TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 289, Number 2, June 1985

SIX STANDARD DEVIATIONS SUFFICE

BY

JOEL SPENCER[1]

ABSTRACT. Given n sets on n elements it is shown that there exists a two-coloring such that all sets have discrepancy at most $Kn^{1/2}$, K an absolute constant. This improves the basic probabilistic method with which $K = c(\ln n)^{1/2}$. The result is extended to n finite sets of arbitrary size. Probabilistic techniques are melded with the pigeonhole principle. An alternate proof of the existence of Rudin-Shapiro functions is given, showing that they are exponential in number. Given n linear forms in n variables with all coefficients in $[-1, +1]$ it is shown that initial values $p_1,\ldots,p_n \in \{0,1\}$ may be approximated by $\varepsilon_1,\ldots,\varepsilon_n \in \{0,1\}$ so that the forms have small error.

1. We state our main result first in the language of linear forms.

THEOREM 1. *Let*

$$L_i(x_1,\ldots,x_n) = a_{i1}x_1 + \cdots + a_{in}x_n, \qquad 1 \le i \le n, \tag{1.1}$$

be n linear forms in n variables with all $|a_{ij}| \le 1$. *Then there exist* $\varepsilon_1,\ldots,\varepsilon_n \in \{-1, +1\}$ *such that*

$$|L_i(\varepsilon_1,\ldots,\varepsilon_n)| \le K\sqrt{n} \tag{1.2}$$

for all i, $1 \le i \le n$. *Here K is an absolute constant.*

When all $a_{ij} \in \{0,1\}$ we may consider $A = (a_{ij})$ as the incidence matrix for a family of n sets on n elements. That is, we may set $A_i = \{j: a_{ij} = 1\}$. Given a two-coloring, say Red and Blue, of $\{1,\ldots,n\}$, the discrepancy $\operatorname{disc}(X)$ of a set $X \subset \{1,\ldots,n\}$ is defined as the number of Red points in X minus the number of Blue points in X. If we interpret $\varepsilon_i = +1$ as meaning i is to be colored Red and $\varepsilon_i = -1$ as Blue then we obtain the following result.

COROLLARY 2. *Let* $A_1,\ldots,A_n \subset \{1,\ldots,n\}$. *Then there exists a two-coloring of* $\{1,\ldots,n\}$ *so that*

$$|\operatorname{disc}(A_i)| \le K\sqrt{n} \tag{1.3}$$

for all i, $1 \le i \le n$.

Received by the editors July 1, 1984.

1980 *Mathematics Subject Classification*. Primary 05B20; Secondary 41A28, 42A61.

[1] This work was initiated while the author was an IREX exchange fellow at the Mathematical Institute, Budapest and completed with support of the National Science Foundation. For the many kindnesses and for the creative research environment provided by Institute staff and colleagues—*köszönöm szépen*!

©1985 American Mathematical Society
0002-9947/85 $1.00 + $.25 per page

In §4, using known techniques, we extend Corollary 2 as follows.

THEOREM 3. *Let $A_1,\dots,A_n \subset \Omega$ be arbitrary finite sets, Ω finite but of arbitrary size. Then there exists a two-coloring of Ω such that* (1.3) *holds for all i, $1 \le i \le n$.*

This resolves a question of Paul Erdös. References [**1**, **4**] contain earlier results.

A second application is to classical Fourier analysis. Let

$$f(z) = \varepsilon_1 z + \cdots + \varepsilon_n z^n, \tag{1.4}$$

where all $\varepsilon_i \in \{-1, +1\}$. Define a norm

$$\|f\| = \max_{|z|=1} |f(z)|, \tag{1.5}$$

where z is a complex variable. A Rudin-Shapiro function (with respect to a given K) is an f of the form (1.4) with $\|f\| \le K\sqrt{n}$. In §5 we give an alternate proof of the existence of Rudin-Shapiro functions with respect to a fixed sufficiently large constant K. Furthermore we show that the number of such functions is at least $(2 - \delta_K)^n$ with δ_K approaching zero as K approaches infinity.

Let $\|\cdot\|$ denote the L^∞ norm in R^n, $\|(x_1,\dots,x_n)\| = \max|x_i|$. We may reformulate Theorem 1 as follows. Let $v_1,\dots,v_n \in R^n$ with all $\|v_i\| \le 1$. Then there exist $\varepsilon_1,\dots,\varepsilon_n \in \{-1, +1\}$ such that

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n\| \le K\sqrt{n}\,. \tag{1.6}$$

Here, in the notation of Theorem 1, $v_j = (a_{1j},\dots,a_{nj})$ is the jth column vector of the matrix $A = (a_{ij})$ of coefficients. (This was the original formulation of Theorem 1.) Let $|\cdot|$ denote the L^2 (usual Euclidean) norm in R^n. Let $v_1,\dots,v_s \in R^n$ with $|v_i| \le 1$. János Komlós conjectures that there exist $\varepsilon_1,\dots,\varepsilon_s \in \{-1, +1\}$ such that

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_s v_s\| \le K. \tag{1.7}$$

Here K is an absolute constant with s, n arbitrary. While we have not succeeded in proving this tantalizing conjecture we show in §7 the existence of $\varepsilon_1,\dots,\varepsilon_s \in \{-1, 0, +1\}$ satisfying (1.7) with only a bounded proportion of the ε_i equal to zero. We give strong, albeit inconclusive, evidence for the full conjecture.

At its heart, our result is an extension of the probabilistic method. Let $a_1,\dots,a_n$ be arbitrary real numbers and let σ be such that

$$\sigma^2 = a_1^2 + \cdots + a_n^2. \tag{1.8}$$

Let $\boldsymbol{\varepsilon}_1,\dots,\boldsymbol{\varepsilon}_n$ be independent random variables with

$$\Pr[\boldsymbol{\varepsilon}_i = +1] = \Pr[\boldsymbol{\varepsilon}_i = -1] = 1/2 \tag{1.9}$$

and set

$$\mathbf{X} = \boldsymbol{\varepsilon}_1 a_1 + \cdots + \boldsymbol{\varepsilon}_n a_n. \tag{1.10}$$

Observe that $\mathbf{X}$ has mean zero and standard deviation σ. We shall use repeatedly the following result (see, e.g. [**6**]), valid for all $\lambda \ge 0$:

$$\Pr[|\mathbf{X}| > \lambda\sigma] < 2e^{-\lambda^2/2}. \tag{1.11}$$

We apply the basic probabilistic method to give a weak form of Theorem 1. Under the assumptions of Theorem 1

$$a_{i1}^2 + \cdots + a_{in}^2 \le n \tag{1.12}$$

for each i and

$$\Pr\left[|\mathbf{L}_i| > \lambda\sqrt{n}\right] < 2e^{-\lambda^2/2}, \tag{1.13}$$

where $\mathbf{L}_i = L_i(\boldsymbol{\varepsilon}_1,\ldots,\boldsymbol{\varepsilon}_n)$. When $\lambda = \sqrt{2}\sqrt{\ln(2n)}$

$$\Pr\left[|\mathbf{L}_i| > \lambda\sqrt{n}\right] < 1/n \tag{1.14}$$

for $1 \le i \le n$ and

$$\Pr\left[|\mathbf{L}_i| \le \lambda\sqrt{n} \text{ for all } i, 1 \le i \le n\right] > 1 - n(1/n) = 0. \tag{1.15}$$

Hence there exist $\varepsilon_1,\ldots,\varepsilon_n \in \{-1, +1\}$ such that

$$|L_i| \le \sqrt{2}\sqrt{n}\sqrt{\ln(2n)} \tag{1.16}$$

for $1 \le i \le n$.

Our improvement will thus consist of removing the $\ln(2n)^{1/2}$ factor.

A word on constants. In §§2 and 3 we prove our basic results for a specific value of K, one that is sufficiently large to give us plenty of room. We feel this makes the argument most clear for the reader—it certainly does for the author. In the final section we make some attempts at finding the best constant K. We do show $K \le 6$, giving our work its title. In the remainder of the paper we concern ourselves only with the existence of constants K having the desired properties.

2. The basic idea. In this section we prove a basic result that gives the key ideas of this work.

LEMMA 4. *Let*

$$L_i(x_1,\ldots,x_n) = a_{i1}x_1 + \cdots + a_{in}x_n, \qquad 1 \le i \le n, \tag{2.1}$$

be n linear forms in n variables with all $|a_{ij}| \le 1$. Then if n is sufficiently large, there exist $\varepsilon_1,\ldots,\varepsilon_n \in \{-1, 0, +1\}$ such that

$$|\{i: \varepsilon_i = 0\}| < 4 \times 10^{-10}n, \tag{2.2}$$

$$|L_i(\varepsilon_1,\ldots,\varepsilon_n)| < 10\sqrt{n}, \qquad 1 \le i \le n. \tag{2.3}$$

PROOF. Define a map

$$T: \{-1, +1\}^n \to Z^n \tag{2.4}$$

by

$$T(\varepsilon_1,\ldots,\varepsilon_n) = (b_1,\ldots,b_n), \tag{2.5}$$

where b_i is the *nearest integer* to $L_i(\varepsilon_1,\ldots,\varepsilon_n)/20\sqrt{n}$. That is,

$b_i = 0$ if and only if $|L_i| \le 10\sqrt{n}$,

$b_i = +1$ if and only if $10\sqrt{n} < |L_i| \le 30\sqrt{n}$,

$b_i = -1$ if and only if $-30\sqrt{n} \le |L_i| < -10\sqrt{n}$, etc.

The "new idea" in the proof is the following definition of a subset $B \subset Z^n$ of the range. Set

$$B = \{(b_1,\dots,b_n) \in Z^n: \quad |\{i: |b_i| \geq 1\}| < n(2e^{-50})4, \quad |\{i: |b_i| \geq 2\}| < n(2e^{-450})8 \tag{2.6}$$

and, in general,

$$|\{i: |b_i| \geq s\}| < n(2e^{-(2s-1)^2 50})2^{s+1}$$

for all positive integers $s\}$.

We shall show

$$|T^{-1}(B)| \geq \tfrac{1}{2}2^n, \tag{2.7}$$

$$|B| \leq 2^{cn}, \qquad c = 1.1 \times 10^{-19}. \tag{2.8}$$

Let $\boldsymbol{\varepsilon}_1,\dots,\boldsymbol{\varepsilon}_n \in \{-1, +1\}$ be independent and uniform and let $\mathbf{L}_1,\dots,\mathbf{L}_n$, $\mathbf{b}_1,\dots,\mathbf{b}_n$ be the values they generate.

$$\Pr[|\mathbf{b}_i| \geq 1] = \Pr[|\mathbf{L}_i| \geq 10\sqrt{n}] < 2e^{-50} \tag{2.9}$$

(by (1.13)) for each i. As expectation is linear

$$E[|\{i: |\mathbf{b}_i| \geq 1\}|] = n(2e^{-50}). \tag{2.10}$$

(Note that the $\mathbf{b}_i$ are not necessarily independent.) Hence

$$\Pr[|\{i: |\mathbf{b}_i| \geq 1\}| \geq n(2e^{-50})4] \leq 1/4. \tag{2.11}$$

Similarly

$$\Pr\left[\left|\{i: |\mathbf{b}_i| \geq s\}\right| \geq n\left(2e^{-(2s-1)^2 50}\right)2^{s+1}\right] \leq 1/2^{s+1}.$$

Thus

$$\Pr[(\mathbf{b}_1,\dots,\mathbf{b}_n) \notin B] \leq \sum_{s=1}^{\infty} 1/2^{s+1} = 1/2. \tag{2.12}$$

That is, at least half of all $(\varepsilon_1,\dots,\varepsilon_n) \in \{-1, +1\}^n$ are in $T^{-1}(B)$, yielding (2.7).

Claim (2.8) will follow from crude counting arguments. In general, suppose

$$\tfrac{1}{2} > \alpha_1 > \alpha_2 > \cdots \tag{2.13}$$

and

$$B = \left\{(b_1,\dots,b_n) \in Z^n: \left|\{i: |b_i| \geq s\}\right| \leq \alpha_s n, s = 1,2,\dots\right\}, \tag{2.14}$$

where $1/2 > \alpha_1 > \alpha_2 > \cdots$. Then

$$|B| \leq \prod_{s=1}^{\infty} \left[\left[\sum_{i=0}^{\alpha_s n} \binom{n}{i}\right] 2^{\alpha_s n}\right]. \tag{2.15}$$

Indeed $\{i: |b_i| = s\}$ can be chosen in at most $\sum_{i=0}^{\alpha_s n}\binom{n}{i}$ ways and, having been selected, can be split into $\{i: b_i = s\}$ and $\{i: b_i = -s\}$ in at most $2^{\alpha_s n}$ ways. We bound

$$\sum_{i=0}^{\alpha n} \binom{n}{i} \leq 2^{nH(\alpha)}, \tag{2.16}$$

where $H(\alpha)$ is the entropy function

$$H(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha). \tag{2.17}$$

Therefore

$$|B| \le 2^{cn}, \quad \text{where } c = \sum_{s=0}^{\infty} \left[H(\alpha_s) + \alpha_s\right]. \tag{2.18}$$

In our case $\alpha_1 = 8e^{-50}$, $\alpha_2 = 16e^{-450}$ and, in general, $\alpha_s = 2^{s+2}e^{-50(2s-1)^2}$. The sequence (2.18) for c clearly converges and is dominated by the first term.

$$c \sim H[8e^{-50}] < 1.1 \times 10^{-19} \tag{2.19}$$

as claimed. (Note: If the value of the constant in (2.2) is not pertinent, then we need only $c < 1$.)

Applying the "pigeonhole principle" to (2.7), (2.8) there exists $(b_1,\ldots,b_n) \in B$ such that, setting

$$\mathscr{A} = \left\{(\varepsilon_1,\ldots,\varepsilon_n) \in \{-1,+1\}^n : T(\varepsilon_1,\ldots,\varepsilon_n) = (b_1,\ldots,b_n)\right\} \tag{2.20}$$

we may bound

$$|\mathscr{A}| \ge |T^{-1}(B)|/|B| \ge 2^{n(1-c)-1}. \tag{2.21}$$

We use the following result, due to D. Kleitman [3].

THEOREM 5. *Let* $\mathscr{A} \subset \{-1,+1\}^r$, $s < r/2$, $|\mathscr{A}| \ge \sum_{i=0}^{s}\binom{r}{i}$. *Then* $\operatorname{diam}(\mathscr{A}) \ge 2s$.

That is, there exist two vectors in $\mathscr{A}$ which differ in at least $2s$ coordinates. (This result is "best possible" since $\mathscr{A}$ may be a ball of radius s around an arbitrary point using the Hamming metric.) We rewrite Theorem 5 in a form appropriate for our need.

If $|\mathscr{A}| \ge 2^{rH(1/2-p)}$ with $p > 0$, then $\operatorname{diam}(\mathscr{A}) \ge (1-2p)r$. (2.22)

Let p_0 be that positive real such that

$$H\left(\frac{1}{2} - p_0\right) = 1 - c \tag{2.23}$$

and let p be such that $p > p_0$. For x small

$$H\left(\frac{1}{2} - x\right) \sim 1 - \left(\frac{2}{\ln 2}\right)x^2 \tag{2.24}$$

so that

$$p_0 \sim \left[\frac{\ln 2}{2}c\right]^{1/2} < 2 \cdot 10^{-10} \tag{2.25}$$

and we may take $p = 2 \cdot 10^{-10}$. We require n to be sufficiently large so that

$$n(1-c) - 1 \ge nH\left(\frac{1}{2} - p\right). \tag{2.26}$$

(This is our only condition on n.) Since $|\mathscr{A}| \ge 2^{n(1-c)-1} \ge 2^{nH(1/2-p)}$, $\operatorname{diam}(\mathscr{A}) \ge (1-2p)n$. Let $\vec{\varepsilon}' = (\varepsilon_1',\ldots,\varepsilon_n')$, $\vec{\varepsilon}'' = (\varepsilon_1'',\ldots,\varepsilon_n'') \in \mathscr{A}$ with $\rho(\vec{\varepsilon}', \vec{\varepsilon}'') = \operatorname{diam}(\mathscr{A})$, where we let ρ denote the Hamming metric. Set

$$\vec{\varepsilon} = (\varepsilon_1,\ldots,\varepsilon_n) = (\vec{\varepsilon}' - \vec{\varepsilon}'')/2. \tag{2.27}$$

That is,

$$\varepsilon_i = \left(\varepsilon_i' - \varepsilon_i''\right)/2, \qquad 1 \le i \le n. \tag{2.28}$$

All $\varepsilon_i \in \{-1, 0, +1\}$ and $\varepsilon_i = 0$ if and only if $\varepsilon_i' = \varepsilon_i''$ so

$$|\{i: \varepsilon_i = 0\}| = n - \rho(\varepsilon', \varepsilon'') = n - \operatorname{diam}(\mathscr{A}) \le 2pn \tag{2.29}$$

so that (2.2) is satisfied. For all i

$$L_i(\vec{\varepsilon}) = \left[L_i(\vec{\varepsilon}') - L_i(\vec{\varepsilon}'')\right]/2. \tag{2.30}$$

Since T is identical on $\vec{\varepsilon}'$ and $\vec{\varepsilon}''$, $L_i(\vec{\varepsilon}')$ and $L_i(\vec{\varepsilon}'')$ lie on a common interval of length $20\sqrt{n}$. Thus $|L_i(\vec{\varepsilon})| \le 10\sqrt{n}$, i.e., (2.3) is satisfied, completing the proof.

REMARK. We have not been able to find an algorithm that will yield $\varepsilon_1, \ldots, \varepsilon_n$ satisfying Lemma 4 (or the later results of this work employing the same basic methodology) in polynomial time. The stumbling block appears to be in the use of the pigeonhole principle. The paragraphs at the conclusion of §3 suggest, but surely do not prove, that such an algorithm may not exist.

3. Proof of Theorem 1. The following lemma generalizes Lemma 4 and differs mainly in technical details.

LEMMA 6. *Let $r \le n$ and let*

$$L_i(x_1, \ldots, x_r) = a_{i1}x_1 + \cdots + a_{ir}x_r, \qquad 1 \le i \le n, \tag{3.1}$$

be n linear forms in r variables with all $|a_{ij}| \le 1$. Then, if r is sufficiently large, there exist $\varepsilon_1, \ldots, \varepsilon_r \in \{-1, 0, +1\}$ with

$$|\{i: \varepsilon_i = 0\}| < cr, \qquad c = 6 \times 10^{-7}, \tag{3.2}$$

$$|L_i(\varepsilon_1, \ldots, \varepsilon_r)| < 10\sqrt{r}\sqrt{\ln(2n/r)}, \qquad 1 \le i \le n. \tag{3.3}$$

PROOF. Define a map T: $\{-1, +1\}^r \to Z^n$ by

$$T(\varepsilon_1, \ldots, \varepsilon_r) = (b_1, \ldots, b_n), \tag{3.4}$$

where b_i is the nearest integer to $L_i(\varepsilon_1, \ldots, \varepsilon_r)/20\sqrt{r}\sqrt{\ln(2n/r)}$. Define $B \subset Z^n$ by

$$\begin{aligned} B = \{(b_1, \ldots, b_n) \in Z^n:& \\ &|\{i: |b_i| \ge 1\}| \le n(2n/r)^{-50}4, \\ &|\{i: |b_i| \ge 2\}| \le n(2n/r)^{-450}8 \\ &\text{and, in general,} \\ &|\{i: |b_i| \ge s\}| \le n(2n/r)^{-50(2s-1)^2}2^{s+1} \\ &\text{for all positive integers } s\}. \end{aligned} \tag{3.5}$$

We shall show

$$|T^{-1}(B)| \ge \tfrac{1}{2}2^r, \tag{3.6}$$

$$|B| < 2^{\gamma r}, \qquad \gamma = 250 \cdot 2^{-50}. \tag{3.7}$$

Let $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_r \in \{-1, +1\}$ be independent and uniform and let $\mathbf{L}_1, \ldots, \mathbf{L}_n$, $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be the values they generate.

$$\Pr\left[|\mathbf{b}_i| \ge 1\right] = \Pr\left[|\mathbf{L}_i| \ge 10\sqrt{r}\sqrt{\ln(2n/r)}\right] < (2n/r)^{-50} \tag{3.8}$$

(by (1.13)) for each i. As in the proof of Lemma 4

$$E\left[\left|\{i: |\mathbf{b}_i| \geq 1\}\right|\right] < n(2n/r)^{-50}, \tag{3.9}$$

$$\Pr\left[\left|\{i: |\mathbf{b}_i| \geq 1\}\right| > n(2n/r)^{-50}4\right] < 1/4. \tag{3.10}$$

Similarly

$$\Pr\left[\left|\{i: |\mathbf{b}_i| \geq s\}\right| > n(2n/r)^{-50(2s-1)^2}2^{s+1}\right] < 1/2^{s+1} \tag{3.11}$$

so

$$\Pr[(\mathbf{b}_1, \ldots, \mathbf{b}_n) \notin B] < \sum_{s=1}^{\infty} 1/2^{s+1} = 1/2 \tag{3.12}$$

yielding (3.6).

As with (2.18) we bound $|B| \leq 2^{\beta n}$, where

$$\beta = \sum_{s=1}^{\infty} (H(\alpha_s) + \alpha_s), \qquad \alpha_s = (2n/r)^{-50(2s-1)^2}2^{s+1}. \tag{3.13}$$

As $(2n/r) \geq 2$, $\alpha_{s+1} \leq 2^{-49}\alpha_s$ for all s and all $\alpha_s \leq \alpha_1 \leq 2^{-48}$ so

$$H(\alpha_{s+1}) + \alpha_{s+1} \leq 2^{-47}[H(\alpha_s) + \alpha_s]. \tag{3.14}$$

Thus β is dominated by the first term

$$\begin{aligned} |B| &< (1 + 2^{-46})(H(\alpha_1) + \alpha_1) \\ &< 1.1\alpha_1(-\log_2 \alpha_1) \quad (\text{since } \alpha_1 < 2^{-48}) \\ &< 5(2n/r)^{-50}[48 + 50\log_2(n/r)]. \end{aligned} \tag{3.15}$$

Then $|B| \leq 2^{\beta n} \leq 2^{\gamma r}$, where

$$\begin{aligned} \gamma &= (n/r)5(2n/r)^{-50}[48 + 50\log_2(n/r)] \\ &= 5 \cdot 2^{-50}(n/r)^{-49}[48 + 50\log_2(n/r)] \\ &< 250 \cdot 2^{-50} \end{aligned} \tag{3.16}$$

since, letting $n/r = y$, the inequality $y^{-49}(48 + 50\log_2 y) \leq 50$ is valid for all $y \geq 1$.

Applying the "pigeonhole principle" to (3.6), (3.7), precisely as in Lemma 4, we find $\mathscr{A}$ on which T is constant with

$$|\mathscr{A}| \geq |T^{-1}(B)|/|B| \geq 2^{r(1-\gamma)-1}. \tag{3.17}$$

Let p_0 be that positive real such that $H(\frac{1}{2} - p_0) = 1 - \gamma$ and let p be such that $p > p_0$. In our case

$$p_0 \sim \left[\frac{\ln 2}{2}\gamma\right]^{1/2} < 3 \times 10^{-7} \tag{3.18}$$

so we take $p = 3 \times 10^{-7}$. Let r be sufficiently large so that

$$r(1 - \gamma) - 1 \geq rH(\tfrac{1}{2} - p). \tag{3.19}$$

Then $|\mathscr{A}| \geq 2^{r(1-\gamma)-1} \geq 2^{rH(1/2-p)}$ so, by (2.22), $\operatorname{diam}(\mathscr{A}) \geq (1 - 2p)r$.

Let $\vec{\varepsilon}', \vec{\varepsilon}'' \in \mathscr{A}$ with $\rho(\vec{\varepsilon}', \vec{\varepsilon}'') = \operatorname{diam}(\mathscr{A})$ and set

$$\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_r) = (\vec{\varepsilon}' - \vec{\varepsilon}'')/2. \tag{3.20}$$

Then, precisely as with Lemma 4, $\vec{\varepsilon}$ satisfies (3.2), (3.3) with $c = 2p$.

THEOREM 7. *Let $r \leq n$ and let*

$$L_i(x_1, \ldots, x_r) = a_{i1}x_1 + \cdots + a_{ir}x_r, \qquad 1 \leq i \leq n, \tag{3.21}$$

be n linear forms in r variables with all $|a_{ij}| \leq 1$. Then for r sufficiently large there exist $\varepsilon_1, \ldots, \varepsilon_r \in \{-1, +1\}$ with

$$|L_i(\varepsilon_1, \ldots, \varepsilon_r)| \leq 11\sqrt{r}\sqrt{\ln(2n/r)} \tag{3.22}$$

for $1 \leq i \leq n$.

Setting $n = r$, Theorem 1 is derived as a special case with $K = 11\sqrt{\ln 2} < 9.2$. Improvement of K to less than six (yielding our title) is given in the final section. These K apply only if $n \geq n_0$, where n_0 is some absolute constant. Since $|L_i| \leq n$ always Theorem 1 holds for *all* n by redefining $K = \max(9.2, \sqrt{n_0})$. Then if $n < n_0$, $|L_i| \leq n \leq K\sqrt{n}$.

REMARK. The elementary use of the probabilistic method, as given in §1, can be used to show Theorem 7 with $11\sqrt{r}\sqrt{\ln(2n/r)}$ replaced by $\sqrt{2}\sqrt{r}\sqrt{\ln(2n)}$. When $n > r^{1+c}$, c positive fixed, these results lie within a constant factor of each other. Thus Theorem 7, while valid for all $r \leq n$, improves previous results only when $n = r^{1+o(1)}$.

PROOF. Let k be that absolute constant so that Lemma 6 applies for all $r \geq k$. Set $r = r_0$ and apply Lemma 6 to find values $\varepsilon_i \in \{-1, +1\}$ to all but r_1 variables with $r_1 \leq cr_0$. Let $L_i^{(1)}$ be the ith linear form restricted to the r_1 still undetermined (i.e., still equal zero) variables and apply Lemma 6 again, leaving r_2 variables undetermined with $r_2 \leq c^2 r_0$. Iterate this process, giving a sequence $r = r_0 > r_1 > \cdots > r_{u+1}$ terminating when $r_{u+1} < k$. At this stage the undetermined variables are set equal to $+1$ or -1 arbitrarily, effecting $|L_i|$ by at most r_{u+1}. With all $\varepsilon_i \in \{-1, +1\}$

$$\begin{aligned} |L_i| &\leq \sum_{t=0}^{u} 10\sqrt{r_t}\sqrt{\ln(2n/r_t)} + r_{u+1} \\ &\leq k + \sum_{t=0}^{\infty} 10\sqrt{rc^t}\sqrt{\ln(2n/rc^t)}. \end{aligned} \tag{3.23}$$

While the final steps are simple calculations we do them in detail. They show that when Lemma 6 is iterated the later terms are secondary. Set $A = \ln(2n/r)$ so that $A \geq \ln(2)$. Recall $c = 6 \times 10^{-7}$. We use the inequality $(x + y)^{1/2} \leq x^{1/2} + y^{1/2}$, valid for all $x, y \geq 0$. Then

$$\sum_{t=0}^{\infty} \sqrt{c^t}\sqrt{\ln(Ac^{-t})} \leq \sum_{t=0}^{\infty} \sqrt{c}^t \left[\sqrt{\ln A} + \sqrt{t \ln c^{-1}}\right], \tag{3.24}$$

$$\sum_{t=0}^{\infty} \sqrt{c^t} = 1/(1 - \sqrt{c}) < 1 + 10^{-3}, \tag{3.25}$$

$$\sum_{t=0}^{\infty} \sqrt{c^t}\sqrt{t \ln c^{-1}} < 3 \times 10^{-3}. \tag{3.26}$$

So

$$(3.27)\qquad \sum_{t=0}^{\infty} \sqrt{c^t}\sqrt{\ln(Ac^{-t})} < \sqrt{\ln A}\,(1 + 10^{-3}) + (3 \times 10^{-3}) < \sqrt{\ln A}\,(1.005)$$

and

$$(3.28)\qquad |L_i| \le k + 10\sqrt{r} \sum_{t=0}^{\infty} \sqrt{c^t}\sqrt{\ln(Ac^{-t})} \le k + 10\sqrt{r}\sqrt{\ln A}\,(1.005)$$

$$\le k + 10.05\sqrt{r}\sqrt{\ln(2n/r)}\,.$$

This result holds for *all* r. When r is sufficiently large the constant k may be absorbed into the main term, giving Theorem 7.

The juxtaposition of the probabilistic method and the pigeonhole principle allows us to prove the existence of an appropriate $\vec{\varepsilon} \in \{-1, +1\}^n$ when, as we shall show, only an exponentially small proportion of the $\vec{\varepsilon}$ have the desired property. Let $\mathscr{A} = \{-1, +1\}^n$, the set of possible $\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$ and let $\mathscr{M}$ be the set of $n \times n$ matrices $A = (a_{ij})$ with all $a_{ij} \in \{-1, +1\}$. Let K be a fixed but arbitrary positive constant. Let

$$(3.29)\qquad \mathscr{U} = \mathscr{U}_K = \left\{(A, \vec{\varepsilon}) : \vec{\varepsilon} \in \mathscr{A}, A \in \mathscr{M}, \|A\vec{\varepsilon}\| \le K\sqrt{n}\right\},$$

where $\|\cdot\|$ is the L^∞ norm. That is, with $L_1, \ldots, L_n$ given by (1.1), $(A, \vec{\varepsilon}) \in \mathscr{U}_K$ if and only if (1.2) holds for all i, $1 \le i \le n$. Set

$$(3.30)\qquad \mathscr{M}_{\vec{\varepsilon}} = \{A \in \mathscr{M} : (A, \vec{\varepsilon}) \in \mathscr{U}\}, \qquad \vec{\varepsilon} \in \mathscr{A},$$

$$\mathscr{A}_A = \{\vec{\varepsilon} \in \mathscr{A} : (A, \vec{\varepsilon}) \in \mathscr{U}\}, \qquad A \in \mathscr{M},$$

so that

$$(3.31)\qquad |\mathscr{U}| = \sum_{\vec{\varepsilon} \in \mathscr{A}} |\mathscr{M}_{\vec{\varepsilon}}| = \sum_{A \in \mathscr{M}} |\mathscr{A}_A|.$$

Fix $\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n) \in \mathscr{A}$. Let $\mathbf{a}_{ij} \in \{-1, +1\}$, $1 \le i, j \le n$, be uniform and independent random variables. Then for $1 \le i \le n$

$$(3.32)\qquad \mathbf{L}_i = \mathbf{a}_{i1}\varepsilon_1 + \cdots + \mathbf{a}_{in}\varepsilon_n$$

has the distribution of $\mathbf{U}_n$ of the sum of n independent variables, each chosen uniformly from $\{-1, +1\}$. (The particular values of $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}^n$ do not effect the distribution!) By the Central Limit Theorem

$$(3.33)\qquad \lim_{K \to \infty} \Pr\left[|\mathbf{L}_i| \ge K\sqrt{n}\right] = 2\Phi(-K),$$

where

$$(3.34)\qquad \Phi(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,dt$$

is the cdf of the standard normal distribution. The mutual independence of the variables $\mathbf{a}_{ij}$ guarantees that the $\mathbf{L}_i$ are mutually independent. Thus

$$(3.35)\qquad \Pr\left[|\mathbf{L}_i| \le K\sqrt{n}, 1 \le i \le n\right] = \prod_{i=1}^{n} \Pr\left[|\mathbf{L}_i| \le K\sqrt{n}\right]$$

$$= (1 - 2\Phi(-K) - o(1))^n,$$

where $o(1)$ is, for arbitrary fixed K, an "error term" approaching zero as n approaches infinity. Then

$$|\mathscr{M}_{\vec{\varepsilon}}| = 2^{n^2}\Pr[(\mathbf{A}, \vec{\varepsilon}) \in \mathscr{U}] = 2^{n^2}[1 - 2\Phi(-K) - o(1)]^n. \tag{3.36}$$

Double counting

$$\sum_{A \in \mathscr{M}} |\mathscr{A}_A| = \sum_{\vec{\varepsilon} \in \mathscr{A}} |\mathscr{M}_{\vec{\varepsilon}}| = 2^n 2^{n^2}(1 - 2\Phi(-K) - o(1))^n \tag{3.37}$$

$$= |\mathscr{M}|2^n(1 - 2\Phi(-K) - o(1))^n.$$

Therefore there exists $A \in \mathscr{M}$, i.e. a particular choice of a_{ij}, so that

$$|\mathscr{A}_A| \leq 2^n(1 - 2\Phi(-K) - o(1))^n. \tag{3.38}$$

For this particular A if $\vec{\varepsilon} \in \mathscr{A}$ is chosen uniformly,

$$\Pr[\vec{\varepsilon} \in \mathscr{A}_A] \leq (1 - 2\Phi(-K) - o(1))^n \tag{3.39}$$

which is exponentially small.

We have actually shown

$$E\left[|\mathscr{A}_{\mathbf{A}}|\right] = 2^n(1 - 2\Phi(-K) - o(1))^n, \tag{3.40}$$

where $\mathbf{A}$ is the matrix with coefficients $\mathbf{a}_{ij}$.

$$E\left[\Pr[\vec{\varepsilon} \in \mathscr{A}_{\mathbf{A}}]\right] = (1 - 2\Phi(-K) - o(1))^n, \tag{3.41}$$

where the Expected Value is taken with respect to $\mathbf{A}$ and the inner probability with respect to $\vec{\varepsilon}$ for a fixed A. Thus, for example, for any $c \geq 1$

$$\Pr[\vec{\varepsilon} \in \mathscr{A}_A] \geq c(1 - 2\Phi(-K) - o(1))^n \tag{3.42}$$

for at most a proportion c^{-1} of the $A \in \mathscr{M}$.

We cannot say that $|\mathscr{A}_A|$ is exponentially small, or even small, for all A. Indeed, when all rows of A are identical the variables $L_i(\varepsilon_1, \ldots, \varepsilon_n)$ are identical so

$$\Pr[\vec{\varepsilon} \in \mathscr{A}_A] = 1 - 2\Phi(-K) - o(1). \tag{3.43}$$

We have not been able to give a specific A for which we can prove (3.39), or even that $\Pr[\vec{\varepsilon} \in \mathscr{A}_A]$ is exponentially small. We believe that (3.39) holds if A is a Hadamard matrix. While there is strong evidence in that direction we do not have a proof.

4. Reductions and discrepancies. The results of this section will allow us to make reductions when there are more variables than linear forms. The methods are "well known".

COROLLARY 8. *Let*

$$L_i(x_1, \ldots, x_n) = a_{i1}x_1 + \cdots + a_{in}x_n, \qquad 1 \leq i \leq n, \tag{4.1}$$

be n linear forms in n variables with all $|a_{ij}| \leq 1$. Let $p_1, \ldots, p_n \in [0, 1]$. Then there exist $\varepsilon_1, \ldots, \varepsilon_n \in \{0, 1\}$ such that

$$|L_i(\varepsilon_1, \ldots, \varepsilon_n) - L_i(p_1, \ldots, p_n)| \leq K\sqrt{n} \tag{4.2}$$

for all i, $1 \leq i \leq n$. Here K is the same absolute constant as in Theorem 1.

PROOF. Assume $p_1,\dots,p_n$ have finite binary expansions with maximal length T. Let J be the set of indices j for which p_j has a "1" as its Tth binary digit. Set

$$L_i^* = \sum_{j\in J} a_{ij}x_j. \tag{4.3}$$

By Theorem 1 there exist $\varepsilon_j \in \{-1, +1\}, j \in J$, so that $|L_i^*| \le K\sqrt{n}$. Set

$$p_j^* = \begin{cases} p_j + 2^{-T} & \text{if } \varepsilon_j = +1, \\ p_j - 2^{-T} & \text{if } \varepsilon_j = -1, \\ p_j & \text{if } j \notin J. \end{cases} \tag{4.4}$$

Then

$$\left|L_i(p_1^*,\dots,p_n^*) - L_i(p_1,\dots,p_n)\right| \le 2^{-T}|L_i^*| \le 2^{-T}K\sqrt{n} \tag{4.5}$$

and $p_1^*,\dots,p_n^*$ have binary expansions with maximal length $T-1$. Applying this procedure $(T-1)$ more times we replace $p_1,\dots,p_n$ with $\varepsilon_1,\dots,\varepsilon_n \in \{0,1\}$ such that

$$\left|L_i(\varepsilon_1,\dots,\varepsilon_n) - L_i(p_1,\dots,p_n)\right| \le \sum_{s=1}^{T} 2^{-s}K\sqrt{n} < K\sqrt{n}. \tag{4.6}$$

Finally, if $p_1,\dots,p_n \in [0,1]$ are arbitrary the existence of $\varepsilon_1,\dots,\varepsilon_n$ follows from a simple Compactness Argument.

Corollary 8 has a geometric reformulation. Let $v_1,\dots,v_n \in R^n$, $\|v_j\| \le 1$. Let $p_1,\dots,p_n \in [0,1]$ and set $w = p_1v_1 + \cdots + p_nv_n$. Then there exists a vertex $v = \varepsilon_1v_1 + \cdots + \varepsilon_nv_n$ of the parallelepiped generated by the v_i such that $\|v - w\| \le K\sqrt{n}$. Here, as in §1, $v_j = (a_{1j},\dots,a_{nj})$ is the jth column vector of the matrix $A = (a_{ij})$ of coefficients. This variant is reminiscent of, but apparently not derivable from, the Minkowski geometry of numbers.

THEOREM 9. *Let*

$$L_i(x_1,\dots,x_r) = a_{i1}x_1 + \cdots + a_{ir}x_r, \qquad 1 \le i \le n, \tag{4.7}$$

be n linear forms in r variables with $r \ge n$. Then there exist $x_1,\dots,x_r \in [-1, +1]$ such that

$$AL_i(x_1,\dots,x_r) = 0, \qquad 1 \le i \le n, \tag{4.8}$$

$$\left|\{j: x_j \notin \{-1, +1\}\}\right| \le n. \tag{4.9}$$

PROOF. Let $x_1,\dots,x_r \in [0,1]$ satisfy (4.7) such that $|\{j: x_j \notin \{-1, +1\}|$ is minimal. Suppose, reordering for convenience, that $x_1,\dots,x_s \in (-1, +1)$, $x_{s+1},\dots,x_r \in \{-1, +1\}$ with $s > n$ (else there is nothing to prove). Let $(\lambda_1,\dots,\lambda_s)$ be a nonzero solution to the underdetermined homogeneous system

$$a_{i1}y_1 + \cdots + a_{is}y_s = 0, \qquad 1 \le i \le n. \tag{4.10}$$

Let α be the smallest real value, in absolute value, so that some $x_j + \lambda_j\alpha \in \{-1, +1\}$. Then set $x_j' = x_j + \lambda_j\alpha$, $1 \le j \le s$; $x_j' = x_j$, $s < j$. The values $x_1',\dots,x_r'$ satisfy (4.7) but there are fewer than s indices j with $x_j' \notin \{-1, +1\}$, contradicting the minimality of s.

THEOREM 10. *Let $r \geq n$ and let*

$$L_i(x_1,\ldots,x_r) = a_{i1}x_1 + \cdots + a_{ir}x_r, \qquad 1 \leq i \leq n, \tag{4.11}$$

be n linear forms in r variables with all $|a_{ij}| \leq 1$. Then there exist $\varepsilon_1,\ldots,\varepsilon_r \in \{-1, +1\}$ such that

$$|L_i(\varepsilon_1,\ldots,\varepsilon_r)| \leq 2K\sqrt{n}, \qquad 1 \leq i \leq n, \tag{4.12}$$

where K is the absolute constant of Theorem 1.

PROOF. By Theorem 9, renumbering for convenience, there exist $p_1,\ldots,p_n \in [-1, +1]$, $\varepsilon_{n+1},\ldots,\varepsilon_r \in \{-1, +1\}$ so that

$$L_i(p_1,\ldots,p_n, \varepsilon_{n+1},\ldots,\varepsilon_r) = 0, \qquad 1 \leq i \leq n. \tag{4.13}$$

Let L_i^* denote L_i restricted to the first n variables. That is,

$$L_i^*(x_1,\ldots,x_n) = a_{i1}x_1 + \cdots + a_{in}x_n \tag{4.14}$$

for $1 \leq i \leq n$. We apply Corollary 8 to find $\varepsilon_1,\ldots,\varepsilon_n \in \{-1, +1\}$ with

$$|L_i^*(\varepsilon_1,\ldots,\varepsilon_n) - L_i^*(p_1,\ldots,p_n)| \leq 2K\sqrt{n} \tag{4.15}$$

for $1 \leq i \leq n$. (We require a linear transformation between $[0, 1]$ in Corollary 8 and $[-1, +1]$ here causing an additional factor of two.) These $\varepsilon_1,\ldots,\varepsilon_n, \varepsilon_{n+1},\ldots,\varepsilon_r$ have the desired property.

We apply Theorem 10 directly to discrepancies of sets. Let $A_1,\ldots,A_n \subset \Omega$, all finite sets. For convenience write $\Omega = \{1,\ldots,r\}$ and, for $1 \leq i \leq n$, $1 \leq j \leq r$, set $a_{ij} = 1$ if $j \in A_i$, zero otherwise. Theorem 10 gives $\varepsilon_1,\ldots,\varepsilon_r \in \{-1, +1\}$ such that all $|L_i(\varepsilon_1,\ldots,\varepsilon_r)| \leq 2K\sqrt{n}$. We associate a two-coloring $\chi\colon \Omega \to \{-1, +1\}$ by $\chi(j) = \varepsilon_j$. Then $L_i(\varepsilon_1,\ldots,\varepsilon_r)$ is the discrepancy of set A_i. Hence Corollary 3 is proven.

5. Rudin-Shapiro functions. We consider the norm

$$\|f\| = \max_{|z|=1} |f(z)|, \tag{5.1}$$

where f is a function of a complex variable z.

THEOREM 11. *There exist $\varepsilon_1,\ldots,\varepsilon_n \in \{-1, +1\}$ so that, setting*

$$f(z) = \varepsilon_1 z + \cdots + \varepsilon_n z^n \tag{5.2}$$

we have

$$\|f\| \leq K\sqrt{n}. \tag{5.3}$$

Here K is an absolute constant.

The first examples of such functions were given by H. Shapiro and later rediscovered independently by W. Rudin [5]. We shall call such an f a Rudin-Shapiro function for K. The classic inequality of Bernstein states $|f'(z)| \leq n\|f\|$ for all z on the unit circle. Let z be that point such that $|f(z)| = \|f\|$. Let ω be a primitive $(4n)$th root of unity. There exists j, $0 \leq j < 4n$, $|z - \omega^j| \leq 2\pi/8n$. Then

$$|f(\omega^j)| \geq |f(z)| - |z - \omega^j|(n\|f\|) \geq \|f\|(1 - (2\pi/8))$$

and hence

$$\|f\| \le 6 \max_{0 \le j < 4n} |f(\omega^j)|. \tag{5.4}$$

These constants could easily be improved. The essential point is that bounding $\|f\|$ now becomes a discrete problem. (A proof not using (5.4) is given at the end of §6.) For $0 \le j < 4n$ set

$$L_j(\varepsilon_1, \ldots, \varepsilon_n) = \mathrm{Re}\big[f(\omega^j)\big] = \sum_{i=1}^{n} \varepsilon_i \mathrm{Cos}(2\pi ij/2n), \tag{5.5}$$

$$L_{j+4n}(\varepsilon_1, \ldots, \varepsilon_n) = \mathrm{Im}\big[f(\omega^j)\big] = \sum_{i=1}^{n} \varepsilon_i \mathrm{Sin}(2\pi ij/2n). \tag{5.6}$$

By Theorem 1 (with $8n$ forms) there exist $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ such that all

$$|L_j| \le K\sqrt{8n}, \tag{5.7}$$

with K the constant of Theorem 1. This implies

$$|f(\omega^j)| \le \left[L_j^2 + L_{j+4n}^2\right]^{1/2} \le 4K\sqrt{n} \tag{5.8}$$

so that

$$\|f\| \le 24K\sqrt{n} \tag{5.9}$$

as desired.

Further, we prove the existence of "many" Rudin-Shapiro functions. Our result is best stated in the language of Theorem 1. (Throughout this section $o(1)$ denotes a function approaching zero in n for any fixed K.)

THEOREM 12. *Under the conditions of Theorem* 1 *let* $s_K(n)$ *denote the number of* $(\varepsilon_1, \ldots, \varepsilon_n) \in \{-1, +1\}$ *such that*

$$|L_i(\varepsilon_1, \ldots, \varepsilon_n)| < K\sqrt{n} \tag{5.10}$$

for $1 \le i \le n$. *For all* $K \ge K_0$, *where* K_0 *is an absolute constant,*

$$s_K(n) > (2 - \delta_K + o(1))^n, \tag{5.11}$$

where $0 < \delta_K < 1$. *Furthermore*

$$\lim_{K \to \infty} \delta_K = 0. \tag{5.12}$$

We outline the argument, which requires examination of the proof of Lemma 4. We found $\mathscr{A} \subset \{-1, +1\}^n$ on which T was constant. With $K = 10$ we had $|\mathscr{A}| \ge 2^{n(1-c)(1+o(1))}$ with $c \sim 1.1 \times 10^{-19}$. For $K \ge 10$ we get $|\mathscr{A}| \ge 2^{n(1-c(K))(1+o(1))}$, where $c(K) \to 0$ as $K \to \infty$. Let $p(K) > 0$ satisfy $H(\frac{1}{2} - p(K)) = 1 - c(K)$ so that $p(K) \to 0$ as $K \to \infty$. Set $\mathscr{A} = \mathscr{A}^{(0)}$ and for $1 \le t \le |\mathscr{A}|/4$, having defined $\mathscr{A}^{(t-1)}$, let $x^{(t)}, y^{(t)} \in \mathscr{A}^{(t-1)}$ be a pair of vectors at maximal distance and set $\mathscr{A}^{(t)} = \mathscr{A}^{(t-1)} - \{x^{(t)}, y^{(t)}\}$. As $|\mathscr{A}^{(t)}| \ge |\mathscr{A}|/2$,

$$\rho\big(x^{(t)}, y^{(t)}\big) \ge n(1 - 2p(K) - o(1)),$$

where ρ denotes the Hamming distance. Set $z^{(t)} = (x^{(t)} - y^{(t)})/2$ so that $z^{(t)} \in$

$\{-1, 0, +1\}^n$, at most $n(2p(K) + o(1))$ of its coefficients are zero, and

$$|L_i(z^{(t)})| \leq K\sqrt{n}, \qquad 1 \leq i \leq n.$$

By Lemma 6 we "extend" $z^{(t)}$ to $w^{(t)} \in \{-1, +1\}^n$ with $|L_i(w^{(t)})| \leq K'\sqrt{n}$, $1 \leq i \leq n$. For $K \geq 10$ we may take $K' = K + 1$.

On those coefficients where $x^{(t)}$, $y^{(t)}$ differ, $x^{(t)}$ and $z^{(t)}$, hence $x^{(t)}$ and $w^{(t)}$, are the same. Thus $\rho(x^{(t)}, w^{(t)}) \leq n(2p(K) + o(1))$. If $w^{(s)} = w^{(t)}$, then by the triangle inequality $\rho(x^{(s)}, x^{(t)}) \leq n(4p(K) + o(1))$. Let $q(K) = H(4p(K))$ so that $q(K) \to 0$ as $K \to \infty$. For each s there are at most $2^{n(q(K)+o(1))}$ indices t with $w^{(s)} = w^{(t)}$. There are $|\mathscr{A}|/4 \geq 2^{n(1-c(K)+o(1))}$ indices t. Let δ_K be defined by the equation

$$2 - \delta_K = 2^{1-c(K)-q(K)} \tag{5.13}$$

so that $\operatorname{Lim}_{K\to\infty} \delta_K = 0$. The number of distinct $w^{(t)} \in \{-1, +1\}^n$ with $|L_i(w^{(t)})| \leq K'\sqrt{n}$ for $1 \leq i \leq n$ is at least

$$2^{n(1-c(K)+o(1))}/2^{n(q(K)+o(1))} = (2 - \delta_K + o(1))^n \tag{5.14}$$

giving the theorem.

We apply Theorem 12 directly to count Rudin-Shapiro functions.

COROLLARY 13. *There are at least* $(2 - \delta_K + o(1))^n$ *functions* $f(z) = \varepsilon_1 z + \cdots + \varepsilon_n z^n$ *with* $\|f\| \leq K\sqrt{n}$. *Here* δ_K *is defined for all* $K \geq K_0$, K_0 *an absolute constant. For these* K, $0 < \delta_K < 1$. *Furthermore* $\operatorname{Lim}_{K\to\infty} \delta_K = 0$.

The remarks at the end of §3 suggest, but do not prove, that the number of Rudin-Shapiro functions may be bounded from above by $(2 - \varphi_K + o(1))^n$ for some $\varphi_K > 0$. If this is the case, then Rudin-Shapiro functions are plentiful and rare—both in an exponential sense!

ACKNOWLEDGEMENT. The author would like to thank Gabor Halász for bringing the connection between discrepancy results and the Rudin-Shapiro functions to his attention and for numerous stimulating discussions.

6. Bounding the *i*th form. In this section we wish to find $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ such that $|L_i|$ is bounded by a function of i, independent of the number m of linear forms.

THEOREM 14. *Let*

$$L_i(x_1, \ldots, x_n) = a_{i1}x_1 + \cdots + a_{in}x_n, \qquad 1 \leq i \leq m, \tag{6.1}$$

be m linear forms in n variables with all $|a_{ij}| \leq 1$. *Then there exist* $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ *such that*

$$|L_i(\varepsilon_1, \ldots, \varepsilon_n)| < K\sqrt{i} \tag{6.2}$$

for $1 \leq i \leq m$. *Here K is an absolute constant.*

In [2], Jozsef Beck and this author proved Theorem 14 with $K\sqrt{i}$ replaced by $K\sqrt{i}\,(\ln i)$ in (6.2). The following lemma plays a role analogous to Lemmas 4 and 6 and has a similar proof.

LEMMA 15. *Given m linear forms on n variables as in Theorems* 14 *there exist* $\varepsilon_1,\dots,\varepsilon_n \in \{-1,0,+1\}$ *such that*

$$|\{i: \varepsilon_i = 0\}| < cn, \tag{6.3}$$

$$|L_i(\varepsilon_1,\dots,\varepsilon_n)| \le K\sqrt{i}\,/\ln^2(2n/i), \qquad 1 \le i \le n, \tag{6.4}$$

$$|L_i(\varepsilon_1,\dots,\varepsilon_n)| \le K\sqrt{i}\,/\ln^2(2i/n), \qquad n \le i < m. \tag{6.5}$$

Here $K > 0$ and $c < 1$ are absolute constants.

PROOF OF LEMMA 15. Define T: $\{-1,+1\}^n \to Z^m$ by

$$T(\varepsilon_1,\dots,\varepsilon_n) = (b_1,\dots,b_m), \tag{6.6}$$

where

$$b_i = \begin{cases} \text{nearest integer to } L_i/\left[K\sqrt{i}\,/\ln^2(2n/i)\right], & 1 \le i \le n, \\ \text{nearest integer to } L_i/\left[K\sqrt{i}\,/\ln^2(2i/n)\right], & n < i \le m. \end{cases}$$

We define a subset $B \subset Z^m$ by conditions on $(b_1,\dots,b_m) \in B$. Unlike in Lemmas 4 and 6 the conditions on the coefficients b_i here shall vary with i. We first require that for all positive integers u, s

$$\left|\left\{i: un < i \le (u+1)n, |b_i| \ge s\right\}\right| \le n\xi_{u,s}, \tag{6.7}$$

where we define

$$\xi_{u,s} = 2\exp\left[-\left(\frac{2s+1}{2}\right)K^2u^2\ln^{-4}(2(u+1))\right](100s^2u^2). \tag{6.8}$$

For $un < i \le (u+1)n$, if $|b_i| \ge s$ then $|L_i| \ge ((2s+1)/2)K\sqrt{i}\ln^{-2}(2n/i)$ and thus $|L_i| \ge ((2s+1)/2)K\sqrt{n}\sqrt{u}\ln^{-2}(2(u+1))$. Hence the expected number of i, $un < i \le (u+1)n$, with $|\mathbf{b}_i| \ge s$ is less than $\xi_{u,s}(100s^2u^2)^{-1}$. ($\mathbf{b}_i$ and $\mathbf{L}_i$ are generated from independent uniform $\boldsymbol{\varepsilon}_1,\dots,\boldsymbol{\varepsilon}_n$ as in the previous lemmas.) As

$$\sum_{s=1}^{\infty}\sum_{u=1}^{\infty}(100s^2u^2)^{-1} < .03$$

condition (6.7) is satisfied by at least 97% of the $(\varepsilon_1,\dots,\varepsilon_n) \in \{-1,+1\}^n$.

Define an auxiliary roundoff function R: $R^n \to Z^n$ by $R(L_1,\dots,L_n) = (b_1,\dots,b_n)$, where b_i is the nearest integer to $L_i/[K\sqrt{i}\,/\ln^2(2n/i)]$. Define $D \subset R^n$ by

$$D = \{(L_1,\dots,L_n): \left|\left\{i: 1 \le i \le n, |L_i| > v\sqrt{n}\right\}\right| \le n\left(2e^{-v^2/2}\right)(100v^2) \tag{6.9}$$

for all positive integers $v\}$.

We require of $(b_1,\dots,b_m) \in B$ that

$$(b_1,\dots,b_n) \in R(D). \tag{6.10}$$

Conditions (6.7), (6.10) define B. Our usual argument shows

$$\Pr[(\mathbf{L}_1,\dots,\mathbf{L}_n) \notin D] < \sum_{v=1}^{\infty}(100v^2)^{-1} < .02 \tag{6.11}$$

so

$$\Pr[(\mathbf{b}_1,\dots,\mathbf{b}_n) \notin R(D)] \le \Pr[(\mathbf{L}_1,\dots,\mathbf{L}_n) \notin D] < .02 \tag{6.12}$$

and condition (6.10) is satisfied by at least 98% of the $(\varepsilon_1,\ldots,\varepsilon_n) \in \{-1, +1\}^n$. Together

$$|T^{-1}(B)| \geq .95 \times 2^n. \tag{6.13}$$

Now we bound $|B|$. This is somewhat complicated but elementary and the reader may wish to jump to the result (6.44) that $|B| \leq 2^{n\nu}$, where $\operatorname{Lim}_{K\to\infty} \nu = 0$ on first reading.

Let B_2 be the set of $(b_{n+1},\ldots,b_m)$ satisfying (6.7). Then

$$|B_2| \leq \prod_{u=1}^{\infty} \prod_{s=1}^{\infty} \binom{n}{\xi_{u,s} n} 2^{\xi_{u,s} n} \leq 2^{n\beta}, \tag{6.14}$$

where

$$\beta = \sum_{u=1}^{\infty} \sum_{s=1}^{\infty} (H(\xi_{u,s}) + \xi_{u,s}). \tag{6.15}$$

If we consider β as a function of K, then $\operatorname{Lim}_{K\to\infty} \beta = 0$. We write this

$$\beta = o_K(1). \tag{6.16}$$

(To show (6.16) we may note, for example, that for K sufficiently large $\xi_{u,s} \leq \exp[-K^2u^2s^2/100]$ for all u, s.)

Let B_1 be the set of $(b_1,\ldots,b_n)$ satisfying (6.10). Bounding $|B_1|$ is more complex.

Let $C(x)$ denote the number of choices for the nearest integer to y where y ranges over $[-x, +x]$. Then

$$C(x) = 2n + 1, \qquad n - \tfrac{1}{2} \leq x < n + \tfrac{1}{2}. \tag{6.17}$$

We shall require two quite rough inequalities:

$$C(x) \leq 6x, \tag{6.18}$$

valid whenever $C(x) > 1$, and

$$C(x\alpha) \leq 3\alpha C(x), \tag{6.19}$$

valid for all $x > 0$ and $\alpha \geq 1$. Let $g(i, v)$ denote the number of choices for b_i given that $|L_i| \leq v\sqrt{n}$.

$$g(i, v) = C(x), \qquad x = v\sqrt{n}\,/\big[K\sqrt{i}\,/\ln^2(2n/i)\big]. \tag{6.20}$$

Let $M = M(K)$ be integral so that

$$\lim_{K\to\infty} M(K)/K = 0, \tag{6.21}$$

$$\lim_{K\to\infty} M(K) = \infty. \tag{6.22}$$

For definiteness we may take

$$M = [\sqrt{K}]. \tag{6.23}$$

For all i, $1 \leq i \leq n$, and all $v \geq M$

$$g(i, v) \leq 3(v/M)g(i, M). \tag{6.24}$$

Let $\mathcal{F}$ be the family of all sequences

$$\mathcal{S} = (S_M, S_{M+1}, \dots) \tag{6.25}$$

such that all $S_v \subset \{1, \dots, n\}$ and

$$|S_v| \le \alpha_v n, \qquad \alpha_v = 2e^{-v^2/2}(100v^2), \tag{6.26}$$

for all $v \ge M$. Given $\mathcal{S} \in \mathcal{F}$ define

$$D(\mathcal{S}) = \left\{(L_1, \dots, L_n): \left\{i: 1 \le i \le n, |L_i| > v\sqrt{n}\right\} = S_v \right. \tag{6.27}$$
$$\left. \text{for all integers } v \ge M\right\}.$$

Then

$$B_1 = \bigcup_{\mathcal{S} \in \mathcal{F}} R[D(\mathcal{S})]. \tag{6.28}$$

We shall bound $|B_1|$ by bounding $|\mathcal{F}|$ and placing a uniform bound on $|R[D(\mathcal{S})]|$.

$$|\mathcal{F}| \le \prod_{v=M}^{\infty} \binom{n}{\alpha_s n} \le 2^{n\gamma}, \tag{6.29}$$

where

$$\gamma = \sum_{v=M}^{\infty} H(\alpha_v). \tag{6.30}$$

We consider γ a function of K. Then, using (6.21),

$$\gamma = o_K(1). \tag{6.31}$$

Now we fix $\mathcal{S} = (S_M, S_{M+1}, \dots) \in \mathcal{F}$ and bound $|R[D(\mathcal{S})]|$. If $i \notin S_M$, then $|L_i| \le M\sqrt{n}$ and there are at most $g(i, M)$ choices for b_i. If $i \notin S_v$, then $|L_i| \le v\sqrt{n}$ and there are at most $g(i, v) \le 3(v/M)g(i, M)$ choices for b_i (using (6.24)). Thus

$$|R[D(\mathcal{S})]| \le \left[\prod_{i \notin S_M} g(i, M)\right]\left[\prod_{v=M+1}^{\infty} \prod_{\substack{i \notin S_v \\ i \in S_{v-1}}} 3(v/M)g(i, M)\right] \tag{6.32}$$
$$= \left[\prod_{i=1}^{n} g(i, n)\right] \prod_{v=M+1}^{\infty} [3(v/M)]^{|S_{v-1}|}.$$

(The left factor gives the number of $(b_1, \dots, b_n)$ given that all $|L_i| \le M\sqrt{n}$. The right factor and $|\mathcal{F}|$ allow for the possibility that some $|L_i|$ are larger.)

$$\prod_{v=M+1}^{\infty} [3(v/M)]^{|S_{v-1}|} \le \prod_{v=M+1}^{\infty} [3(v/M)]^{n\alpha_{v-1}} = 2^{n\delta}, \tag{6.33}$$

where

$$\delta = \sum_{v=M+1}^{\infty} \alpha_{v-1} \log_2(3v/M) = o_K(1) \tag{6.34}$$

by (6.22) and the rapid decay of the α_v. Set

$$A = \prod_{i=1}^{n} g(i, M) = \prod_{i=1}^{n} C\left[(M/K)(n/i)^{1/2} \ln^2(2n/i)\right]. \tag{6.35}$$

Let $T = T(K)$ be that real number such that

$$(M/K)T^{1/2}\ln^2(2T) = 1/2. \tag{6.36}$$

Relation (6.21) implies

$$\lim_{K\to\infty} T = \infty. \tag{6.37}$$

For $i \le n/T$, $(M/K)(n/i)^{1/2}\ln^2(2n/i) \le 1/2$ so that $C = 1$. Combining (6.35) with this observation and (6.18)

$$A \le \prod_{i=1}^{n/T} 6(M/K)(n/i)^{1/2}\ln^2(2n/i). \tag{6.38}$$

By elementary methods we bound

$$\prod_{i=1}^{n/T} i^{-1/2} = (n/T)!^{-1/2} \le \left[(n/eT)^{n/T}\right]^{-1/2} \tag{6.39}$$

and

$$\prod_{i=1}^{n/T} \ln^2(2n/i) \le \left[c_1\ln^2(2T)\right]^{n/T}, \tag{6.40}$$

where c_1 is an absolute constant. (The $\ln^2(2n/i)$ term plays a negligible role. It is required when Lemma 15 is iterated to produce Theorem 14.) Together

$$A \le \left[6(M/K)(n/T)^{1/2}(n/eT)^{-1/2}c_1\ln^2(2T)\right]^{n/T} = 2^{n\mu}, \tag{6.41}$$

where (using (6.37))

$$\mu = T^{-1}\log_2\left[6(M/K)e^{1/2}c_1\ln^2(2T)\right] = o_K(1). \tag{6.42}$$

We combine (6.41) with (6.32), (6.28), (6.14) to bound

$$\begin{aligned} |B| &\le |B_1|\,|B_2| \le |B_1|\,|\mathscr{F}|\max|R(D(\mathscr{S}))| \\ &\le 2^{n\beta}2^{n\gamma}2^{n(\delta+\mu)} = 2^{n\nu}, \end{aligned} \tag{6.43}$$

where, by (6.16), (6.31), (6.34) and (6.42)

$$\nu = \beta + \gamma + \delta + \mu = o_K(1). \tag{6.44}$$

Let K be any constant sufficiently large so that $\nu < 1 + \log_2(.95)$ and set $\nu' = \nu - \log_2(.95)$. (The introduction of ν' is a purely technical device to allow Lemma 15 to hold for *all* $n \ge 1$.) As with Lemma 4 we find $\mathscr{A}$ on which T is constant with

$$|\mathscr{A}| \ge |T^{-1}(B)|/|B| \ge .95 \times 2^{n(1-\nu)} \tag{6.45}$$

using (6.13), (6.43). As $n \ge 1$ we may write

$$|\mathscr{A}| \ge 2^{n(1-\nu')}. \tag{6.46}$$

Let $p < 1/2$ satisfy $H(\frac{1}{2} - p) = 1 - \nu'$. Then, by (2.22), $\operatorname{diam}(\mathscr{A}) \ge (1 - 2p)n$. Select $\vec{\varepsilon}', \vec{\varepsilon}'' \in \mathscr{A}$ at maximal Hamming distance and set

$$\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n) = (\vec{\varepsilon}' - \vec{\varepsilon}'')/2. \tag{6.47}$$

The conditions (6.3), (6.4), (6.5) of Lemma 15 are then met with the above K and with $c = 2p$.

PROOF OF THEOREM 14. We iterate Lemma 15, beginning with n variables. Suppose $n = n_0, n_1, \ldots, n_s$ are the values for the number of remaining variables during the iterations. Then, by (6.3), $n_{j+1} < cn_j$, all j. For any i

$$|L_i(\varepsilon_1, \ldots, \varepsilon_n)| \le \sum_{n_j < i} K\sqrt{i}/\ln^2(2i/n_j) + \sum_{n_j \ge i} K\sqrt{i}/\ln^2(2n_j/i). \tag{6.48}$$

Let J be that index such that $n_{J+1} < i \le n_J$. Then $n_{J+1+u} \le ic^u$ for $u \ge 0$ and $n_{J-u} \ge ic^{-u}$ for $u \ge 0$. Then

$$\begin{aligned} |L_i| &\le K\sqrt{i}\left[\sum_{u=0}^{\infty} \ln^{-2}(2i/ic^u) + \sum_{u=0}^{\infty} \ln^{-2}(2i/ic^{-u})\right] \\ &\le K\sqrt{i}\left[2\sum_{u=0}^{\infty} [u \ln c^{-1} + \ln 2]^{-2}\right]. \end{aligned} \tag{6.49}$$

Let k denote the bracketed sum, which is finite as Σu^{-2} converges. (The convergence of this sequence was, in fact, the purpose of the $\ln^2(2n/i)$ and $\ln^2(2i/n)$ factors in (6.4), (6.5).) Then Theorem 14 holds with Kk as the constant "K" of (6.2).

We apply Theorem 14 to give a proof of the existence of Rudin-Shapiro functions (Theorem 11) which does not require (5.4). Set

$$f(z) = \varepsilon_1 z + \cdots + \varepsilon_n z^n \tag{6.50}$$

and let $\omega = e^{2\pi i/n}$. As in §5 we set

$$L_j = \operatorname{Re}[f(\omega^j)], \qquad L_{j+n} = \operatorname{Im}[f(\omega^j)] \tag{6.51}$$

for $0 \le j < n$. We define further forms by

$$L_{j+2n} = \operatorname{Re}[f'(\omega^j)]/n, \qquad L_{j+3n} = \operatorname{Im}[f'(\omega^j)]/n \tag{6.52}$$

and, more generally, for $0 \le s \le n$, $0 \le j < n$

$$L_{j+2sn} = \operatorname{Re}[f^{(s)}(\omega^j)]/n^s, \qquad L_{j+(2s+1)n} = \operatorname{Im}[f^{(s)}(\omega^j)]/n^s. \tag{6.53}$$

The denominator n^s assures that all coefficients have absolute value at most one. By Theorem 14 there exist $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ with $|L_i| \le K\sqrt{i}$ for all i. In terms of f

$$\begin{aligned} |\operatorname{Re}[f^{(s)}(\omega^j)]| &\le Kn^s\sqrt{j + 2sn} \le Kn^s\sqrt{n}\sqrt{2(s+1)}, \\ |\operatorname{Im}[f^{(s)}(\omega^j)]| &\le Kn^s\sqrt{j + (2s+1)n} \le Kn^s\sqrt{n}\sqrt{2(s+1)} \end{aligned} \tag{6.54}$$

so that

$$|f^{(s)}(\omega^j)| \le 2Kn^s\sqrt{n}\sqrt{s+1}. \tag{6.55}$$

Let $z = e^{2\pi i\theta}$ be an arbitrary point on the unit circle. For some j, $|z - \omega^j| \le \pi/n$. We bound $f(z)$ by taking the Taylor series about ω^j. (Since f itself is a polynomial of degree n the series has only n terms.)

$$f(z) = \sum_{s=0}^{n} f^{(s)}(\omega^j)(z - \omega^j)^s/s! \tag{6.56}$$

Then

$$|f(z)| \le \sum_{s=0}^{n} |f^{(s)}(\omega^j)|\,|z-\omega^j|^s/s! \tag{6.57}$$

$$\le \sum_{s=0}^{n} 2Kn^s\sqrt{n}\sqrt{s+1}\,(\pi/n)^s/s!$$

$$= K\sqrt{n}\sum_{s=0}^{n} 2\sqrt{s+1}\,\pi^s/s! \le K'\sqrt{n}$$

for $K' = cK$ where the definition of c,

$$c = \sum_{s=0}^{\infty} 2\sqrt{s+1}\,\pi^s/s!, \tag{6.58}$$

is given by a clearly convergent series.

7. The Komlós Conjecture. Let $|\cdot|$ denote the Euclidean norm and $\|\cdot\|$ the L^∞ norm. Let $v_1,\dots,v_s \in R^n$ with $|v_i| \le 1$. János Komlós has conjectured that there exist $\varepsilon_1,\dots,\varepsilon_s \in \{-1,+1\}$ such that

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_s v_s\| \le K, \tag{7.1}$$

where K is an absolute constant with s, n arbitrary positive integers. When all coefficients of each v_i are $\pm n^{-1/2}$, Theorem 1 or, when $s > n$, Theorem 9 yield directly the Komlós Conjecture. The full Komlós Conjecture has resisted our efforts but the following result is close.

THEOREM 16. *Let $v_1,\dots,v_r \in R^n$ with all $|v_i| \le 1$ and $r \le n$. Then there exist $\varepsilon_1,\dots,\varepsilon_r \in \{-1,0,+1\}$ such that*

$$|\{i\colon \varepsilon_i = 0\}| < cr, \tag{7.2}$$

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_r v_r\| \le K. \tag{7.3}$$

Here $c < 1$ and K are positive absolute constants.

PROOF. Let $v_j = (a_{1j},\dots,a_{rj})$, $1 \le j \le r$, and set

$$\sigma_i = \left[a_{i1}^2 + \cdots + a_{ir}^2\right]^{1/2}, \qquad 1 \le i \le n, \tag{7.4}$$

so that

$$\sigma_1^2 + \cdots + \sigma_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{r} a_{ij}^2 = \sum_{j=1}^{r}\left[\sum_{i=1}^{n} a_{ij}^2\right] \le \sum_{j=1}^{r} 1 = r. \tag{7.5}$$

Set

$$L_i = L_i(\varepsilon_1,\dots,\varepsilon_r) = a_{i1}\varepsilon_1 + \cdots + a_{ir}\varepsilon_r, \qquad 1 \le i \le n. \tag{7.6}$$

Then $\varepsilon_1 v_1 + \cdots + \varepsilon_r v_r = (L_1,\dots,L_n)$ and σ_i is the standard deviation of $\mathbf{L}_i$ generated by uniform independent $\boldsymbol{\varepsilon}_1,\dots,\boldsymbol{\varepsilon}_r \in \{-1,+1\}$.

Our proof will be patterned after the proof of Lemma 15.

Define a map T: $\{-1,+1\}^r \to Z^n$ by $T(\varepsilon_1,\dots,\varepsilon_r) = (b_1,\dots,b_n)$, where b_i is the nearest integer to $L_i/2K$. Order the row indices i by the values σ_i in decreasing

order. Let t be such that

$$\sigma_1^2 \geq \cdots \geq \sigma_t^2 \geq 1 > \sigma_{t+1}^2 \geq \cdots \geq \sigma_n^2. \tag{7.7}$$

For each integer $v \geq 0$ set

$$Q_v = \{i: 2^{-v} > \sigma_i \geq 2^{-v-1}\}. \tag{7.8}$$

We define a subset $B \subset Z^n$ by conditions on $(b_1,\ldots,b_n)$. For the "small" rows we require

$$|\{i \in Q_v: |b_i| \geq s\}| \leq \xi_{v,s}|Q_v| \tag{7.9}$$

for each $v \geq 0$, $s \geq 1$ where we set

$$\xi_{v,s} = 20 \cdot 2^{v+1}2^s e^{-(2s-1)^2K^2 2^{v-1}} \tag{7.10}$$

for convenience. For the "large" rows define an Auxiliary Roundoff function R: $R^t \to Z^t$ by $R(L_1,\ldots,L_t) = (b_1,\ldots,b_t)$, where b_i is the nearest integer to $L_i/2K$. Let M (as in (6.21), (6.22)) be an integral function of K such that

$$\lim_{K\to\infty} M(K)/K = 0, \tag{7.11}$$

$$\lim_{K\to\infty} M(K) = \infty. \tag{7.12}$$

Again, for definiteness we may take $M = [\sqrt{K}]$. Define $D \subset R^t$ by

$$D = \{(L_1,\ldots,L_t): |\{i: 1 \leq i \leq t, |L_i| \geq M\sigma_i(2s-1)\}| \leq \alpha_s t \text{ for all } s \geq 1\}, \tag{7.13}$$

where we define

$$\alpha_s = 20 \cdot 2^s e^{-M^2(2s-1)^2/2} \tag{7.14}$$

for convenience. We require

$$(b_1,\ldots,b_t) \in R(D). \tag{7.15}$$

Conditions (7.9), (7.15) define B.

Let $\varepsilon_1,\ldots,\varepsilon_r \in \{-1, +1\}$ be uniform and independent. We have defined $\xi_{v,s}$ so that

$$E\left[|\{i \in Q_v: |\mathbf{b}_i| \geq s\}|\right] \leq \xi_{v,s}|Q_v|(20 \cdot 2^{v+1}2^s)^{-1} \tag{7.16}$$

so that the probability (7.9) does not hold for a particular v, s is bounded by $(20 \cdot 2^{v+1}2^s)^{-1}$. Similarly, we have defined α_s so that

$$E\left[|\{i: 1 \leq i \leq t, |\mathbf{L}_i| \geq M\sigma_i(2s-1)\}|\right] \leq \alpha_s t(20 \cdot 2^s)^{-1} \tag{7.17}$$

so that the probabilty (7.15) does not hold because of a particular s is bounded by $(20 \cdot 2^s)^{-1}$. Thus

$$\Pr[(\mathbf{b}_1,\ldots,\mathbf{b}_n) \notin B] \leq \sum_{v=0}^{\infty}\sum_{s=1}^{\infty}(20 \cdot 2^{v+1}2^s)^{-1} + \sum_{s=1}^{\infty}(20 \cdot 2^s)^{-1} = .1 \tag{7.18}$$

so

$$|T^{-1}(B)| \geq .9 \cdot 2^r. \tag{7.19}$$

Now we bound $|B|$. This is somewhat complicated but elementary and the reader may wish to jump to the result (7.39) that $|B| \leq 2^{r\nu}$ where $\operatorname{Lim}_{K\to\infty} \nu = 0$ at first reading.

Let B_2 be the set of $(b_{t+1},\ldots,b_n)$ satisfying (7.9) for all $v \geq 0$, $s \geq 1$. Each $i \in Q_v$ has $\sigma_i^2 \geq 2^{-2v-2}$ so, using (7.5), $|Q_v| \leq r2^{2v+2}$. Then

$$\begin{aligned} |B_2| &\leq \prod_{v=0}^{\infty}\prod_{s=1}^{\infty}\binom{|Q_v|}{\xi_{v,s}|Q_v|}2^{\xi_{v,s}|Q_v|} \\ &\leq \prod_{v=0}^{\infty}\prod_{s=1}^{\infty}\binom{r2^{2v+2}}{r2^{2v+2}\xi_{v,s}}2^{r2^{2v+2}\xi_{v,s}} = 2^{\beta r}, \end{aligned} \tag{7.20}$$

where

$$\beta = \sum_{v=0}^{\infty}\sum_{s=1}^{\infty} 2^{2v+2}\left[H(\xi_{v,s}) + \xi_{v,s}\right]. \tag{7.21}$$

Then β is a function of K and, noting in (7.10) the dominance of an $\exp(-2s^2K^2 2^v)$ term when either s or v is large,

$$\beta = o_K(1). \tag{7.22}$$

Let B_1 be the set of $(b_1,\ldots,b_t)$ satisfying (7.15). Bounding $|B_1|$ is more complex. Since $\sigma_i^2 \geq 1$ for $1 \leq i \leq t$, (7.5) implies $t \leq r$. Let $\mathcal{F}$ be the collection of sequences

$$\mathcal{S} = (S_1, S_2, \ldots) \tag{7.23}$$

such that all $S_s \subset \{1,\ldots,t\}$ with $|S_s| \leq \alpha_s t$. For each $\mathcal{S} \in \mathcal{F}$ set

$$D(\mathcal{S}) = \left\{(L_1,\ldots,L_t)\colon \left\{i\colon 1 \leq i \leq t, |L_i| \geq M\sigma_i(2s-1)\right\} = S_s \text{ for all } s \geq 1\right\} \tag{7.24}$$

so that

$$B_1 = \bigcup_{\mathcal{S}\in\mathcal{F}} R(D(\mathcal{S})). \tag{7.25}$$

We shall bound $|B_1|$ by bounding $|\mathcal{F}|$ and placing a uniform bound on $|R(D(\mathcal{S}))|$.

$$|\mathcal{F}| \leq \prod_{s=1}^{\infty}\binom{t}{\alpha_s t} \leq \prod_{s=1}^{\infty}\binom{r}{\alpha_s r} \leq 2^{r\gamma}, \tag{7.26}$$

where

$$\gamma = \sum_{s=1}^{\infty} H(\alpha_s) \tag{7.27}$$

and, considering γ as a function of K,

$$\gamma = o_K(1). \tag{7.28}$$

We fix $\mathcal{S}$ and bound $|R(D(\mathcal{S}))|$. Let C be defined as in §6, with properties (6.18), (6.19) most pertinent. If $i \notin S_{s+1}$, then $|L_i| \leq (2s+1)M\sigma_i$ so there are at most

$C((2s+1)\sigma_i M/2K)$ choices for b_i. Thus

$$(7.29)\quad |R(D(\mathscr{S}))| \le \prod_{i\notin S_1} C(\sigma_i M/2K) \prod_{s=1}^{\infty} \prod_{\substack{i\in S_s \\ i\notin S_{s+1}}} C((2s+1)\sigma_i M/2K)$$

$$\le \prod_{i\notin S_1} C(\sigma_i M/2K) \prod_{s=1}^{\infty} \prod_{\substack{i\in S_s \\ i\notin S_{s+1}}} 3(2s+1)C(\sigma_i M/2K)$$

$$= \prod_{i=1}^{t} C(\sigma_i M/2K) \prod_{s=1}^{\infty} (3(2s+1))^{|S_s|}.$$

(The left factor gives the number of $(b_1,\dots,b_t)$ if all $|L_i| \le M\sigma_i$. The other factor and $|\mathscr{F}|$ allow for the possibility that some $|L_i|$ are larger.)

$$(7.30)\quad \prod_{s=1}^{\infty} (3(2s+1))^{|S_s|} \le \prod_{s=1}^{\infty} (3(2s+1))^{\alpha_s t} = 2^{t\delta} \le 2^{r\delta},$$

where

$$(7.31)\quad \delta = \sum_{s=1}^{\infty} \alpha_s \log_2(3(2s+1))$$

and, considering δ as a function of K

$$(7.32)\quad \delta = o_K(1).$$

Now we bound

$$(7.33)\quad A = \prod_{i=1}^{t} C(\sigma_i M/2K).$$

Let v be the maximal index, $1 \le v \le t$, such that $\sigma_v M/2K > 1/2$. For $i > v$, $C(\sigma_i M/2K) = 1$ and the factor may be deleted. Using (6.18)

$$(7.34)\quad A \le \prod_{i=1}^{v} \sigma_i (3M/K).$$

LEMMA 17. *Let $\varepsilon > 0$ and suppose $x_1 + \cdots + x_v \le W$. Then*

$$(7.35)\quad \prod_{i=1}^{v} (x_i \varepsilon) \le e^{W\varepsilon/e}.$$

PROOF. First assume $\varepsilon = 1$. By elementary calculus $(\ln z)/z \le 1/e$ for all $z > 0$ so $\ln x_1 + \cdots + \ln x_v \le (x_1/e) + \cdots + (x_v/e) \le W/e$ which yields (7.33) by exponentiating both sides. Arbitrary $\varepsilon > 0$ may be reduced to the $\varepsilon = 1$ case by setting $y_i = x_i\varepsilon$.

Apply Lemma 17 with $\varepsilon = 9M^2/K^2$ and $W = r$. Squaring (7.34) and using (7.5)

$$(7.36)\quad A^2 = \prod_{i=1}^{v} \sigma_i^2 (9M^2/K^2) \le e^{r(9M^2/K^2)/e}$$

so that

$$A \le 2^{r\mu}, \qquad \mu = (9/2e\ln 2)(M/K)^2, \tag{7.37}$$

and, using (7.11)

$$\mu = o_K(1). \tag{7.38}$$

Combining (7.20), (7.26), (7.37)

$$|B| \le |B_2|\,|B_1| \le 2^{r(\beta+\gamma+\mu)} = 2^{r\nu}, \tag{7.39}$$

where

$$\nu = o_K(1). \tag{7.40}$$

Let K be any constant sufficiently large so that $\nu < 1 + \log_2(.9)$ and set $\nu' = \nu - \log_2(.9)$. (The introduction of ν' is a purely technical device to allow Theorem 16 to hold for *all* $r \ge 1$.) As with Lemma 4 we find $\mathscr{A}$ on which T is constant with

$$|\mathscr{A}| \ge |T^{-1}(B)|/|B| \ge .9 \cdot 2^{r(1-\nu)} \tag{7.41}$$

using (7.19), (7.39). As $r \ge 1$ we may write

$$|\mathscr{A}| \ge 2^{r(1-\nu')}. \tag{7.42}$$

Let $p < \frac{1}{2}$ satisfy $H(\frac{1}{2} - p) = 1 - \nu'$. Then by (2.22) $\operatorname{diam}(\mathscr{A}) \ge (1 - 2p)r$. Select $\vec{\varepsilon}', \vec{\varepsilon}'' \in \mathscr{A}$ at maximal Hamming distance and set

$$\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_r) = (\vec{\varepsilon}' - \vec{\varepsilon}'')/2. \tag{7.43}$$

The conditions (7.2), (7.3) of Theorem 16 are met with the above K and with $c = 2p$.

The procedure given by Theorem 16 can certainly be iterated. However, at each stage the norm $\|\varepsilon_1 v_1 + \cdots + \varepsilon_r v_r\|$ is bounded by the same absolute constant. (Notice that K is independent of r so that reduction of r to cr does not affect it.) If we begin with $r \le n$ we need apply Theorem 16 only t times, where $nc^t < 1$, until all ε_i have been determined. Therefore there exist $\varepsilon_1, \ldots, \varepsilon_r \in \{-1, +1\}$ with

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_r v_r\| \le K' \ln n, \tag{7.44}$$

where $K' = K/\ln(1/c)$. Suppose $r > n$. By Theorem 9 there exist, after reordering, $p_1, \ldots, p_n \in [-1, +1]$, $\varepsilon_{n+1}, \ldots, \varepsilon_r \in \{-1, +1\}$ such that

$$p_1 v_1 + \cdots + p_n v_n + \varepsilon_{n+1} v_{n+1} + \cdots + \varepsilon_r v_r = 0 \tag{7.45}$$

and, applying the method of Corollary 8, there exist $\varepsilon_1, \ldots, \varepsilon_r \in \{-1, +1\}$ so that

$$\|(\varepsilon_1 - p_1)v_1 + \cdots + (\varepsilon_n - p_n)v_n\| \le 2K' \ln n. \tag{7.46}$$

We combine these results.

COROLLARY 18. *Let $v_1, \ldots, v_r \in R^n$, $|v_i| \le 1$. Then there exist $\varepsilon_1, \ldots, \varepsilon_r \in \{-1, +1\}$ such that*

$$\|\varepsilon_1 v_1 + \cdots + \varepsilon_r v_r\| \le K'' \ln n. \tag{7.47}$$

Here K'' is an absolute constant.

REMARK. A surprising aspect of Theorem 16 is that n can be arbitrarily large compared to r. The "worst case" is when $\sigma_1^2 = \cdots = \sigma_r^2 = 1$ and $\sigma_{r+1} = \cdots = \sigma_n = 0$. We apply Theorem 16 and reduce to, say, cr variables with new $\sigma_1^*, \ldots, \sigma_n^*$. Here the "worst case" would be if $\sigma_1^* = \cdots = \sigma_{cr}^* = 1$, the rest equal zero. If we could find a determination of all but cr of the variables which "split" the rows in the sense that each $\sigma_i^* \sim c\sigma_i$, then perhaps we could show the full Komlós Conjecture.

REMARK. Let $\mathscr{F} = \{A_1, \ldots, A_m\}$ be a family of subsets of $\{1, \ldots, n\}$ such that every point is in at most d sets. The Komlós Conjecture, applied to the column vectors of the incidence matrix, would imply the existence of a two-coloring χ of $\{1, \ldots, n\}$ such that $\operatorname{disc}(A_i) \leq K\sqrt{d}$ for all i, $1 \leq i \leq m$. This would improve the result of J. Beck and T. Fiala [1] that such a χ exists with $\operatorname{disc}(A_i) \leq 2d - 1$ for all i, $1 \leq i \leq m$. We are able in the special case when $\mathscr{F}$ is the set of lines of a projective plane of order p to use the methods of this section to show the existence of a two-coloring χ such that every line has discrepency at most $K\sqrt{p+1}$. The proof will appear elsewhere.

8. Best possible. Here we give two proofs that Theorem 1 is "best possible" up to the constant factor. Similar results can be shown when the number of linear forms does not equal the number of variables (Theorem 7), discrepancies of sets (Corollary 2), and the Rudin-Shapiro functions (Theorem 11).

THEOREM 19. *There exist* $a_{ij} \in \{-1, +1\}$, $1 \leq i, j \leq n$, *with the property that for all* $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$

$$|L_i(\varepsilon_1, \ldots, \varepsilon_n)| > k\sqrt{n}\,(1 + o(1)) \tag{8.1}$$

for some i, $1 \leq i \leq n$. *Here* L_i *is given by* (1.1), k *is an absolute constant, and* $o(1)$ *is with respect to* n.

PROOF 1. Let $k \sim .67$ be that real number such that $1 - 2\Phi(-k) = .5$. Let $\delta > 0$ be arbitrarily small and set $k' = k + \delta$. From (3.38) (and using the notation of that section) there exists A such that $|\mathscr{A}_A| \leq 2^n(1 - 2\Phi(-k') - o(1))^n$. For n sufficiently large (so that δ outweighs the $o(1)$ term), $|\mathscr{A}_A| < 2^n(.5)^n$ so that $\mathscr{A}_A = \varnothing$ as desired.

PROOF 2. Let $k = 1$. Let $A = (a_{ij})$ be a Hadamard matrix of order n with $v_1, \ldots, v_n$ its orthogonal columns. Each $|v_i| = n^{1/2}$ so for all $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$

$$|\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n| = \left[|v_1|^2 + \cdots + |v_n|^2\right]^{1/2} = n. \tag{8.2}$$

But $\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n = (L_1, \ldots, L_n)$ so that $L_1^2 + \cdots + L_n^2 = n^2$, hence some $|L_i| \geq n^{1/2}$.

While Hadamard matrices do not exist for all orders they are asymptotically dense (simply from the values $n = 4^a 12^b$) in the sense that for all n there is an $n' = n(1 - o(1)) \leq n$ for which a Hadamard matrix does exist. Let A be a Hadamard matrix of order n' bordered by $n - n'$ rows and columns of zeros. For all $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}$ some $|L_i| \geq \sqrt{n'} = \sqrt{n}\,(1 + o(1))$.

9. Six standard deviations suffice. Here we show that Theorem 1 is valid with a moderate value of K, $K = 5.32$. Let Φ be the cdf of the standard normal distribution, as given in (3.34). Let $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_r \in \{-1, +1\}$ be uniform and independent and

let $a_1,\dots,a_r \in [-1,+1]$. The Central Limit Theorem implies

$$\Pr\left[|\varepsilon_1 a_1 + \cdots + \varepsilon_r a_r| > k\sqrt{r}\right] < 2\Phi(-k) + o(1), \tag{9.1}$$

where k is fixed and r approaches infinity. Using (9.1) instead of (1.13) we effectively replace $\exp(-k^2/2)$ with $\Phi(-k)$. (When $k = 4$ these values are roughly 3.3×10^{-4} and 3.2×10^{-5} respectively.) Lemma 6 is generalized as follows.

LEMMA 20. *Let $\alpha \le 1$ be fixed. Let K, an infinite sequence $\gamma_1,\gamma_2,\dots$, β and p be given satisfying*

$$\beta = \alpha^{-1}\sum_{s=1}^{\infty} H\big(2\Phi(-K(2s-1))\gamma_s\big) + 2\Phi(-K(2s-1))\gamma_s,$$
$$\sum_{s=1}^{\infty} \gamma_s^{-1} < 1, \qquad H\left(\tfrac{1}{2} - p\right) > 1 - \beta. \tag{9.2}$$

Then for n sufficiently large the following holds. Given n linear forms $L_i(x_1,\dots,x_r) = a_{i1}x_1 + \cdots + a_{ir}x_r$, $1 \le i \le r$, in r variables with $r \le \alpha n$ and with all $|a_{ij}| \le 1$ there exist $\varepsilon_1,\dots,\varepsilon_r \in \{-1,0,+1\}$ such that

$$\left|\{i\colon \varepsilon_i = 0\}\right| \le 2p(\alpha n), \tag{9.3}$$

$$|L_i(\varepsilon_1,\dots,\varepsilon_r)| \le K\sqrt{r} \le K\sqrt{\alpha}\sqrt{n}, \qquad 1 \le i \le n. \tag{9.4}$$

PROOF. Adding additional variables with zero coefficients if necessary, it is convenient to assume $r = \alpha n$. As in Lemma 6 we define $T\colon \{-1,+1\}^r \to Z^n$ by $T(\varepsilon_1,\dots,\varepsilon_r) = (b_1,\dots,b_n)$, where b_i is the nearest integer to $L_i/K\sqrt{r}$. Define $B \subset Z^n$ by

$$B = \Big\{(b_1,\dots,b_n) \in Z^n\colon \left|\{i\colon |b_i| \ge s\}\right| \le n\big(2\Phi(-K(2s-1))\big)\gamma_s \text{ for all } s \ge 1\Big\}. \tag{9.5}$$

With $\varepsilon_i \in \{-1,+1\}$ uniform and independent

$$\Pr\left[|\mathbf{b}_i| \ge s\right] = \Pr\left[|\mathbf{L}_i| \ge K(2s-1)\sqrt{r}\right] \le 2\Phi(-K(2s-1)) \tag{9.6}$$

by (9.2). The expected number of i with $|\mathbf{b}_i| \ge s$ is at most $n(2\Phi(-K(2s-1)))$ and so

$$\Pr\left[\left|\{i\colon |\mathbf{b}_i| \ge s\}\right| \ge n\big(2\Phi(-K(2s-1))\big)\gamma_s\right] < \gamma_s^{-1}. \tag{9.7}$$

Setting $c = 1 - \sum_{s=1}^{\infty}\gamma_s^{-1} > 0$

$$\Pr\left[(\mathbf{b}_1,\dots,\mathbf{b}_n) \in B\right] > c, \quad \text{i.e. } \left|T^{-1}(B)\right| > c2^r. \tag{9.8}$$

The method used in Lemma 4 shows

$$|B| \le 2^{(\alpha\beta)n} = 2^{\beta r}. \tag{9.9}$$

We find $\mathscr{A}$ on which T is constant with

$$|\mathscr{A}| \ge c2^r/2^{\beta r} > 2^{rH(1/2-p)} \tag{9.10}$$

for n sufficiently large (to absorb c). Then $\operatorname{diam}(\mathscr{A}) \ge r(1-2p)$. Let $\vec{\varepsilon}', \vec{\varepsilon}'' \in \mathscr{A}$ at maximal distance. We set $\vec{\varepsilon} = (\vec{\varepsilon}' - \vec{\varepsilon}'')/2$ and complete the proof identically with Lemma 4.

We wish to bound a function $G(\alpha)$, $\alpha \in (0,1]$, such that given n linear forms L_i in $r < \alpha n$ variables with all $|a_{ij}| \leq 1$ there exist $\varepsilon_1, \ldots, \varepsilon_r \in \{-1, +1\}$ such that all $|L_i| \leq G(\alpha)\sqrt{n}$.

If α, K, p satisfy (9.2) for some $\beta, \gamma_1, \gamma_2, \ldots$, then

$$G(\alpha) \leq K\sqrt{\alpha} + G(2p\alpha) \tag{9.11}$$

by Lemma 19. Theorem 7 gives

$$G(\alpha) \leq 11\sqrt{\alpha}\sqrt{\ln(2\alpha^{-1})} \tag{9.12}$$

for all $\alpha \leq 1$.

We use (9.11), (9.12) to bound $G(1)$, the absolute constant of Theorem 1. Equation (9.11) leads to a tradeoff between K and p which we do not here fully optimize. The terms $\Phi(-K(2s-1))$ decrease so rapidly in s (asymptotically $\Phi(-t) \sim \exp(-t^2/2)/t\sqrt{2\pi}$) that all summations over s are dominated by the $s = 1$ term. When ε is small $H(\varepsilon) \sim \varepsilon \log_2(1/\varepsilon) \gg \varepsilon$ so that $H(\varepsilon) + \varepsilon \sim \varepsilon \log_2(1/\varepsilon)$. We shall always choose $\gamma_1 = 1.1$, $\gamma_s = 20^{s-1}$ for $s \geq 2$. Then $\beta \sim \alpha^{-1}H(2\Phi(-K)(1.1))$. For ε small $H(\frac{1}{2} - \varepsilon) \sim 1 - (2/\ln(2))\varepsilon^2$ so when β is small, $p \sim (\beta(\ln 2)/2)^{1/2}$.

Calculations. Let $\alpha = 1$. Take $K = 4$. Then

$$\beta \sim H(2.2\Phi(-4)) \sim H(7 \cdot 10^{-5}) > .001, \quad p \sim .02, \quad 2p\alpha < .04.$$

(By way of illustration, the term $s = 2$ for β here is $2\Phi(-12) \times 20 \sim 8 \times 10^{-32}$.) Let $\alpha = .04$. Take $K = 5$. Then

$$\beta \sim (.04)^{-1}H(2.2\Phi(-5)) > 25H(7 \cdot 10^{-7}) < 4 \cdot 10^{-4}, \quad p \sim .012, \quad 2p\alpha < .001.$$

Let $\alpha = .001$. Take $K = 6$. Then

$$\beta \sim (.001)^{-1}H(2.2\Phi(-6)) \sim 10^3 H(2.5 \cdot 10^{-9}) < 7 \cdot 10^{-5},$$
$$p \sim .005, \quad 2p\alpha < .00001.$$

Let $\alpha = .00001$. By (9.12), $G(\alpha) < .122$. Thus

$$\begin{aligned} G(1) &\leq 4 + G(.04) \leq 4 + 5(.04)^{1/2} + G(.001) \\ &\leq 4 + 1 + 6(.001)^{1/2} + G(.00001) \\ &\leq 4 + 1 + .19 + .13 = 5.32. \end{aligned} \tag{9.13}$$

The controlling factor in the calculation was the choice of $K = 4$ when $\alpha = 1$. One may choose K smaller but at the cost of increasing p and hence the later terms. For example, if $K = 2$, $\beta \sim H(2\Phi(-2)(1.1)) + 2\Phi(-2)(1.1) \sim .34$ and $p \sim .325$. On the next step 65% of the ε_i still remain to be determined. While optimizing these calculations would certainly reduce the value 5.32 it is doubtful that, say, $G(1) \leq 3$ could be obtained without essentially new techniques.

We close by emphasizing the asymptotic nature of Theorem 1 and, indeed, all our results. The "simple" probabilistic method given at the end of the first section gives a bound of $(2n \ln(2n))^{1/2}$. This is clearly larger than $5.32n^{1/2}$ for sufficiently large n but it is in fact smaller than $5.32n^{1/2}$ for all n up to more than one half million!

REFERENCES

1. J. Beck and T. Fiala, *Integer-making theorems*, Discrete Appl. Math. **3** (1981), 1–8.
2. J. Beck and J. Spencer, *Integral approximation sequences*, Math. Programming **30** (1984), 88–98.
3. D. Kleitman, *On a combinatorial conjecture of Erdös*, J. Combin. Theory **1** (1966), 209–214.
4. J. Olson and J. Spencer, *Balancing families of sets*, J. Combin. Theory Ser. A **25** (1978), 29–37.
5. W. Rudin, *Some theorems on Fourier coefficients*, Proc. Amer. Math. Soc. **10** (1959), 855–859.
6. J. Spencer, *Sequences with small discrepency relative to n events*, Compositio Math. **47** (1982), 365–392.

DEPARTMENT OF MATHEMATICS, STATE UNIVERSITY OF NEW YORK, STONY BROOK, NEW YORK 11794