Balancing Sets of Vectors

*Noga Alon**

Department of Mathematics
Tel Aviv University
Ramat Aviv, Tel Aviv
and
Bell Communications Research
Morristown, NJ 07960

E. E. Bergmann

AT&T Bell Laboratories
Allentown, PA 18103

D. Coppersmith

IBM Research
Yorktown Heights, NY 10598

A. M. Odlyzko

AT&T Bell Laboratories
Murray Hill, NJ 07974

1. Introduction

Let $K(n, d)$ denote the minimal k for which there exist ± 1 vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of length n such that for any ± 1 vector $\mathbf{w}$ of length n, there is an i, $1 \le i \le k$, such that $|\mathbf{v}_i \cdot \mathbf{w}| \le d$, where $\mathbf{v} \cdot \mathbf{w}$ denotes the usual inner product of two vectors. Since $\mathbf{v} \cdot \mathbf{w} \equiv n \pmod 2$ for any two ± 1 vectors $\mathbf{v}$ and $\mathbf{w}$ of length n, $K(n, 0)$ is defined only for even n, while $K(n, d)$ for $d \ge 1$ is well-defined for all n. A very simple and surprising construction of Knuth [9] shows that $K(n, 0) \le n$ for even n. The main result of this note is that for n even, $K(n, 0) = n$. More generally, Knuth's construction shows that $K(n, d) \le \lfloor cn/(d+1) \rfloor c$ for $n \equiv d \pmod 2$, and we show that this construction is best possible.

The reason Knuth's result is surprising is that it might appear that if one tries to select ± 1 vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ so as to minimize

$$\max_{w} \min_{1 \le i \le n} |\mathbf{v}_i \cdot \mathbf{w}| \tag{1.1}$$

(where the maximum is over all ± 1 vectors $\mathbf{w}$ of length n), then the best choice to make is to take the $\mathbf{v}_i$ to be pairwise orthogonal; i.e., as the rows of a Hadamard matrix of order n. Hadamard matrices are conjectured to exist for $n = 1, 2$, and all those n divisible by 4, but this is not known to be true in general [8]. When the $\mathbf{v}_i$ are chosen as the rows of a Hadamard matrix, it is easy to show that the maximum in (1.1) is $\le n^{1/2}$.

In many cases this bound is not best possible (for one thing, $n^{1/2}$ is usually not an integer, and one can even show that some Hadamard matrices give bounds for (1.1) that are smaller than the largest integer $\le n^{1/2}$ that is $\equiv n \pmod 2$). However, for Sylvester matrices of order $n = 4^m$ [8] it can be shown that the $n^{1/2}$ bound is tight, since they have a ± 1 eigenvector.

Our upper bound for $K(n, d)$ is obtained from a simple modification of Knuth's construction [9] which shows $K(n, 0) \le n$ for n even. It is so simple that we present it in the introduction.

Theorem 1. If $n > 0$, $d \ge 0$, $n \equiv d \pmod 2$, then

$$K(n, d) \le \left\lceil \frac{n}{d+1} \right\rceil , \tag{1.2}$$

where $\lceil x \rceil$ is the least integer $\ge x$.

Proof. We first consider $d = 0$, $n \equiv 0 \pmod 2$. Consider the following $n+1$ by n matrix

$$\begin{array}{cccccc} 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & -1 & 1 & \cdots & 1 & 1 \\ & \vdots & & & & \\ -1 & -1 & -1 & \cdots & -1 & 1 \\ -1 & -1 & -1 & \cdots & -1 & -1 \end{array} \tag{1.3}$$

and let $\mathbf{v}_i$, $1 \le i \le n+1$, denote the vector that forms the i-th row of (1.3). We claim that for any ± 1 vector $\mathbf{w}$ of length n, $\mathbf{w} \cdot \mathbf{v}_i = 0$ for some $i \le n$. To see this, note that $\mathbf{w} \cdot \mathbf{v}_1 = -\mathbf{w} \cdot \mathbf{v}_{n+1}$, while $\mathbf{w} \cdot \mathbf{v}_i = \mathbf{w} \cdot \mathbf{v}_{i+1} \pm 2$ for each $i \le n$, so, since $\mathbf{w} \cdot \mathbf{v}_i \equiv 0 \pmod 2$ for all i, there is at least one $i \le n$ such that $\mathbf{w} \cdot \mathbf{v}_i = 0$ (this is something like a discrete mean value theorem). If $\mathbf{w} \cdot \mathbf{v}_{n+1} = 0$, then also $\mathbf{w} \cdot \mathbf{v}_1 = 0$, which establishes our claim that $\mathbf{w} \cdot \mathbf{v}_i = 0$ for some $i \le n$. This completes the proof that $K(n, 0) \le n$ for $n \equiv 0 \pmod 2$.

The bounds for the other $K(n, d)$ are obtained by selecting the vectors $\mathbf{v}_1$, $\mathbf{v}_{d+2}$, $\mathbf{v}_{2d+3}$, . . . , $\mathbf{v}_{r(d+1)+1}$, where $r = \lceil n/(d+1) \rceil - 1$. Since for any ± 1 vector $\mathbf{w}$ and any j, $0 \le i < r$,

$$|\mathbf{w} \cdot \mathbf{v}_{j(d+1)+1} - \mathbf{w} \cdot \mathbf{v}_{(j+1)(d+1)+1}| \le 2d + 2 ,$$

an argument similar to the one above shows that $|\mathbf{w} \cdot \mathbf{v}_{j(d+1)+1}| \le d$ for at least one j, $0 \le j \le r$.

■

Our main result is a proof, given in Section 2, that the construction of Theorem 1 is optimal.

Theorem 2. If $n > 0$, $d \ge 0$, $n \equiv d$ (mod 2), then

$$K(n, d) \geq lc\frac{n}{d+1}rc\ , \tag{1.4}$$

so that $K(n, d) = lcn/(d+1)rc$.

The proof we give uses only elementary linear algebra, and is similar to the proof used in Appendix of [1]. Another, but still related, proof can be given using commutation algebra and Hilberts' Nullstellinsatz [7].

The proof of Theorem 2 can be used to prove the following mass general result.

Theorem 3. Suppose $n > 0$, *and let D be an arbitrary set of integers, where each* $d \in D$ *satisfies* $n \equiv d \pmod 2$. *Let V be a set of* ± 1 *vectors of length n, such that for every* ± 1 *vector* $\mathbf{u}$ *of length n, there is a* $\mathbf{v} \in V$ *with* $\mathbf{u} \cdot \mathbf{v} \in D$. *Then*

$$|V| \geq n/|D|\ . \tag{1.5}$$

Theorem 2 is the special case of this theorem, where $D = \{0, \pm 2, \pm 4\ , \ldots\ , \pm d\}$, as $D = \{\pm 1, \pm 3\ , \ldots\ , \pm d\}$. In general, (1.5) can be very far from best possible. It would be interesting to find the best possible bound for various sets *D*.

It is possible to generalize the problem considered here and consider balancing families of vectors whose components are ± 1, $\pm i$, or more generally, with roots of unity for some fixed *r*. Some preliminary results in this direction have been obtained by the last two authors and P. Shor.

Our balancing vector problem can be rephrased in terms of an extremal problem for subsets of a set, with a ± 1 vector $(u_1\ , \ldots\ , u_n)$ of length *n* corresponding to a subset *A* of $\{1\ , \ldots\ , n\}$, with $j\ \varepsilon\ A$ if and only if $u_j = 1$. F. Galvin recently posed a problem in this setting that is similar to ours. He asked for the determination of the minimal integer

$m = m(n)$ such that there exist subsets $A_1, \ldots, A_m$ of $\{1, \ldots, 4n\}$, each of size $2n$, such that for any subset $B \subseteq \{1, \ldots, 4n\}$ with $2n$ elements there is at least one i, $1 \leq i \leq m$, with $A_i \cap B$ having n elements. Frankl and Rödl [5] have noticed that if one defines $A_i = \{i, i+1, \ldots, i + 2n-1\}$ for $1 \leq i \leq 2n$, then it is easy to see that these A_i have the right property, and so $m \leq 2n$. Frankl and Rödl proved that $m > \varepsilon n$ for some fixed $\varepsilon > 0$. They conjectured initially that $m = 2n$, but this was shown to be false by Markert and West (unpublished), who found using Madamard matrices that $m(2) \leq 3$ and $m(4) \leq 7$. On the other hand, the recent proof [4] of Itos' conjecture (that every linear subspace of dimension $2n + 1$ of the space $GF(2)^{4n}$ has a vector of Manning weight exactly $2n$) leads to a proof that $m(n) = 2n$ for n odd. Although the Galvin problem is somewhat similar to that of determining $K(2n, 0)$, there does not seem to be any simple dependence between them.

A still different conjecture that is related to the determination of $K(2n, 0)$ is due to S. Poljak (unpublished). It states that if G is the graph with vertices equal to the 2^n binary vectors of length n, and edges between vertices that are at Hamming distance 1, then the minimal number of hyperplanes that cut each edge is exactly n. It is easy to see that n hyperplanes can be found that cut each edge, but that $\geq n$ hyperplanes as changes needed has only been shown for a few values of n by Z. Füredi and S. Poljak.

Our research on balancing sets of vectors was motivated by a problem in optical data communication which also arises in other communication areas. This problem is described in Section 3.

2. Proof of Theorem 2

Put $N = \{1, 2, \ldots, n\}$, and let U be the set of all ± 1 vectors of length n. A vector $\mathbf{u} \in U$ is *even* if it has an even number of –1's, otherwise it is *odd*.

Lemma 2.1. Let $P(\mathbf{y}) = P(y_1, y_2, \ldots, y_n)$ *be a multilinear polynomial of degree less than* $n/2$ *over the reals, i.e.,* $P = \Sigma\left\{\alpha_X \cdot \prod_{i \in X} y_i : X \subset N, |X| < n/2\right\}$ *where* $\alpha_X \in R$. *If* $P(\mathbf{y}) = 0$ *for every even vector* $\mathbf{y} \in U$ *than* $P \equiv 0$ *(i.e.,* $\alpha_X = 0$ *for all X). Similarly, if* $P(\mathbf{y}) = `0$ *for every odd vector* $\mathbf{y} \in U$, *then* $P \equiv 0$.

Proof. We can prove the even case. (The odd case is analogous.) By the hypotheses, for every even subset $Y \subset N$ we have

$$\Sigma\left\{\alpha_X(-1)^{|Y \cap X|} : X \subset N, |X| < n/2\right\} = 0 .$$

It thus suffices to check that the columns of the matrix

$$A = `(\alpha_{Y,X}) = ((-1)^{|Y \cap X|}) \; Y \subset N, \; |Y| \text{even}; \; X \subset N, |X| < n/2$$

are linearly independent (over the reals). One can easily check that

$$(A^T A)_{X_1, X_2} = \Sigma\left\{(-1)^{|Y \cap X_1| + |Y \cap Y_2|} : Y \subset N, |Y| \text{ even}\right\}$$

$$= \Sigma\left\{(-1)^{|Y \cap (X_1 \oplus X_2)|} : Y \subset N, |Y| \text{ even}\right\} .$$

The last sum is 2^{n-1} for $X_1 = X_2$. For $X_1 \neq X_2$, since $X_1 \oplus X_2 \neq N, \varnothing$ this sum is

$$\Sigma\left\{(-1)^{|A|}\ 2^{n-|X_1\oplus X_2|-1}:\ A\subset X_1\oplus X_2\right\}=\text{‘}0\ .$$

We conclude that A^TA is nonsingular and hence that A has full column rank. This completes the proof of the lemma.

We can now prove Theorem 2. For simplicity we consider the case $n \equiv 0 \pmod 4$ and $d \equiv 0 \pmod 4$. (The proofs for the other cases are analogous.) Let $V \subset U$ be a set of vectors such that for every $u \in U$ there is a $v \in V$ with $|\mathbf{v}\cdot\mathbf{u}| \le d$. We must show that $|V| \ge n/(d+1)$. Let V_0 be the set of all even vectors of V and let V_1 be the set of all odd vectors of V.

Consider the following polynomial in $\mathbf{y} = (y_1, y_2, \ldots, y_n)$;

$$P(\mathbf{y}) = \prod_{v\in V_0}(\mathbf{v}\cdot\mathbf{y})\prod_{v\in V_1}(2^2-(\mathbf{v}\cdot\mathbf{y})^2)\prod_{v\in V_0}(4^2-(\mathbf{v}\cdot\mathbf{y})^2)\cdot\cdots\cdot\prod_{v\in V_0}(d^2-(\mathbf{v}\cdot\mathbf{y})^2)\ .$$

Since $n \equiv 0 \pmod 4$, $(\mathbf{v}_1\cdot\mathbf{v}_2) \equiv 0 \pmod 2$ for all $\mathbf{v}_1, \mathbf{v}_2 \in U$. Also, as is easily checked, for every $\mathbf{v}_1, \mathbf{v}_2 \in U$, $(\mathbf{v}_1\cdot\mathbf{v}_2) \equiv 0 \pmod 4$ if and only if both $\mathbf{v}_1$ and $\mathbf{v}_2$ are even or both are odd. Otherwise, $(\mathbf{v}_1\cdot\mathbf{v}_2) \equiv 2 \pmod 4$. By assumption, for every even vector $\mathbf{y} = (y_1, \ldots, y_n) \in U$ there is a $\mathbf{v} \in V$ with $|\mathbf{v}\cdot\mathbf{y}| \le d$, i.e., $\mathbf{v}\cdot\mathbf{y} \in \{0, \pm2, \pm4, \ldots, \pm d\}$. Since $\mathbf{v}\cdot\mathbf{y}$ can be in $\{0, \pm4, \ldots, \pm d\}$ only if $\mathbf{v}$ is even, and can be in $\{\pm2, \pm6, \ldots, \pm(d-2)\}$ only if $\mathbf{v}$ is odd, we conclude that for every even $\mathbf{y} \in U$, $P(\mathbf{y}) = 0$.

Let $\overline{P}(\mathbf{y})$ be the multilinear polynomial obtained from $P(\mathbf{y})$ by replacing repeatedly each occurrence of y_i^2 in the standard representation of P as a sum of monomials by 1. Clearly $\overline{P}(\mathbf{y}) = P(\mathbf{y})$ for all $\mathbf{y} \in U$, and deg $\overline{P} \le$ deg P. Since $\overline{P}(\mathbf{y}) = 0$ for every

even $\mathbf{y} \in U$ and $\overline{P} \neq 0$, (since $\overline{P}(\mathbf{y}) \neq 0$ for every odd $\mathbf{y} \in U$), we conclude, by Lemma 2.1, that deg $P \geq$ deg $\overline{P} \geq n/2$. Therefore

$$\deg P = |V_0| + \frac{d}{2}(|V_0| + |V_1|) \geq n/2 . \tag{2.1}$$

We now repeat the above process for odd vectors $\mathbf{y} \in U$. Define

$$G(\mathbf{y}) = \prod_{v \in V_1} (v \cdot \mathbf{y}) \prod_{v \in V_0} (2^2 - (v \cdot \mathbf{y})^2) \cdot \cdots \cdot \prod_{v \in V_1} (d^2 - (v \cdot \mathbf{y})^2) .$$

Observe, as before, that $G(\mathbf{y}) = 0$ for every odd $\mathbf{y} \in U$ and $G(\mathbf{y}) \neq 0$ for every even $\mathbf{y} \in U$. This implies, as before, that

$$|V_1| + \frac{d}{2}(|V_0| + |V_1|) \geq n/2 . \tag{2.2}$$

The summation of (2.1) and (2.2) implies that $(d + 1)\,|V| \geq n$ and completes the proof of Theorem 1.1.

■

Remarks: Another proof of Theorem 2 can be given using the fact that the polynomial $P(\mathbf{y})$ vanishes on the zero set of the ideal generated by $y_1^2 - 1, \ldots, y_n^2 - 1, y_1 y_2, \ldots, y_n - 1$. Hilbert's Nullstellinsatz can then be used to show that P vanishes identically.

3. Applications

The problem which motivated the research reported in this note arose in optical data communications, and is similar to the problem that motivated Knuth's investigation [9]. We will regard the basic signal bits used in an optical channel as ±1's. In many applications, such as data links, where the data rates are very high but the electronics

have to be simple to be economical, it is desirable to encode information so that the transmitted stream will have a low d.c. component; i.e., in any long block, there will be about as many $+1$'s as -1's. For other problems and solutions in the design of simple balanced signal sets, see [2, 3, 6, 10, 11].

It is possible to encode each -1 as $(+1\ -1)$ and each $+1$ as $(-1\ +1)$, but this code (called Manchester code) has information rate of $1/2$ and so is very inefficient. A more efficient method is to take blocks of m bits and encode them in blocks of $n = m + k$ bits, which have exactly as many $+1$'s as -1's. For example, since $2^{17} = 131,072$ and $\binom{20}{10} = 184,756$, we could take $m = 17$ and $n = 20$. Similarly, since $2^{33} = 8,589,934,592$ while $\binom{36}{18} = 9,075,135,300$, we could take $m = 33$ and $n = 36$. However, no way is known to efficiently encode or decode information using these codes [12]. The best algorithm for assigning a unique integer between 1 and $\binom{2r}{r}$ to each subset of size r of a set of size $2r$ takes either on the order of r operations on integers of about r bits when about r^2 stored r-bit values are available, or about r^2 operations when about r stored values are available. Thus this scheme would probably be practical only for very small values of m and n, where direct memory lookup is feasible, for example for $m = 6$ and $n = 8$.

Proposals have been made to use scramblers to transform the data stream into a more random-looking sequence which will therefore have a small average d.c. component. Scramblers using a linear feedback shift register can be implemented very efficiently. However, there are data patterns for which scramblers will fail, and in data communications, where long repetitions of a particular data sequence are fairly common,

it seems reasonable to try to protect against even such pathological cases.

Knuth's and our proposed solution to the d.c. component problem is to use balancing sets of vectors. In the very simplest case, corresponding to the construction for $R(m, \lfloor m/2 \rfloor)$, we could take blocks of m data bits $x_1 \cdots x_m$ and encode them into $n = m+1$ bits $y_1 \cdots y_n$ as follows: If $|\sum x_i| \le m/2$, then $y_i = x_i$ for $1 \le i \le m$ and $y_n = 1$, while if $|\sum x_i| > m/2$, then $y_i = -x_i$ for $1 \le i \le \lfloor m/2 \rfloor$, $y_i = x_i$ for $\lfloor m/2 \rfloor < i \le m$, and $y_n = -1$. It then follows that $|\sum y_i| \le (n+1)/2$, so the d.c. component is substantially decreased. At the other extreme, we can use (for m even) the construction that achieves $R(m, 0)$. In this case the data block $(x_1, \ldots, x_m)$ would be encoded as $(x_1\, v_{i1}, \ldots, x_m\, v_{im}, z_1, \ldots, z_k)$, where $(v_{i1}, \ldots, v_{im})$ is the vector from (1.3) that's orthogonal to $(x_1, \ldots, x_m)$, and $z_1, \ldots, z_k$ are ± 1's indicating the value of i. Since we need to distinguish only between m different values of i, we can use k on the order of $\log_2 n$, even if we select the z_i so that $z_1 + \cdots + z_k = 0$ in order for each transmitted block to have zero d.c. component. (For $m \le 924$, for example, we could take $k = 12$, as $\binom{12}{6} = 924$, so that encoding and decoding could be implemented by easy table lookup, with addresses at most 12 bits wide.) Some other modifications of the basic scheme are discussed in [9]. If the d.c. component does not have to be zero, we could use some of the other schemes derived in Theorem 1 to obtain bounds for $K(m, d)$. The balancing vector scheme has the advantage that it can be implemented with very simple electronics even for large block lengths, which leads to a high information rate. Furthermore, by making small adjustments in the basic scheme, it is possible to achieve additional desirable properties, such as a guaranteed number of internal transitions (i.e.,

differences between neighboring bits) inside each transmitted block to aid in clock synchronization.

REFERENCES

[1] N. Alon, Sh. Friedland, and G. Kalai, Regular subgraphs of almost regular graphs, *J. Combinatorial Theory Ser. B 37* (1984), 79-91.

[2] E. E. Bergmann, A. M. Odlyzko, and S. Sangani, Weight 1/2 block codes for optical communications, to be published.

[3] B. S. Basik, The spectral density of a coded binary signal, *Bell System Tech. J. 51* (1972), 921-932.

[4] H. Enomoto, P. Frankl, N. Ito, and K. Nomura, manuscript in preparation.

[5] P. Frankl and V. Rödl, Forbidden intersections, to be published.

[6] J. N. Franklin and J. R. Pierce, Spectra and efficiency of binary codes without dc, *IEEE Trans. Comm., COM-21* (1972), 1438-1440.

[7] W. Fulton, *Algebraic Curves*, Benjamin/Cummings, 1969.

[8] A. V. Geramita and J. Seberry, *Orthogonal Designs*, Lecture Notes in Pure and Applied Mathematics, Vol. 45, Marcel Dekker, 1979.

[9] D. E. Knuth, Efficient balanced codes, *IEEE Trans. Information Theory IT-32* (1986), 51-53.

[10] G. L. Pierobon, Codes for zero spectral density at zero frequency, *IEEE Trans. Information Theory IT-30* (1984), 435-439.

[11] K. A. Schouhamen Immink and G. F. M. Beenker, Binary transmission codes

with higher order spectral zeros at zero frequency, to be published.

[12] H. S. Wilf, A unified setting for sequencing, ranking, and selection algorithms for combinatorial objects, *Advances in Math. 24* (1977), 281-291.

Balancing Sets of Vectors

*Noga Alon**

Department of Mathematics
Tel Aviv University
Ramat Aviv, Tel Aviv
and
Bell Communications Research
Morristown, NJ 07960

E. E. Bergmann

AT&T Bell Laboratories
Allentown, PA 18103

D. Coppersmith

IBM Research
Yorktown Heights, NY 10598

A. M. Odlyzko

AT&T Bell Laboratories
Murray Hill, NJ 07974

ABSTRACT

For $n > 0$, $d \geq 0$, $n \equiv d (\mathrm{mod} 2)$, let $K(n, d)$ denote the minimal cardinality of a family V of ± 1 vectors of length n, such that for any ± 1 vector $\mathbf{u}$ of length n there is a $\mathbf{v} \in V$ such that $|\mathbf{v} \cdot \mathbf{u}| \leq d$, where $\mathbf{v} \cdot \mathbf{u}$ is the usual scalar product of $\mathbf{v}$ and $\mathbf{u}$. A generalization of a very simple construction due to Knuth shows that $K(n, d) \leq \lfloor cn/(d+1) \rfloor c$. A proof is given here that this construction is optimal, so that $K(n, d) = \lfloor cn/(d+1) \rfloor c$ for all $n \equiv d \pmod 2$. This construction and its extensions have applications to communication theory, especially to construction of signal sets for optical data links.

* Research supported in part by Alon Fellowship and by The Fund for Basic Research Administered by the Israel Academy of Sciences.