satisfy $\mathbf{a}^j \cdot \mathbf{x}^5 = \mathbf{a}^j \cdot \mathbf{y}^3$ and $\mathbf{a}^k \cdot \mathbf{y}^2 = \mathbf{a}^k \cdot \mathbf{y}^3$. Let $\mathbf{y}^4$ satisfy $\mathbf{a}^k \cdot \mathbf{x}^2 = \mathbf{a}^k \cdot \mathbf{y}^4$ and $\mathbf{a}^j \cdot \mathbf{x}^4 = \mathbf{a}^j \cdot \mathbf{y}^4$. Or alternatively, define $\mathbf{y}^1$ as above and translate the parallelogram with vertices $\mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5, \mathbf{y}^1$ in the direction $\mathbf{b}^j$ so that $\mathbf{x}^3$ projects onto $\mathbf{x}^2$. Then $\mathbf{y}^1$ projects onto $\mathbf{y}^2$, $\mathbf{x}^5$ projects onto $\mathbf{y}^3$, and $\mathbf{x}^4$ projects onto $\mathbf{y}^4$. This brick may reduce to an r-hexagon. However, this has no real effect on the argument. Adding (or subtracting) this brick does not increase M. If M is not decreased, then $\mathbf{x}^1 \neq \mathbf{y}^2$ and $\mathbf{x}^6 \neq \mathbf{y}^3$. In this case $\mathbf{x}^1, \mathbf{y}^2, \mathbf{y}^3, \mathbf{x}^6$ can replace $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, \mathbf{x}^5, \mathbf{x}^6$ in the above cycle. In this way the length of the minimal cycle is decreased.

ACKNOWLEDGMENTS

The first author would like to thank R. Stöcker and H. Zieschang for many discussions.

REFERENCE

1. N. DYN, W. A. LIGHT, AND E. W. CHENEY, Interpolation by piecewise-linear radial basis functions, *J. Approx. Theory* **59** (1989), 202–223.

Printed in Belgium
Uitgever: Academic Press, Inc.
Verantwoordelijke uitgever voor België:
Hubert Van Maele
Altenastraat 20, B-8310 Sint-Kruis