THE CONJECTURES OF RUDICH, TARDOS, AND KUSNER

BY CLIFFORD D. SMYTH

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Michael E. Saks

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2001

ABSTRACT OF THE DISSERTATION

The Conjectures of Rudich, Tardos, and Kusner

by Clifford D. Smyth

Dissertation Director: Michael E. Saks

We prove a dual version of the celebrated inequality of David Reimer (a.k.a. the van den Berg-Kesten conjecture). We use the dual inequality to prove Rudich's conjecture, a conjecture in combinatorial probability made by Steven Rudich in 1984, motivated by his investigation of reductions among cryptographic primitives. We also use the dual inequality to prove a conjecture of Gabor Tardos made in 1989, proving an upper bound on the approximate decision tree complexity of a boolean function that is quadratic in the function's approximate certificate complexity.

We also make progress on a 1983 conjecture of Robert Kusner in combinatorial geometry, improving the upper bound on the size of a one-distance set in certain metric spaces from an exponential in the dimension of the space to a polynomial in the dimension.

Acknowledgements

I'd like to thank: my family and friends for the love and support that made it all possible; my advisor Mike Saks for his generous support and advice; Jeff Kahn for his wonderful courses communicating all those great problems; David Galvin for the many mathematical discussions; Charlie Suffel for extending me the greatest help when I needed it the most; and Roger Pinkham whose example inspired me to become a mathematician.

The work in this thesis was supported in part by a research assistantship paid for by the NSF under contract CCR-9988526.

Dedication

To my parents.

Table of Contents

Chapter 1

A Brief Overview

For full statements of the problems with which this thesis is concerned, see the introductory sections of the individual chapters. Here we just give a brief overview of the thesis as a whole. We prove a dual version of the celebrated correlation inequality of David Reimer (a.k.a. the van den Berg-Kesten conjecture) in Chapter 2. We use the dual inequality to prove Rudich's conjecture in Chapter 3. This 1984 conjecture in combinatorial probability was motivated by Rudich's investigation of reductions among cryptographic primitives. We also use the dual inequality to prove a 1989 conjecture of Gabor Tardos in Chapter 4. Namely, we prove an upper bound on the approximate decision tree complexity of a boolean function that is quadratic in the function's approximate certificate complexity.

In Chapter 5, we make progress on a 1983 conjecture of Robert Kusner in combinatorial geometry. He conjectured that the maximum size of family of points in $\mathbb{R}^n$ such that each lies at distance 1 from the others in the L^p norm ($1 < p < \infty$) should be $n+1$ just as in the Euclidean case $p = 2$. The best known general upper bound was 2^n which we improve to $O(n^{(p+1)/(p-1)})$.

Chapter 2

Reimer's Inequality and The Dual Inequality

2.1 Introduction and Notation

In this chapter we introduce two new correlation inequalities, the dual inequality and strong dual inequality, Theorems 2.2 and 2.3 respectively. These will be later applied in Chapters 3 and 4, to solve two outstanding problems in computer science, Rudich's conjecture and Tardos' conjecture. For convenience we collect in this section most of the terminology that we use in Chapters 2-4.

Let n be a positive integer. We define $[n] = \{1, \ldots, n\}$, as is customary.

Let $V = \{v_1, v_2, \ldots, v_n\}$ be a set of Boolean ($\{0,1\}$-valued) variables. We call the set of variables v_i and their negations $\overline{v}_i$ the literals of V. A term t over V is a conjunction of literals of V. For example, $t_1 = v_1 v_2 v_3$, $t_2 = \overline{v}_3 v_4$, $t_3 = v_2 v_7 \overline{v}_9 \overline{v}_{10}$ are all terms. We do not allow a term to contain both a variable v_i and its negation $\overline{v}_i$.

Let $X = \{0,1\}^n$. We identify each point $x \in X$ with an assignment of values to the variables in V. Namely, v_i is set to x_i and $\overline{v}_i$ is set to $1 - x_i$. An assignment x satisfies the term t, written $x \models t$, if all of t's literals are set to 1 by x. We'll sometimes speak of terms over $X = \{0,1\}^n$ without explicitly introducing the canonical set of variables $V = \{v_1, \ldots, v_n\}$.

We say two terms s and t are dependent, $s \sim t$, if they have at least one variable in common (where we ignore negations). Otherwise they are independent, $s \not\sim t$. If there is an assignment x satisfying both s and t, we say s and t are compatible, $s \sim^c t$. Otherwise they are incompatible, $s \sim^{inc} t$. Clearly if $s \not\sim t$ then $s \sim^c t$.

For example, suppose $t_1 = v_1 v_2 v_3$, $t_2 = \overline{v}_3 v_4$, $t_3 = v_2 v_7 \overline{v}_9 \overline{v}_{10}$, as before. We have $t_1 \sim t_2$, $t_1 \sim t_3$, and $t_1 \not\sim t_3$. We also have $t_1 \sim^{inc} t_2$, $t_1 \sim^c t_3$, and $t_2 \sim^c t_3$.

Let X be a product of finitely many finite sets, $X = \prod_{i\in[n]} X_i$. We call such a set X, a finite product. Let $x \in X$ and $S \subseteq [n]$. The cylinder $C(x, S)$ is the set

$$C(x, S) = \{y \in X \ : \ \forall i \in S \ \ y_i = x_i\}.$$

If $C = C(x, S)$ is a cylinder on X we define $\mathrm{fix}(C) = S$ to be the set of fixed coordinates of C. We say two cylinders C and D are independent, $C \not\sim D$, if $\mathrm{fix}(C) \cap \mathrm{fix}(D) = \emptyset$. Otherwise they are dependent, $C \sim D$. We say C and D are compatible, $C \sim^c D$, if $C \cap D \neq \emptyset$. Otherwise they are incompatible, $C \sim^{inc} D$. Note that if $C \not\sim D$ then $C \sim^c D$.

Suppose $X = \{0, 1\}^n$. If t is a term we define the cylinder of assignments satisfying t,

$$C_t = \{x \in X : x \models t\}.$$

Every cylinder C of X correspond to a term t such that $C = C_t$ and vice-versa. We have $C_s \not\sim C_t$ iff $s \not\sim t$ and $C_s \sim^c C_t$ iff $s \sim^c t$.

Suppose $\Omega = \prod_{i\in[n]} \Omega_i$ is a product of finitely many finite probability spaces, $\Omega_i = (X_i, 2^{X_i}, \mu_i)$, $i \in [n]$. Then, $\Omega = (X, 2^X, \mu)$ where $X = \prod_{i\in[n]} X_i$, $\mu = \prod_{i\in[n]} \mu_i$. We call any such probability space Ω a finite product space. We call any such measure $\mu = \prod_{i\in[n]} \mu_i$ on a finite product $X = \prod_{i\in[n]} X_i$ a product measure.

Note that if μ is a product measure on $X = \prod_{i\in[n]} X_i$ and C and D are two cylinders of X, then $C \not\sim D$ implies $\mu(C \cap D) = \mu(C)\mu(D)$.

If μ is a probability measure on a set X and k is a positive integer, let $\mu^{(k)} = \prod_{i=1}^k \mu$ denote the product measure on $X^k = \prod_{i=1}^k X$.

2.2 Reimer's Inequality and the Dual

We introduce the following motivating example. Let G be a finite graph with edge set $E = \{e_1, \ldots, e_n\}$. Let $p_1, \ldots, p_n \in [0, 1]$. We define "percolation" on G to be the following random process. Independently assign the edge e_i the value 1 with probability p_i and the value 0 with probability $\overline{p}_i$. If $e_i = 1$ we say the edge is "open" in G, otherwise it is "closed". We say a path in G is open if all its edges are open. If v, v' and w, w'

are vertices in G, let $v \longleftrightarrow w$ be the event that there is an open path from v to w in G and let $(v \longleftrightarrow w)\Box(v' \longleftrightarrow w')$ be the event that there are edge-disjoint open paths from v to w and from v' to w' in G. We have

$$\Pr((v \longleftrightarrow w)\Box(v' \longleftrightarrow w')) \leq \Pr(v \longleftrightarrow w)\Pr(v' \longleftrightarrow w').$$

This follows from the following theorem

Theorem 2.1 Reimer's Inequality

Suppose $X = \{0,1\}^n$ and μ is a product measure on X. For all sets of terms S and T over X we have,

$$\mu\left(\bigcup_{s\in S, t\in T, s\not\subset t} C_s \cap C_t\right) \leq \mu^{(2)}\left(\bigcup_{s\in S, t\in T} C_s \times C_t\right). \tag{2.1}$$

To prove the inequality of the example above, we view each edge e_i as a variable. Every path P in G corresponds to a term s over E, namely, the conjunction of the e_i's in P. The desired inequality is then (2.1) where S is the set of all terms representing paths from v to w in G and T is the set of all terms representing paths from v' to w' in G.

We'll say a term t is positive if (like the terms of the previous example) it contains no negated variables. Motivated by problems in percolation, J. van den Berg and H. Kesten proved Theorem 2.1 for sets of positive terms [5]. This became known as the BK inequality. The same authors conjectured that the restriction to positive terms was uneccessary and that Theorem 2.1 held in general. This became known as the van den Berg–Kesten Conjecture or BK Conjecture and attracted a great deal of interest over the years. It was finally proved by D. Reimer [26] using his Butterfly Lemma, Lemma 2.4 below. The BK inequality and BK conjecture are not usually phrased in the language of terms, but in an equivalent formulation. We'll postpone a full discussion of this background information until Section 2.5.

We state Reimer's Butterfly Lemma in Section 2.3 and use it to give proofs of the following theorems in Section 2.4.

Theorem 2.2 The Dual Inequality

Suppose $X = \{0,1\}^n$ and μ is a product measure on X. For all sets of terms S and T

over X we have,

$$\mu^{(2)}\left(\bigcup_{s\in S,t\in T,s\not\sim t} C_s \times C_t\right) \leq \mu\left(\bigcup_{s\in S,t\in T} C_s \cap C_t\right). \tag{2.2}$$

Theorem 2.3 The Strong Dual Inequality

Suppose $X = \{0,1\}^n$ and μ is a product measure on X. For all sets of terms S and T over X we have,

$$\mu^{(2)}\left(\bigcup_{s\in S,t\in T,s\sim^c t} C_s \times C_t\right) \leq \mu\left(\bigcup_{s\in S,t\in T} C_s \cap C_t\right). \tag{2.3}$$

Since $s \not\sim t$ implies $s \sim^c t$ we have

$$\bigcup_{s\in S,t\in T,s\not\sim t} C_s \times C_t \subset \bigcup_{s\in S,t\in T,s\sim^c t} C_s \times C_t$$

so that Theorem 2.3 implies Theorem 2.2.

We give examples of these inequalities in terms of percolation. Let G' be the graph resulting from G after percolation and G'' the graph resulting from G after another independent run of percolation. Let $(v \longleftrightarrow w) \times^{ind} (v' \longleftrightarrow w')$ be the event that there are edge disjoint open paths from v to w in G' and from v' to w' in G''. Theorem 2.2 with the sets of terms S and T as before gives

$$\Pr^{(2)}((v \longleftrightarrow w) \times^{ind} (v' \longleftrightarrow w')) \leq \Pr(v \longleftrightarrow w \text{ and } v' \longleftrightarrow w')$$

Theorem 2.3 gives

$$\Pr(v \longleftrightarrow w)\Pr(v' \longleftrightarrow w') \leq \Pr(v \longleftrightarrow w \text{ and } v' \longleftrightarrow w')$$

This second statement is a trivial consequence of Kleitman's inequality on the positive correlation of monotone events [19]. When S and T are sets of arbitrary (not necessarily positive) terms Theorems 2.2 and 2.3 become non-trivial inequalities.

Although the inequalities of Theorem 2.2 and 2.3 are perhaps not very natural they find application in the resolution of combinatorial problems that had remained open for some time, namely Rudich's conjecture and Tardos' conjecture, which we address in Chapters 3 and 4 respectively.

We discuss generalizations of Theorems 2.1, 2.2 and 2.3 to arbitrary finite product spaces in Section 2.5.

2.3 Reimer's Butterfly Lemma

Let $X = \{0,1\}^n$. For $x, y \in \{0,1\}^n$, we define the cylinder spanned by x and y, $[x,y]$, to be the intersection of all the cylinders of X containing x and y. If $x \in X$ let $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n)$ be the complement of x. Let C be a cylinder. If $x \in C$, the complement of x relative to C, $\overline{x}^C$, is the unique point $y \in C$ such that $[x,y] = C$.

Let C be a cylinder, $a, b \in C$. The butterfly $B = B_{a,b,C}$ of C is the ordered triple (a,b,C). Let

$$\begin{array}{lllllll} \mathrm{Body}(B) & = & a & \quad \mathrm{Tip}(B) & = & b \\ \mathrm{Red}(B) & = & [a,b] & \mathrm{Yellow}(B) & = & [a,\overline{b}^C] \end{array}$$

If $\mathcal{B}$ is a family of butterflies of C, let

$$\mathrm{Red}(\mathcal{B}) = \bigcup_{B \in \mathcal{B}} \mathrm{Red}(B) \qquad \mathrm{Yellow}(\mathcal{B}) = \bigcup_{B \in \mathcal{B}} \mathrm{Yellow}(B).$$

If no two butterflies of $\mathcal{B}$ have the same tip, we say $\mathcal{B}$ has distinct tips.

Lemma 2.4 (Butterfly Lemma, [26]) *Suppose $X = \{0,1\}^n$ and C is a cylinder of X. If $\mathcal{B}$ is a family of butterflies of C with distinct tips then,*

$$|\mathcal{B}| \leq |\mathit{Red}(\mathcal{B}) \cap \mathit{Yellow}(\mathcal{B})|.$$

∎

2.4 Proof of the Dual Inequality

Proof of Theorems 2.2 and 2.3. Let $X = \{0,1\}^n$ and let μ be a product measure on X. I.e. fix $p_1, \ldots, p_n \in [0,1]$ and let $\mu(x) = (\prod_{x_i=1} p_i)(\prod_{x_i=0} \overline{p}_i)$ for all $x \in X$. Let S and T be two sets of terms over X. Let

$$\mathcal{X} = \bigcup_{s \in S, t \in T, s \not\sim t} C_s \times C_t, \quad \mathcal{Y} = \bigcup_{s \in S, t \in T} C_s \cap C_t \times \{0,1\}^n.$$

We need to show

$$\mu^{(2)}(\mathcal{X}) \leq \mu^{(2)}(\mathcal{Y}).$$

Let $z \in X, R \subset [n], C = C(z,R)$. Let

$$p_C := (\prod_{i \in R, z_i = 1} p_i^2)(\prod_{i \in R, z_i = 0} \overline{p}_i^2)(\prod_{i \in [n] \setminus R} p_i \overline{p}_i).$$

Note $\mu^{(2)}(x,y) = \mu(x)\mu(y) = p_C$ iff $[x,y] = C$. Thus

$$\mu^{(2)}(\mathcal{X}) = \sum_C |\mathcal{X}_C| p_C, \quad \mu^{(2)}(\mathcal{Y}) = \sum_C |\mathcal{Y}_C| p_C$$

where $\mathcal{X}_C = \{(x,y) \in \mathcal{X} \ : \ [x,y] = C\}$, $\mathcal{Y}_C = \{(x,y) \in \mathcal{X} \ : \ [x,y] = C\}$ and the sums range over all cylinders C of X. We'll show $|\mathcal{X}_C| \leq |\mathcal{Y}_C|$ to complete the theorem.

For each $(x,y) \in \mathcal{X}_C$, we define the butterfly $B(x,y)$ of C as follows. Pick $s \in S$ and $t \in T$ such that $x \in C_s$, $y \in C_t$, and $s \not\sim t$. Choose some $a = a(x,y) \in C_s \cap C_t \cap C$. This is possible because a trivial "Helly property" holds for cylinders: if $\{C_i\}$ is a family of pairwise non-disjoint cylinders then $\bigcap C_i \neq \emptyset$. In our case, $x \in C_s \cap C$, $y \in C_t \cap C$, and $C_s \cap C_t \neq \emptyset$ because $s \not\sim t$. Thus $C_s \cap C_t \cap C \neq \emptyset$. We let $B = B(x,y) = B_{a,x,C}$. Observe that $\mathrm{Red}(B) = [a,x] \subset C_s \cap C$ and $\mathrm{Yellow}(B) = [a,y] \subset C_t \cap C$. Let $\mathcal{B} = \{B(x,y) \ : \ (x,y) \in \mathcal{X}_C\}$. Clearly $\mathrm{Red}(\mathcal{B}) \subset \bigcup_{s \in S} C_s \cap C$ and $\mathrm{Yellow}(\mathcal{B}) \subset \bigcup_{t \in T} C_t \cap C$. So

$$\mathrm{Red}(\mathcal{B}) \cap \mathrm{Yellow}(\mathcal{B}) \subset \bigcup_{s \in S, t \in T} C_s \cap C_t \cap C.$$

Note that each point $x \in \mathrm{Red}(\mathcal{B}) \cap \mathrm{Yellow}(\mathcal{B})$ corresponds to a unique point $(x, \overline{x}^C) \in \mathcal{Y}_C$. Since $\mathcal{B}$ has distinct tips, Lemma 2.4 implies

$$|\mathcal{X}_C| = |\mathcal{B}| \leq |\mathrm{Red}(\mathcal{B}) \cap \mathrm{Yellow}(\mathcal{B})| \leq |\mathcal{Y}_C|.$$

Note that in the course of the proof we only used the implication $s \not\sim t \Rightarrow C_s \cap C_t \neq \emptyset$. Thus it would have sufficed if $s \sim^c t$. Thus the above proof also gives Theorem 2.3.

∎

2.5 Generalizations

Let n be a positive integer. Let $X = \prod_{i \in [n]} X_i$ be a product of finite sets. Let μ be a product measure on X.

Let $A, B \subset X$. We define

$$\begin{aligned}
A \cap^{ind} B &= \{x \ : \ \exists \text{ cylinders } C, D \ \ x \in C \subset A, x \in D \subset B, C \not\sim D\} \\
A \times^{ind} B &= \{(x,y) \ : \ \exists \text{ cylinders } C, D \ \ x \in C \subset A, y \in D \subset B, C \not\sim D\} \\
A \times^{c} B &= \{(x,y) \ : \ \exists \text{ cylinders } C, D \ \ x \in C \subset A, y \in D \subset B, C \sim^c D\}
\end{aligned}$$

We call $A \cap^{ind} B$ the independent intersection of A and B. Similarly $A \times^{ind} B$ is the independent product of A and B and $A \times^{c} B$, the compatible product of A and B. The usual notation for $A \cap^{ind} B$ is $A \square B$ but we introduce the notations $A \cap^{ind} B$ to highlight the similarity between this set and $A \times^{ind} B$. Note $A \cap^{ind} B \subset A \cap B$ and $A \times^{ind} B \subset A \times^{c} B \subset A \times B$.

Theorem 2.5 Reimer's Inequality

Let $\Omega = (X, 2^X, \mu)$ be a finite product space. For all $A, B \subseteq X$ we have,

$$\mu(A \cap^{ind} B) \leq \mu^{(2)}(A \times B). \tag{2.4}$$

Equivalently, for all families of cylinders $\mathcal{F}$ and $\mathcal{G}$ of X we have,

$$\mu(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}, C \not\sim D} C \cap D) \leq \mu^{(2)}(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}} C \times D). \tag{2.5}$$

Theorem 2.6 The Dual Inequality

Let $\Omega = (X, 2^X, \mu)$ be a finite product space. For all $A, B \subseteq X$ we have,

$$\mu^{(2)}(A \times^{ind} B) \leq \mu(A \cap B). \tag{2.6}$$

Equivalently, for all families of cylinders $\mathcal{F}$ and $\mathcal{G}$ of X we have,

$$\mu^{(2)}(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}, C \not\sim D} C \times D) \leq \mu(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}} C \cap D). \tag{2.7}$$

Theorem 2.7 The Strong Dual Inequality

Let $\Omega = (X, 2^X, \mu)$ be a finite product space. For all $A, B \subseteq X$ we have,

$$\mu^{(2)}(A \times^{c} B) \leq \mu(A \cap B). \tag{2.8}$$

Equivalently, for all families of cylinders $\mathcal{F}$ and $\mathcal{G}$ of X we have,

$$\mu^{(2)}(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}, C \sim^{c} D} C \times D) \leq \mu(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}} C \cap D). \tag{2.9}$$

We'll show (2.4) and (2.5) are equivalent. The equivalences of (2.6) and (2.7) and of (2.8) and (2.9) can be shown similarly. Suppose we have (2.5). Let $A, B \subseteq X$. If $\mathcal{F}, \mathcal{G}$ are the families of cylinders contained in A, B, respectively, then

$$\mu(A \cap^{ind} B) = \mu(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}, C \not\sim D} C \cap D) \leq \mu^{(2)}(\bigcup_{C \in \mathcal{F}, D \in \mathcal{G}} C \times D) = \mu^{(2)}(A \times B).$$

Suppose we have (2.4). Let $\mathcal{F}, \mathcal{G}$ be families of cylinders of X. Let $A = \bigcup_{C\in\mathcal{F}} C$, $B = \bigcup_{D\in\mathcal{G}} D$, and let $\mathcal{F}'$ and $\mathcal{G}'$ be the families of cylinders contained in A and B, respectively. Note $\mathcal{F} \subseteq \mathcal{F}'$ and $\mathcal{G} \subseteq \mathcal{G}'$, thus

$$\mu(\bigcup_{C\in\mathcal{F},D\in\mathcal{G},C\not\sim D} C\cap D) \leq \mu(\bigcup_{C\in\mathcal{F}',D\in\mathcal{G}',C\not\sim D} C\cap D) = \mu(A\cap^{ind} B)$$

$$\leq \mu^{(2)}(A\times B) = \mu^{(2)}(\bigcup_{C\in\mathcal{F},D\in\mathcal{G}} C\times D).$$

A brief background on Theorems 2.5, 2.6, and 2.7 is now in order. We define the Boolean case of these theorems to be the case where $X = \{0,1\}^n$ and the uniform case to be the case that μ is the uniform measure. Note Theorems 2.1, 2.2, and 2.3 are just the Boolean cases of Theorems 2.5, 2.6, and 2.7, namely (2.5), (2.7), and (2.9), respectively.

We identify $\{0,1\}^n$ and the set of subsets of $[n]$ in the usual way. We say $A \subseteq \{0,1\}^n$ is a filter if $x \in A$ and $y \supset x$ imply $y \in A$. As mentioned earlier J. van den Berg and H. Kesten proved the Boolean case of (2.4) when A and B are filters [5]. This became known as the BK inequality. The same authors conjectured that the Boolean case of (2.4) should hold in general. This conjecture and its generalization to arbitrary finite product spaces, namely, Theorem 2.5, became known as the van den Berg-Kesten conjecture or BK conjecture.

J. van den Berg and U. Fiebig [4] showed that in order to prove Theorem 2.5, it suffices to prove the uniform Boolean case. The van den Berg-Kesten conjecture attracted a great deal of interest over the years and was finally resolved by D. Reimer [26] who proved the uniform Boolean case using his Butterfly Lemma, Lemma 2.4.

Proof of Theorems 2.6 and 2.7. We'll show that the Boolean cases (Theorems 2.2 and 2.3) suffice. Let $X = \prod_{i\in[n]} X_i$ be a finite product space with product measure $\mu = \prod_{i\in[n]} \mu_i$. Let $X_i = \{x_{i,1}, \ldots, x_{i,m_i}\}$ and $p_{i,j} = \mu_i(x_{i,j})$. Without loss of generality, we may assume $p_{i,j} > 0$.

Let $Y = \prod_{i\in[n]} Y_i$ where $Y_i = \{0,1\}^{m_i}$. Let $Y_{i,j}$ be the subcube of all $y \in Y_i$ with initial prefix $0^{j-1}1$. Let $f : Y \to X$ be defined by

$$f^{-1}(x) = \prod_{i\in[n]} Y_{i,x_i}.$$

Note that since each $Y_{i,j}$ is a cylinder, f^{-1} maps cylinders of X to cylinders of Y and that this map preserves independence and compatibility of cylinders.

We pick a product measure $\nu = \prod_{i\in[n]} \nu_i$ on $Y = \prod_{i\in[n]} Y_i$ so that $\mu = \nu \circ f^{-1}$. Let $V_i = \{v_{i,1}, \ldots, v_{i,m_i}\}$ be the set of variables associated with $Y_i = \{0,1\}^{m_i}$. Let $q_{i,j} = \nu_i(v_{i,j})$. We must have $p_{i,j} = \mu_i(x_j) = \nu_i(Y_{i,j}) = (1-q_{i,1})\cdots(1-q_{i,j-1})q_{i,j}$. Thus if we take $q_{i,j} = p_{i,j}/(1-(p_{i,1}+\cdots+p_{i,j-1}))$, $\mu = \nu \circ f^{-1}$.

Applying the Boolean case of Theorem 2.6 to Y, ν we get

$$\mu(A\cap B) = \nu(f^{-1}(A\cap B)) = \nu(f^{-1}(A)\cap f^{-1}(B)) \geq \nu^{(2)}(f^{-1}(A) \times^{ind} f^{-1}(B)).$$

It remains to show that

$$f^{-1}(A) \times^{ind} f^{-1}(B) \supset (f\times f)^{-1}(A \times^{ind} B),$$

for then

$$\mu(A\cap B) \geq \nu^{(2)}(f^{-1}(A) \times^{ind} f^{-1}(B)) \geq \nu^{(2)}((f\times f)^{-1}(A \times^{ind} B)) = \mu^{(2)}(A \times^{ind} B)$$

as desired.

Suppose $(y,y') \in (f\times f)^{-1}(A \times^{ind} B)$. Then $(x,x') \in A \times^{ind} B$ where $x = f(y)$ and $x' = f(y')$. Pick independent cylinders C, C' of X so that $x \in C \subset A$, $x' \in C' \subset B$. Thus we get independent cylinders D, D', $y \in D = f^{-1}(C) \subset f^{-1}(A)$, $y' \in D' = f^{-1}(C') \subset f^{-1}(B)$ and $(y,y') \in f^{-1}(A) \times^{ind} f^{-1}(B)$ as desired. Similarly $f^{-1}(A) \times^{c} f^{-1}(B) \supset (f\times f)^{-1}(A \times^{c} B)$ and Theorem 2.7 also reduces to the Boolean case.

∎

Chapter 3

Rudich's Conjecture

3.1 Introduction

Let T be a set of terms over the set $V = \{v_1, v_2, \ldots, v_n\}$ of Boolean variables. Let $(T, \sim)$ be the graph with vertex set T and an edge between terms s and t if and only if $s \sim t$. We call $(T, \sim)$ the dependency graph of T. If t is in T, we call $N_T(t) = \{s \in T \ : \ s \sim t\}$, the neighborhood of t in T. We just write $N(t)$ if T is understood.

We pick an assignment $x \in \{0,1\}^n$ uniformly at random. For $S \subseteq T$ let $U(S)$ be the probability that exactly one term of S is satisfied by x and $\Pr(S)$, the probability that at least one term is satisfied. We may think of the following conjecture as saying that one cannot have $U(T)$ close to 1 without some significant dependence among the terms of T.

Conjecture 3.1 Rudich's Conjecture

There exist $\epsilon, \delta > 0$ independent of n such that

$$\textit{if } U(T) \geq 1 - \epsilon \textit{ then there exists } t \in T \textit{ such that } \Pr(N(t)) \geq \delta.$$

Consider the following (known) examples. Let $T = \{v_1 v_2 \cdots v_n, \ldots, \overline{v}_1 \overline{v}_2 \cdots \overline{v}_n\}$ be the set of all 2^n terms of length n over $V = \{v_1, \ldots, v_n\}$. Then $U(T) = 1$. However for any term $t \in T$, $N(t) = T$. Thus $\Pr(N(t)) = 1$.

Let $W = \bigcup_{i=1}^{2^n} \{w_1^i, \ldots, w_n^i\}$ be a set of $n2^n$ variables. Let $S = \{s_1, \ldots, s_m\}$ be a set of $m = 2^n$ independent terms of length n over W, e.g. $s_k = w_1^k \cdots w_n^k$ for $1 \leq k \leq m$. Then $N(s) = \{s\}$ for each $s \in S$ and $\Pr(N(s)) = 2^{-n}$ is not bounded away from 0. However, we have $\lim_{n\to\infty} U(S) = 1/e$. Thus Rudich's conjecture cannot hold with $\epsilon \geq 1 - 1/e$.

We believe this last example is extremal.

Conjecture 3.2

Let T_n be a set of terms on $\{v_1, \ldots, v_n\}$ for $n \geq 1$. If $\max_{t \in T_n} \Pr(N_{T_n}(t)) = o(1)$ as $n \to \infty$, then $U(T_n) < 1/e + o(1)$ as $n \to \infty$.

We now restate Rudich's conjecture. Let T be a set of terms. Unless otherwise stated, indices s and t run over T. For $t \in T$ we set

$$C'_t = C_t \setminus \bigcup_{s \neq t} C_s.$$

Set $f(\epsilon) = (1-\epsilon) - 4\epsilon/(1-\epsilon)$. We have $U(T) = \Pr(\bigcup_t C'_t)$. If $t \in T$ then $\Pr(N(t)) \geq U(N(t)) \geq \Pr(\bigcup_{s \sim t} C'_s)$. Thus the following implies Rudich's conjecture.

Theorem 3.3

Let T be a set of terms and $\epsilon \in [0, 1)$.

$$\textit{If } \Pr\left(\bigcup_t C'_t\right) \geq 1 - \epsilon, \textit{ then } \max_t \Pr\left(\bigcup_{s \sim t} C'_s\right) \geq f(\epsilon).$$

Since $f(1/7) = 4/21 > 1/7$, this gives Rudich's conjecture with $\epsilon = \delta = 1/7$. Since $f(\epsilon) > 0$ for $0 \leq \epsilon < 3 - 2\sqrt{2} \approx .171$, the conjecture holds for ϵ in this range. Also $f(\epsilon) \uparrow 1$ as $\epsilon \downarrow 0$ and $f(0) = 1$.

Let S and T be two sets of terms over $V = \{v_1, \ldots, v_n\}$. We have

Theorem 3.4 The Bipartite Theorem

For every $\epsilon > 0$ there is a $\delta > 0$ independent of n such that if $\Pr(S) \geq \epsilon$ and $\Pr(T) \geq \epsilon$ then

$$\Pr\left(\bigcup_{s \in S, t \in T} C_s \cap C_t\right) \geq \delta \textit{ or } \Pr^{(2)}\left(\bigcup_{s \in S, t \in T, s \sim t} C_s \times C_t\right) \geq \delta.$$

In the next section we use the dual inequality, Theorem 2.2, to give short proofs of Theorems 3.3 and 3.4. We note that the strong dual inequality, Theorem 2.3, implies stronger versions of these theorems, Theorems 3.6 and 3.7. Also since the dual inequality and the strong dual inequality hold for arbitrary product measures on $X = \{0, 1\}^n$, so do Theorems 3.3, 3.4, 3.6, and 3.7.

S. Rudich's conjecture arose in his study of reductions among cryptographic primitives beginning in his 1984 thesis [27]. The bipartite conjecture also proves useful in complexity theory [17]. We state some of these complexity theoretic consequences in Section 3.3. We also apply the bipartite theorem in Chapter 4 to answer a 1989 question of Tardos on complexity measures on Boolean functions.

3.2 Proof of Rudich's Conjecture

Before proving Rudich's Conjecture we require the following lemma:

Lemma 3.5

For any set of terms T, there is a partition $T = T_1 \dot\cup T_2$ such that

$$\Pr^{(2)}\left(\bigcup_{s\in T_1, t\in T_2, s\not\sim t} C'_s \times C'_t\right) \geq (1/4)\Pr^{(2)}\left(\bigcup_{s\in T, t\in T, s\not\sim t} C'_s \times C'_t\right).$$

Proof of Lemma. Consider the complete graph G with vertex set T. We define the weight of an edge to be $w(\{s,t\}) = 2\Pr(C'_s)\Pr(C'_t)$ if $s \not\sim t$ and 0 otherwise. Note that

$$W = \sum_e w(e) = \sum_{s\in T, t\in T, s\not\sim t} \Pr(C'_s)\Pr(C'_t) = \Pr^{(2)}\left(\bigcup_{s\in T, t\in T, s\not\sim t} C'_s \times C'_t\right).$$

As is standard knowledge, we can greedily construct a bipartition $T_1 \dot\cup T_2 = T$ of the vertices of any edge-weighted graph G so that

$$W' = \sum_{s\in T_1, t\in T_2} w(\{s,t\}) \geq (1/2)W.$$

Thus

$$\Pr^{(2)}\left(\bigcup_{s\in T_1, t\in T_2, s\not\sim t} C'_s \times C'_t\right) = \sum_{s\in T_1, t\in T_2, s\not\sim t} \Pr(C'_s)\Pr(C'_t) = (1/2)W' \geq (1/4)W.$$

∎

Proof of Theorem 3.3. Let $T = T_1 \dot\cup T_2$ be a partition of T as in the lemma. We have $\Pr(\bigcup_t C'_t) \geq 1 - \epsilon$. Thus

$$\epsilon \geq 1 - \Pr\left(\bigcup_t C'_t\right) \geq \Pr\left(\bigcup_{s\neq t} C_s \cap C_t\right) \geq \Pr\left(\bigcup_{s\in T_1, t\in T_2} C_s \cap C_t\right).$$

By the dual inequality we have

$$\Pr\left(\bigcup_{s\in T_1, t\in T_2} C_s \cap C_t\right) \geq \Pr^{(2)}\left(\bigcup_{s\in T_1, t\in T_2, s\not\sim t} C_s \times C_t\right).$$

By the lemma, we have

$$\Pr^{(2)}\left(\bigcup_{s\in T_1,t\in T_2,s\not\sim t} C_s\times C_t\right)\geq \Pr^{(2)}\left(\bigcup_{s\in T_1,t\in T_2,s\not\sim t} C'_s\times C'_t\right)$$
$$\geq (1/4)\Pr^{(2)}\left(\bigcup_{s\in T,t\in T,s\not\sim t} C'_s\times C'_t\right).$$

Putting these inequalities together we have

$$4\epsilon\geq \Pr^{(2)}\left(\bigcup_{s\not\sim t} C'_s\times C'_t\right)=\sum_{s\not\sim t}\Pr^{(2)}(C'_s\times C'_t)=\sum_{s\not\sim t}\Pr(C'_s)\Pr(C'_t),$$

or

$$4\epsilon\geq \sum_t\left[\Pr(C'_t)\sum_{s\not\sim t}\Pr(C'_s)\right].$$

Dividing both sides of this inequality by $\lambda=\sum_t\Pr(C'_t)\geq 1-\epsilon$ we get

$$4\epsilon/(1-\epsilon)\geq\sum_t\left[w_t\sum_{s\not\sim t}\Pr(C'_s)\right]$$

where $w_t=\Pr(C'_t)/\lambda$ is a set of non-negative weights with $\sum_t w_t=1$. Thus

$$4\epsilon/(1-\epsilon)\geq \min_t\sum_{s\not\sim t}\Pr(C'_s).$$

Therefore

$$\max_t\sum_{s\sim t}\Pr(C'_s)=\lambda-\min_t\sum_{s\not\sim t}\Pr(C'_s)\geq(1-\epsilon)-(4\epsilon/(1-\epsilon))=f(\epsilon).$$

∎

Proof of Theorem 3.4. Let $\delta=\epsilon^2/2$ and suppose for a contradiction that the conclusion of the theorem fails. Thus, by the dual inequality,

$$\Pr^{(2)}\left(\bigcup_{s\in S,t\in T,s\not\sim t} C_s\times C_t\right)\leq \Pr\left(\bigcup_{s\in S,t\in T} C_s\cap C_t\right)<\delta.$$

By hypothesis we have

$$\Pr^{(2)}\left(\bigcup_{s\in S,t\in T,s\sim t} C_s\times C_t\right)<\delta,$$

so we have

$$\epsilon^2\leq \Pr\left(\bigcup_{s\in S} C_s\right)\Pr\left(\bigcup_{t\in T} C_t\right)=\Pr^{(2)}\left(\bigcup_{s\in S,t\in T} C_s\times C_t\right)$$
$$\leq \Pr^{(2)}\left(\bigcup_{s\in S,t\in T,s\not\sim t} C_s\times C_t\right)+\Pr^{(2)}\left(\bigcup_{s\in S,t\in T,s\sim t} C_s\times C_t\right)<2\delta,$$

a contradiction.

∎

If we use the strong dual inequality, Theorem 2.3, instead of the dual inequality and make other minor modifications to the proofs of Theorems 3.3 and 3.4 we get the following stronger statements:

Theorem 3.6

Let T be a set of terms and $\epsilon \in [0,1)$.

$$\textit{If } \Pr\left(\bigcup_t C'_t\right) \geq 1-\epsilon, \textit{ then } \max_t \Pr\left(C'_t \cup \bigcup_{s \sim^{inc} t} C'_s\right) \geq f(\epsilon).$$

∎

Theorem 3.7 The Strong Bipartite Theorem

For every $\epsilon > 0$ there is a $\delta > 0$ independent of n such that if $\Pr(S) \geq \epsilon$ and $\Pr(T) \geq \epsilon$ then

$$\Pr\left(\bigcup_{s \in S, t \in T} C_s \cap C_t\right) \geq \delta \textit{ or } \Pr^{(2)}\left(\bigcup_{s \in S, t \in T, s \sim^{inc} t} C_s \times C_t\right) \geq \delta.$$

∎

If T is a set of terms and $t \in T$ let $N'_T(t) = \{t\} \cup \{s \in T : s \sim^{inc} t\}$. Theorem 3.6 implies

Theorem 3.8

There exist $\epsilon, \delta > 0$ independent of n such that

$$\textit{if } U(T) \geq 1-\epsilon \textit{ then there exists } t \in T \textit{ such that } \Pr(N'(t)) \geq \delta.$$

∎

where ϵ and δ can be taken as in Theorem 3.3.

Since the dual and strong dual inequalities, Theorems 2.2 and 2.3 hold for general product measures so do Rudich's conjecture and the bipartite theorem:

Corollary 3.9 *Theorems 3.3, 3.4, 3.6, 3.7, 3.8 hold for general product measures.*

∎

3.3 Some Applications

Assuming Rudich's conjecture, Conjecture 3.1, S. Rudich and R. Impagliazzo proved Theorems 3.10 and 3.11 below [17]. Assuming the bipartite theorem, Theorem 3.4, the same authors can prove Theorem 3.12 below [17].

Let $\mathcal{M}$ be the class of polynomial time oracle Turing machines M such that for each oracle R and for every n, M^R maps input strings in $\{0,1\}^n$ to output strings in $\{0,1\}^n$. Roughly speaking, Theorem 3.10 states if one could construct $M \in \mathcal{M}$ such that for a random oracle R, M^R is a one-way permutation, then one would have proved $P \neq NP$. Of course, a random oracle R is a good one-way function with probability 1. Theorem 3.10 indicates that a so called "black-box" reduction M of a one-way function R to a one-way permutation, M^R, may be very hard to construct.

Theorem 3.10 *If $P = NP$ then for every $M \in \mathcal{M}$ there is an $A \in \mathcal{M}$ such that with probability 1 over random oracles R one of the following occurs:*

1. *M^R is not a one to one function,*
2. *there are infinitely many input lengths n so that*

$$\Pr_{x \in \{0,1\}^n}(A^R(M^R(x)) = x) \geq 2/3.$$

∎

Theorem 3.11 *There is an oracle relative to which one-way functions exist but one-way permutations do not.*

∎

We say a language L is in $i.o.AvgP$ if there is a polynomial time oracle Turing machine M such that for each input x, M never incorrectly accepts x (but may refuse to accept or reject x) and there exists a constant fraction $\epsilon > 0$ such that for infinitely many n, M gives the correct answer on an ϵ-fraction of the inputs of length n.

Theorem 3.12 *If $P = NP$ then relative to a random oracle R*

$$NP^R \cap coNP^R \subseteq i.o.AvgP^R.$$

∎

It is hoped [17] that (with the assumption $P = NP$) it can be proved that relative to a random oracle R

$$NP^R \cap coNP^R \subseteq AvgP^R.$$

Chapter 4

Tardos' Conjecture

4.1 Introduction

Let $X = \{0,1\}^n$. We call a function $f : \{0,1\}^n \to \{0,1\}$, a Boolean function. We consider deterministic algorithms A that take $x \in X$ as input and give a number $A(x) \in \{0,1\}$ as output. We say A computes f if $A(x) = f(x)$ for all $x \in X$. We will be interested in the query complexity of these algorithms. Informally, the query complexity of an algorithm $A : X \to \{0,1\}$ is the maximum (taken over all inputs $x \in X$) of the number of bits x_i that A reads from x, (before halting to give $A(x)$ as output). The query complexity of f will then be the minimum query complexity of an algorithm A that computes f.

To formalize the notion of a deterministic algorithm, we introduce decision trees. A decision tree A is a rooted, labeled, directed, binary tree. Every internal node v of A has a label $l(v) \in [n]$, and two children, a left child and a right child. Each leaf v has a label $l(v) \in \{0,1\}$. One evaluates A on $x \in X$ as follows. Let $v = r$ be the root of the tree. Repeat the following process (if necessary) until v is a leaf of the tree. Let $j = l(v)$. Let v_0 and v_1 be the left and right child of v respectively. The tree A queries x_j and branches to $v = v_1$ if $x_j = 1$ and to $v = v_0$ if $x_j = 0$. When a leaf v (uniquely determined by x) is reached, the tree outputs $A(x) = l(v)$. We identify deterministic algorithms and decision trees. The query complexity of a decision tree is defined to be its depth, the maximum length of a path from the root to a leaf.

We say a decision tree A computes (ϵ-approximates) f if $A(x) = f(x)$ for all $x \in X$ (for $\geq (1-\epsilon)$ fraction of the inputs $x \in X$). We define the decision tree complexity of

f, $D(f)$, and the ϵ-approximate version, $D_\epsilon(f)$:

$$\begin{aligned} D(f) &= \text{the minimum depth of a decision tree } A \text{ for } f \\ D_\epsilon(f) &= \text{the minimum depth of a } \epsilon\text{-approximate decision tree } A \text{ for } f \end{aligned}$$

Since the only computational resource we are considering is query complexity, we may assume we have $X_1 = f^{-1}(1)$ and $X_0 = f^{-1}(0)$ at our disposal when constructing a decision tree A computing f. For example, the tree A that queries each bit of x and then answers 1 if $x \in X_1$ and 0 if $x \in X_0$ shows $D(f) \leq n$.

Suppose we are at some intermediate stage in the evaluation of a decision tree A computing $f(x)$ on input x. Suppose the algorithm A has learned $x_i = 1$ for $i \in Y_1$ and $x_i = 0$ for $i \in Y_0$. We define the (current) query term q to be the minimal term consistent with the queries made to x so far.

$$q = \bigwedge_{i \in Y_1} x_i \wedge \bigwedge_{i \in Y_0} \overline{x}_i$$

We call $Q = C_q$ the (current) query cylinder. As A continues querying variables, q grows in length and Q shrinks in size. As soon as

$$Q = C_q \subset f^{-1}(j),$$

for some $j \in \{0, 1\}$, A may halt and answer j. In this event, we say q (or Q) is a "certificate" or "j-certificate" for x, because q certifies $f(x) = j$, i.e. $f(y) = j$ for any $y \models q$. Thus a decision tree A computing $f(x)$ can be viewed as returning a certificate for x, namely the query term q reached by the end of the evaluation.

The size of a certificate is the number of variables queried, i.e. the length of q ($|\text{fix}(Q)|$). If $x \in X$ let $C(x, f) =$ the minimum size of a certificate for x. We define the certificate complexity of f to be

$$C(f) = \max_{x \in X} C(x, f).$$

If T computes f it must make at least $C(x, f)$ queries on x, so

$$D(f) \geq C(f).$$

Let S, T be sets of terms on X. Let $\mathcal{S} = \bigcup_{s\in S} C_s$, $\mathcal{T} = \bigcup_{t\in T} C_t$. We'll say (S,T) is a complete set of certificates for f, if $\mathcal{S} = f^{-1}(0)$ and $\mathcal{T} = f^{-1}(0)$. We define $\text{amb}(S,T) = X \setminus (\mathcal{S}\triangle\mathcal{T})$. We'll say (S,T) is an δ-approximate set of certificates for f if $x \in \text{amb}(S,T)$ or $x \in \mathcal{S} \cap f^{-1}(0)$ or $x \in \mathcal{T} \cap f^{-1}(1)$ for at most $< \delta 2^n$ inputs $x \in X$. The size of (S,T), $m = m(S,T)$ is the maximum length of a term in $S \cup T$. Thus the certificate complexity of a Boolean function f is

$$C(f) = \text{ the minimum size of a complete set of certificates for } f.$$

We define the δ-approximate certificate complexity of f to be

$$C_\delta(f) = \text{ the minimum size of an } \delta\text{-approximate set of certificates for } f.$$

Note $C_0(f) = C(f)$.

We may think of $C(f)$ as the non-deterministic query complexity of f. In fact, $C(f)$ is the minimum query complexity of a non-deterministic algorithm N that computes f. N first correctly guesses x and then makes $C(f)$ queries to get a certificate for x. In principle, no computation path of an algorithm A computing f need make more than $C(f)$ queries, however, deterministically discovering these queries when x is unknown is another matter.

Example 4.1

Consider the following boolean function $f : \{0,1\}^{n^2} \to \{0,1\}$ *where* $\{0,1\}^{n^2} = \{0,1\}^{n\times n}$ *is viewed as the set of* $n \times n$ $\{0,1\}$*-valued matrices.*

$$f(M) = \begin{cases} 1, & (\exists i \in [n])(\forall j \in [n]) M_{i,j} = 1 \\ 0, & \textit{otherwise} \end{cases}$$

Then

$$C(f) = n, \ \textit{but } D(f) = n^2.$$

Indeed, $\{\bigwedge_{j\in[n]} M_{i,j} : i \in [n]\}$ is a complete set of 1-certificates and $\{\bigwedge_{i\in[n]} \overline{M}_{i,g(i)} : g : [n] \to [n]\}$ is a complete set of 0-certificates, all of minimum size. Thus $C(f) = n$. On the other hand, imagine a deterministic algorithm A trying to determine $f(M)$ on input M and an adversary answering A's queries and constructing the matrix M itself

as he goes along. The adversary plays the same strategy on every row $i \in [n]$. He answers 1 for every variable $M_{i,j}$ queried in that row until there is just one M_{i,j_0} left unqueried which he then answers 0. This forces A to query all $M_{i,j}$. Hence $D(f) = n^2$.

This quadratic gap between $D(f)$ and $C(f)$ is worst possible, a result independently discovered in [7, 15, 30].

Theorem 4.2

If $f : \{0,1\}^n \to \{0,1\}$ *then*

$$D(f) \leq C^2(f).$$

We give a proof following [30]. Let S and T be two sets of terms on $X = \{0,1\}^n$. We say an algorithm A on X plays the compatibility game on S and T if it correctly outputs $x \models S$ or $x \models T$ for all $x \notin \text{amb}(S,T)$. For $x \in \text{amb}(S,T)$, A may give arbitrary output.

Theorem 4.3

If S *and* T *are two sets of terms with* $amb(S,T) = \emptyset$ *and* $m = m(S,T)$ *then there is a deterministic algorithm* A *that plays the compatibility game on* S *and* T *with query complexity* $\leq m^2$.

Proof of Theorem 4.2. Let (S,T) be a complete set of certificates for f with $m = m(S,T) = C(f)$ and apply Theorem 4.3.

∎

Proof of Theorem 4.3. The algorithm A will proceed in rounds. Let q_i be the query term at the end of round i. Let $Q_i = C_{q_i}$. The algorithm starts with $i = 0$, $q_0 =$ the empty term, $Q_0 = \{0,1\}^n$, $S_0 = S$, and $T_0 = T$.

At the beginning of round i, the algorithm checks two halting conditions.

1. If $y \models S_{i-1}$ for all $y \in Q_{i-1}$ halt and output "$x \models S$".

2. If $y \models T_{i-1}$ for all $y \in Q_{i-1}$ halt and output "$x \models T$".

If neither halting condition is met, A picks the first term t_i of T_{i-1} and queries all its variables. Update the query term to q_i. Let $S_i = \{s \setminus q_i : s \in S_{i-1}, s \sim^c q_i\}$,

$T_i = \{t \setminus q_i : t \in T_{i-1}, t \sim^c q_i\}$, where $s \setminus t$ is the term obtained from s by removing all variables s has in common with t. Thus for all i, S_i and T_i are sets of terms of Q_i, and for all $y \in Q_i$ we have $y \models S_i$ ($y \models T_i$) iff $y \models S$ ($y \models T$).

At most m variables are queried each round, namely the variables of t_i. Since $\mathrm{amb}(S_{i-1}, T_{i-1}) = Q_{i-1} \cap \mathrm{amb}(S, T) = \emptyset$ we have $s \sim^{inc} t_i$ for all $s \in S_{i-1}$. Thus every $s \in S_{i-1}$ will either be thrown out, or have its size reduced by at least one before being placed in S_i. Thus A can proceed for at most m rounds before S_i is either empty or contains the empty term. If S_i contains the empty term then $y \models S_i$ for all $y \in Q_i$. If S_i is empty, $\mathrm{amb}(S_i, T_i) = \emptyset$ implies that we have $y \models T_i$ for all $y \in Q_i$. In either case, A will halt at the beginning of the $i+1$st round before any new queries are made. Thus the query complexity of A is $\leq m^2$.

∎

Let S and T be two sets of terms on $X = \{0,1\}^n$ as before. We say an algorithm A on X plays the ϵ-compatibility game on S and T, if it answers "$x \models S$" or "$x \models T$" incorrectly for no more than $< \epsilon 2^n$ inputs $x \notin \mathrm{amb}(S, T)$. As before, A is allowed arbitrary output for $x \in \mathrm{amb}(S, T)$. Tardos made the following conjecture about query complexity while investigating separation of complexity classes relative to random oracles [30].

Conjecture 4.4

There exists a positive constant c so that the following holds. For all $\epsilon > 0$ there is a $\delta > 0$ such that there is a deterministic algorithm A to play the ϵ-compatibility game on S and T with query complexity $\leq m^c$ where $|amb(S,T)| < \delta 2^n$ and $m = m(S,T)$.

We show this conjecture holds for $c = 2$. Let $\delta = \delta(S,T) = |\{x : x \models S, x \models T\}|/2^n$.

Theorem 4.5

Let S and T be sets of terms over $X = \{0,1\}^n$, let $m = m(S,T)$, $\delta = \delta(S,T)$. If $\delta < (\epsilon/3)^3$, there is a deterministic algorithm A that plays the ϵ-compatibility game on S and T with query complexity $\leq Km^2$, where $K = (1 - \epsilon/3)/((\epsilon/3)^3 - \delta)$.

This result easily implies

Theorem 4.6

Let f be a Boolean function, and let $\epsilon > 0$. Let ϵ_0 be the unique real root of $p(x) = \epsilon$ where $p(x) = x + (x/3)^3$. For all $0 \leq \delta < \epsilon - \epsilon_0$, we have

$$D_\epsilon(f) \leq K(C_\delta(f))^2$$

where $K = (1 - \gamma/3)/((\gamma/3)^3 - \delta)$ with $\gamma = \epsilon - \delta$).

Proof of Theorem. We'll first show that if $0 \leq \delta < (\gamma/3)^3$ then $D_{\gamma+\delta}(f) \leq K(C_\delta(f))^2$, where $K = (1 - \gamma/3)/((\gamma/3)^3 - \delta)$. Let (S, T) be a δ-approximate set of certificates for f with $m = m(S, T) = C_\delta(f)$. We have $\delta(S, T) \leq |\text{amb}(S, T)|/2^n < \delta < (\gamma/3)^3$. By Theorem 4.5 there is a deterministic algorithm A that plays the γ-compatibility game on S and T. So A correctly outputs $f(x)$ on all but a $< (\gamma + \delta)$-fraction of inputs. So $D_{\gamma+\delta}(f) \leq K(C_\delta(f))^2$ where $K = (1 - \gamma/3)/((\gamma/3)^3 - \delta)$.

Suppose $0 \leq \delta < \epsilon - \epsilon_0$. Let $\gamma = \epsilon - \delta$. Since $\gamma > \epsilon_0$ and $p(x) = x + (x/3)^3$ is strictly increasing, we have $\epsilon = p(\epsilon_0) < p(\gamma) = \gamma + (\gamma/3)^3$ or $\delta = \epsilon - \gamma < (\gamma/3)^3$. Thus $D_\epsilon(f) = D_{\gamma+\delta}(f) \leq K(C_\delta(f))^2$, where $K = (1 - \gamma/3)/((\gamma/3)^3 - \delta)$, as claimed.

∎

Proof of Theorem 4.5. T The ideas used are closely related to methods Rudich and Impagliazzo use to prove Theorem 3.12 [17]. Let $S = \{s_1, \ldots, s_M\}$ and $T = \{t_1, \ldots, t_N\}$ be two sets of terms on $X = \{0, 1\}^n$. Let $m = m(S, T)$, let $\delta = \delta(S, T)$. Let $\epsilon > 0$ be given. Pick $a, b, c > 0$ so that $a + (1 - a)(b + c) \leq \epsilon$. Let $K = b(1 - a)/(a^2 b - \delta)$. If $\delta < a^2 b$, we'll describe an algorithm A that plays the ϵ-compatibility game on S and T with query complexity $\leq (1/c)Km^2$. Taking $a = b = c = \epsilon/3$ gives the statement of the theorem.

We'll modify the algorithm of Theorem 4.3. Let q_i, $Q_i = C_{q_i}$, S_i, T_i, $q_0 =$ the empty term, $Q_0 = \{0, 1\}^n$, $S_0 = S$, and $T_0 = T$ all as in Theorem 4.3. We let S_i, T_i inherit the indexing on terms from S and T. At the beginning of round $i \geq 1$, A checks the following six halting conditions in sequence. Pr denotes the uniform probability on X. We write $q = q_{i-1}$, $Q = Q_{i-1}$, $S = S_{i-1}$, $T = T_{i-1}$.

1. If $\Pr(y \models S | y \in Q) < a$, halt and output "$x \models T$".

2. If $\Pr(y \models S | y \in Q) > 1 - a$, halt and output "$x \models S$".

3. If $\Pr(y \models T | y \in Q) < a$ halt and output "$x \models S$".

4. If $\Pr(y \models T | y \in Q) > 1 - a$ halt and output "$x \models T$".

5. If $\Pr(y \models S \text{ and } y \models T | y \in Q) > (1/b)\delta$ halt and output "failure".

6. If $i \geq (1/c)Km$ halt and output "failure".

If $x \models S$ let $\min_x(S) = s_k$ be the first shortest term of S that x satisfies, that is $x \models s_k \in S$ and for all $s_l \in S$ such that $x \models s_l$, $|s_l| \geq |s_k|$, and if $l < k$, $|s_l| > |s_k|$. For $s \in S$ let

$$C_s = \{x \in Q : x \models s\},\ C''_s = \{x \in C_s : \min_x(S) = s\}.$$

Note that the C''_s are disjoint for $s \in S$ and $\bigcup_{s \in S} C''_s = \bigcup_{s \in S} C_s$. We define C_t and C''_t for $t \in T$ similarly.

If none of the halting conditions were met the following analysis applies. In this paragraph we let Pr be the uniform distribution on Q. The dual inequality, Theorem 2.2, implies

$$\Pr^{(2)}(\textstyle\bigcup_{s \in S, t \in T, s \not\sim t} C''_s \times C''_t) \leq \Pr^{(2)}(\bigcup_{s \in S, t \in T, s \not\sim t} C_s \times C_t)$$

$$\leq \Pr(\textstyle\bigcup_{s \in S, t \in T} C_s \cap C_t) \leq (1/b)\delta$$

Also we have

$$\Pr^{(2)}(\textstyle\bigcup_{s \in S, t \in T, s \not\sim t} C''_s \times C''_t) + \Pr^{(2)}(\bigcup_{s \in S, t \in T, s \sim t} C''_s \times C''_t) = \Pr^{(2)}(\bigcup_{s \in S, t \in T} C''_s \times C''_t)$$

$$= \Pr^{(2)}(\textstyle\bigcup_{s \in S, t \in T} C_s \times C_t) = \Pr(\bigcup_{s \in S} C_s) \Pr(\bigcup_{t \in T} C_t) \geq a^2$$

Putting these last two inequalities together we get

$$\Pr^{(2)}(\textstyle\bigcup_{s \in S, t \in T, s \sim t} C''_s \times C''_t) \geq a^2 - (1/b)\delta$$

or

$$\textstyle\sum_{t \in T} \Pr(C''_t) \left[\Pr(\bigcup_{s \sim t} C''_s)\right] \geq a^2 - (1/b)\delta.$$

Since $\lambda = \sum_{t\in T} Pr(C''_t) = \Pr(\bigcup_{t\in T} C_t) < 1 - a$ we can divide this last inequality by $\lambda < 1 - a$ to get

$$\sum_{t\in T} w_t \Pr(\bigcup_{s\sim t} C''_s) > 1/K$$

where $w_t = \Pr(C'_t)/\lambda \geq 0$, $\sum_{t\in T} w_t = 1$, and $K = b(1-a)/(a^2b - \delta)$. Thus there is a term $t \in T$ such that $\Pr(\bigcup_{s\sim t} C''_s|Q) \geq 1/K$, where Pr once again denotes the uniform distribution on $X = \{0,1\}^n$. Let t be the first such term in T. The algorithm queries all of the variables of t and updates q, S, T, accordingly.

By halting condition 6, A performs $\leq (1/c)Km$ rounds each requiring $\leq m$ queries, so the the query complexity of A is $\leq (1/c)Km^2$ as claimed. We now must show $\Pr(E) < \epsilon$ where

$$E = \{x \in X \ : \ x \notin \text{amb}(S,T), A \text{ makes an error on } x\}$$

For $1 \leq j \leq 6$ let

$$H_j = \{x \in X \ : \ A \text{ halts on } x \text{ due to condition } j\}, \ \ E_j = E \cap H_j$$

Let $p_j = \Pr(H_j)$, $q_j = \Pr(E_j)$. We will show $q_j < ap_j$ for $j = 1, \ldots 4$, $p_5 \leq b$, and $p_6 \leq c$ so that

$$\Pr(E) = \sum_j q_j < a(p_1 + \cdots + p_4) + p_5 + p_6 \leq a + (1-a)(b+c) \leq \epsilon.$$

We think of creating the decision tree for A in the following way. At first the tree is completely "unexpanded". It consists of just the root node, the cylinder $Q = \{0,1\}^n$ with its associated sets of terms S and T. Pick a leaf Q of the current tree. If A halts at cylinder Q due to the halting condition j, label Q by j. It becomes a permanent leaf of the tree. Otherwise, "expand" the cylinder Q by performing one round of the algorithm on it. Using the terms S and T associated with Q, A picks a unique term t of T and queries its variables. Thus for each possible outcome of these queries, we create a branch from Q to the cylinder Q' with its updated sets of terms S' and T'. The process continues until the tree is fully expanded and every leaf is a halting cylinder. We associate a round number with each node of the tree. The round number of a cylinder Q is the number r of expansions or rounds necessary to arrive at Q. Thus the

round number of the root cylinder is 0, those reached by the end of round 1, have round number 1, etc.

For $1 \leq j \leq 4$, we can write $q_j = \Pr(E_j) = \sum_C \Pr(E_j|C)\Pr(C)$ where the sum extends over all leaf cylinders C labeled j. But by halting condition j we have $\Pr(E_j|C) < a$, thus $q_j \leq ap_j$ as claimed. Suppose $p_5 \geq b$ then $\Pr(y \models S \text{ and } y \models T) \geq \sum_C \Pr(y \models S \text{ and } y \models T|C)\Pr(C)$ where the sum is over leaf cylinders labeled 5. But by halting condition 5, the conditional probabilities in the last sum are $> (1/b)\delta$ and so we get $\Pr(y \models S \text{ and } y \models T) > \delta$ a contradiction.

For a partially expanded tree we define the round and length functions r' and l' on X. For $x \in X$, let $Q(x)$ be the unique leaf Q of the current tree such that $x \in Q$. Let $S(x)$ and $T(x)$ be the sets of terms corresponding to $Q(x)$. Let $r'(x)$ be the round number of $Q(x)$. If $x \notin \text{amb}(S,T)$ and $x \models S$, let $l'(x)$ be the length of the shortest term of $S(x)$ that x satisfies. Otherwise let $l'(x) = 0$. Let r'' and l'' denote the depth and length functions of the tree after expanding one cylinder Q.

We have $l''(x) \geq l'(x)$ for all x. Suppose $x \in Q$. Let S, T be the sets of terms corresponding to Q. Suppose $x \notin \text{amb}(S,T)$. Let $s = \min_S(x) \in S$. Since the algorithm does not terminate on Q, a uniquely defined term $t \in T$ is selected, and its variables are queried. The cylinder Q is partitioned into subcylinders according to the answers to these queries. Suppose x is in a particular subcylinder Q' with associated terms S', T'. Since the term $s(x) = \min_S(x)$ is compatible with the new query term q', the term $s\ t$ appears in S'. Thus if $s \sim t$, we have $l''(x) \geq l'(x)+1$. But by the choice of t the fraction of $x \in Q$ that satisfy this condition is large, $\Pr(\cup_{s \in S: s \sim t} C''_s|Q) \geq (1/K)$. Thus $\mathbb{E}[l'(x)|Q] \geq \mathbb{E}[l''(x)|Q] + (1/K)$. Taking expected values we get

$$\mathbb{E}l'(x) = \sum_{C \neq Q} \mathbb{E}[l'(x)|C]\Pr(C) + \mathbb{E}[l'(x)|Q]\Pr(Q)$$

$$\geq \sum_{C \neq Q} \mathbb{E}[l''(x)|C]\Pr(C) + \mathbb{E}[l''(x)|Q]\Pr(Q) + (1/K)\Pr(Q)$$

$$= \mathbb{E}l''(x) \geq (1/K)\Pr(Q)$$

where the sums extend over all leaf cylinders C of the tree before expanding Q. Thus

we have

$$\mathbb{E}l'(x) - \mathbb{E}l''(x) \geq (1/K)\Pr(Q).$$

Since $r''(x) = r'(x) + 1$ for $x \in Q$ and $r''(x) = r'(x)$ otherwise, we have $\mathbb{E}r''(x) - \mathbb{E}r'(x) = \Pr(Q)$. Thus

$$\mathbb{E}l''(x) - \mathbb{E}l'(x) \geq (1/K)(\mathbb{E}r''(x) - \mathbb{E}r'(x)).$$

Let l_0 be the function l' on the unexpanded tree, $Q = \{0,1\}^n$. Let l be the function l' on the fully expanded tree. Define r_0 and r, similarly. Summing the previous inequality over all expansions we get

$$\mathbb{E}l_0(x) - \mathbb{E}l(x) \geq (1/K)(\mathbb{E}r(x) - \mathbb{E}r_0(x)).$$

Initially, $r_0 = 0$ and $l_0 \leq m$ so we have

$$Km \geq \mathbb{E}r(x).$$

Thus by Markov's inequality $p_6 = \Pr(H_6) \leq \Pr(\{x : r(x) \geq (1/c)\mathbb{E}d(x)\}) \leq c$.

∎

One can show (by appropriately choosing a, b, c in the last proof) that for all $\delta \leq f(\epsilon)$ there is a $K > 0$ such that $D_\epsilon(f) \leq K(C_\delta(f))^2$, where $f(\epsilon) = (4/27)\epsilon^3 + O(\epsilon^4)$. It is natural to suppose

Conjecture 4.7 *For all $\epsilon > 0$ and for all $\delta < \epsilon$ there exists a $K > 0$ so that*

$$D_\epsilon(f) \leq K(C_\delta(f))^2.$$

Chapter 5

Progress on Kusner's Conjecture

5.1 Introduction

Let (X, d) be a metric space. We say a subset S of X is a one-distance set or an equilateral set if every pair of points of S determine the same distance, i.e. if there is a number $r \geq 0$ such that $d(s,t) = r$ for all $s, t \in S$ with $s \neq t$. We define

$$e(X) = \text{the maximum cardinality of an equilateral subset of } X,$$

if this maximum exists. For example, it is "obvious" that $e(X) = n + 1$ when X is ordinary Euclidean space $\mathbb{R}^n$. See Theorem 5.5 below for a proof.

In general, let $M^n = (\mathbb{R}^n, \|\cdot\|)$ be a Minkowski space, i.e. a real finite dimensional normed space [32]. We view M^n as a metric space with the metric induced by the norm, i.e. $d(x,y) = \|x - y\|$ for all $x, y \in M^n$. The results of Petty and Brass [25, 9] give

Theorem 5.1

There is a constant C such that

$$C\left(\frac{\log n}{\log\log n}\right)^{1/3} \leq e(M^n) \leq 2^n$$

for every Minkowski space M^n. We have $e(M^n) = 2^n$ if and only if the unit ball of the norm is a parallelotope.

∎

If $p \in [1, \infty]$ and $n \geq 1$, we define $L^p(n) = (\mathbb{R}^n, \|\cdot\|_p)$ where $\|\cdot\|_p$ is the usual L^p norm on $\mathbb{R}^n$, namely

$$\|x\|_p = \begin{cases} (\sum_{i\in[n]} |x_i|^p)^{1/p}, & p \in [1, \infty) \\ \max_{i\in[n]} |x_i|, & p = \infty \end{cases}$$

Consider the following examples of equilateral sets in $L^p(n)$. The set of unit vectors together with an appropriately chosen multiple $\lambda_p \mathbf{1}$ of the all ones vector $\mathbf{1}$ is an equilateral set of size $n+1$ in $L^p(n)$. The set of points $\{0,1\}^n$ is an equilateral set of size 2^n in $L^\infty(n)$. The set of standard basis vectors and their negatives is an equilateral set of size $2n$ in $L^1(n)$. Thus

$$e(L^p(n)) \geq \begin{cases} 2n, & p = 1 \\ n+1, & p \in (1,\infty) \\ 2^n, & p = \infty \end{cases}$$

Note by Theorem 5.1, $e(L^p(n)) \leq 2^n$, and so $e(L^\infty(n)) = 2^n$. It is tempting to conjecture that the other examples are also best possible.

Conjecture 5.2 (Kusner, [14])

For all $n \geq 1$,

$$e(L^1(n)) = 2n.$$

Conjecture 5.3 (Kusner, [14])

For all $p \in (1,\infty)$ and all $n \geq 1$,

$$e(L^p(n)) = n+1.$$

The cases $n = 1, 2$ of Conjecture 5.2 are easy to verify. The case $n = 3$ was handled in [3] and the case $n = 4$ in [21] and also independently by Nussbaum [24]. Otherwise all that is known is what Theorem 5.1 and the trivial lower bound give us, namely,

$$2n \leq e(L^1(n)) \leq 2^n - 1 \text{ for } n > 4.$$

It is easy to verify Conjecture 5.3 for $n = 1, 2$. We have $e(L^2(n)) = n+1$. See Theorem 5.5 below for a proof. Galvin [13] observed that this proof also gives $e(L^p(n)) \leq 1 + (p-1)n$ when p is a positive even integer. Swanepoel proved $e(L^4(n)) = n+1$ [28]. Otherwise all that is known is that when p is an even integer and $p \geq 6$,

$$n+1 \leq e(L^p(n)) \leq \min\{1 + (p-1)n, 2^n - 1\} \text{ for } n > 2,$$

and that when $p \in (1,\infty)$ and p is not an even integer,

$$n+1 \leq e(L^p(n)) \leq 2^n - 1 \text{ for } n > 2.$$

5.2 New Results

If x is a real number and $n \geq 0$ an integer, we define $(x)_n := \prod_{i=0}^{n-1}(x-i)$. We prove

Theorem 5.4

For $p \in (1,\infty)$ and $n \geq 1$,

$$e(L^p(n)) \leq C(p)n^{(p+1)/(p-1)},$$

where $C(p) = ((2^p\lceil p\rceil^p(1+\pi^2/2)^{\lceil p\rceil}(p)_{\lceil p\rceil-1})/\lceil p\rceil!)^{1/(p-1)}$.

Note that $C(p)/p \to 2+\pi^2$ as $p \to \infty$ and so for any $p_0 > 1$ there is a fixed constant C depending on p_0 such that $e(L^p(n)) \leq Cn^{(p+1)/(p-1)}$ for all $p \geq p_0$.

We will first prove

Theorem 5.5

For $n \geq 1$,

$$e(L^2(n)) = n+1.$$

Note that we do not give the shortest or most natural proof but rather one that we will extend to a proof of Theorem 5.4.

Before giving the proof of Theorem 5.5, we require the following standard independence criterion,

Lemma 5.6

Let V be the vector space of real-valued functions on a set X. Let $\{f_1,\ldots,f_m\} \subset V$. Suppose $\{a^1,\ldots,a^m\} \subset X$ and the matrix

$$\left[f_i(a^j)\right]_{i,j=1,\ldots,m}$$

is invertible. Then the functions f_i are linearly independent in V.

Proof of Lemma. Suppose $f = \sum_k \lambda_k f_k = 0$. Evaluating f at each a^i we get the matrix equation $\lambda M = 0$ where $M = \left[f_i(a^j)\right]_{i,j=1,\ldots,m}$ and $\lambda = [\lambda_1 \cdots \lambda_m]$. Since M is invertible, $\lambda = 0$. Thus the f_i are linearly independent.

∎

Proof of Theorem 5.5. Let $S = \{a^1, \ldots, a^m\}$ be an equilateral set in $L^2(n)$ of maximum size. We must show $|S| \leq n + 1$.

By scaling S, we may assume $\|a^i - a^j\|_2 = 1$, for $i \neq j$. For $1 \leq i \leq m$, we define $f_i : \mathbb{R}^n \to \mathbb{R}$,

$$f_i(x) = 1 - \|x - a^i\|_2^2$$

Thus

$$M = \left[f^i(a^j)\right]_{i,j=1,\ldots,m}$$

is the identity matrix. By Lemma 5.6, $f_1, \ldots, f_m$ are linearly independent.

Note that each f_i is a polynomial in $R = \mathbb{R}[x_1, \ldots x_n]$,

$$f_i(x) = 1 - \|x - a^i\|_2^2 = (1 - \sum_k (a_k^i)^2) + \sum_k 2a_k^i x_k - \sum_k x_k^2.$$

In fact, each f_i is in the subspace W of R spanned by $\{1, x_1, \ldots, x_n, \sum x_k^2\}$. Thus

$$|S| = |\{f_1, \ldots, f_m\}| \leq \dim(W) = n + 2.$$

Let $f_0 = 1$. We will show $f_0, \ldots, f_n$ are also linearly independent. Suppose that $f = \lambda_0 f_0 + \cdots + \lambda_m f_m = 0$. Evaluating f at a^i we find $\lambda_i = -\lambda_0$ for $1 \leq i \leq m$. Thus the coefficient of $\sum x_k^2$ in f is $-m\lambda_0$. On the other hand, since $f = 0$, we have $-m\lambda_0 = 0$ and so $\lambda_i = 0$ for all i. Thus

$$|S| + 1 = |\{f_0, \ldots, f_m\}| \leq \dim(W) = n + 2,$$

or $|S| \leq n + 1$.

∎

Note: The technique of using linearly independent polynomials in some space W to establish an upper bound on the size of a set in terms of $\dim(W)$ appears first in [20]. The technique of "augmenting" the set of polynomials to obtain improved bounds appears first in [6]. A standard reference for these "linear algebra" methods is [2].

Note that if $S = \{a^1, \ldots, a^m\}$ is an equilateral set in $L^p(n)$ where p is an even integer, we can define $f_i = 1 - \|x - a^i\|_p^p \in W$ where W is the space in $\mathbb{R}[x_1, \ldots, x_n]$ spanned by $\{1, \sum_p x_k^p\} \cup \bigcup_k \{x_k, x_k^2, \ldots, x_k^{p-1}\}$. The set $f_0, \ldots, f_m$ will be linearly independent, and so $|S| + 1 \leq \dim(W) = 2 + (p-1)n$. This is

Theorem 5.7 (Galvin, [13])

If p is an even integer

$$e(L^p(n)) \leq 1 + (p-1)n.$$

∎

Before proving Theorem 5.4 we require two lemmas. The first comes from the following matrix invertibility criterion. We say a matrix $M \in \mathbb{R}^{m\times m}$ is diagonally dominated if

$$|M_{ii}| > \sum_{j\neq i} |M_{ij}|, \quad \text{for } i = 1, \ldots, m.$$

It is a well known result that

Theorem 5.8

If $M \in R^{m\times m}$ diagonally dominated then M is invertible.

Proof of Theorem. Suppose M is not invertible, i.e. suppose we have $x \in \mathbb{R}^m$, $x \neq 0$ such that $Mx = 0$. Pick i such that $|x_i| = \max_j |x_j| > 0$. Since $\sum_j M_{ij}x_j = 0$, we have

$$|M_{ii}||x_i| = |M_{ii}x_i| = |\sum_{j\neq i} -M_{ij}x_j| \leq \sum_{j\neq i} |M_{ij}x_j| \leq \sum_{j\neq i} |M_{ij}||x_i|$$

or

$$|M_{ii}| \leq \sum_{j\neq i} |M_{ij}|,$$

a contradiction.

∎

See [16] for a discussion of the history of this result. Also [31] cites no fewer than 25 references in which this result was apparently independently discovered, the earliest, [22], dating to 1881. We only require the following easy consequence of this theorem. For $A \in \mathbb{R}^{m\times m}$, let

$$\|A\|_\infty = \max_{1\leq i,j\leq m} |A_{ij}|.$$

Let I denote the identity matrix.

Lemma 5.9

Let $A \in \mathbb{R}^{m\times m}$. If $\|A - I\|_\infty < 1/m$, then A is invertible.

Proof of Lemma. By hypothesis $A_{ii} > 1 - 1/m$ and $|A_{ij}| < 1/m$ for $i \neq j$. Thus A is diagonally dominated and hence invertible by Theorem 5.8.

∎

The second lemma comes from the theory of approximation of functions. Let f be a real-valued function with domain $S \subset \mathbb{R}$. Let $M \geq 0$ and $0 < \alpha \leq 1$. If

$$|f(x) - f(y)| \leq M|x - y|^\alpha, \quad x, y \in S,$$

we say $f \in \mathrm{Lip}_M(\alpha)$. If $I \subset \mathbb{R}$ is a compact interval we define the norm,

$$\|f\|_{\sup} = \max_{x\in I} |f(x)|,$$

for $f \in C^0(I)$. Let $f^{(k)}(x)$ denote the kth derivative of $f(x)$ if it exists. We then have the following theorem as stated in [23], Theorem 45, p.57.

Theorem 5.10 (Jackson, [18])

Let $k \geq 1$. Suppose $f \in C^k([-1, 1])$ and $f^{(k)} \in Lip_M(\alpha)$. For every $d > k$ there is a polynomial g of degree at most d such that

$$\|f - g\|_{\sup} \leq \frac{D(k, \alpha, d)c^{k+1}M}{d^{k+\alpha}}$$

where $D(k, \alpha, d) = d^{k+\alpha}/((d)_k(d - k)^\alpha)$ and $c = 1 + \pi^2/2$.

∎

As a corollary we obtain

Lemma 5.11

Let $p \in (1, \infty)$. Let $f(x) = |x|^p$ for $x \in [-1, 1]$. For every $d \geq \lceil p \rceil$ there is a polynomial g of degree at most d such that

$$\|f - g\|_{\sup} \leq B(p)/d^p$$

where $B(p) = (\lceil p \rceil^p (1 + \pi^2/2)^{\lceil p \rceil}(p)_{\lceil p \rceil - 1})/\lceil p \rceil!$

Proof of Lemma. Let $k = \lceil p \rceil - 1$, $\alpha = p - k \in (0, 1]$. Then $f^{(k)}(x) = \operatorname{sgn}^k(x)(p)_k|x|^\alpha \in \operatorname{Lip}_M(\alpha)$ where $M = (p)_k$. So by Theorem 5.10, if $d \geq k + 1 = \lceil p \rceil$, there is a polynomial g of degree at most d such that $\|f - g\|_{\sup} \leq D(k, \alpha, d)c^{k+1}M/d^p$. If k and α are fixed, then $D = ((1 - 1/d) \cdots (1 - (k - 1)/d)(1 - k/d)^\alpha)^{-1}$ increases as d decreases and so $\|f - g\|_{\sup} \leq B/d^p$ where $B = B(p) = D(k, \alpha, k + 1)c^{k+1}M$.

∎

Proof of Theorem 5.4. Fix $p \in (1, \infty)$ and let $f(x) = |x|^p$ for $x \in [-1, 1]$. Fix $d \geq \lceil p \rceil$. By Lemma 5.11 we can find a polynomial g of degree at most d satisfying $\|f - g\|_{\sup} \leq B(p)/d^p$.

Let $S = \{a^1, a^2, \ldots, a^m\} \subset L^p(n)$ be an equilateral set of maximum size, scaled so that $\|a^i - a^j\|_p = 1$ for $i \neq j$. Define the following functions from $\mathbb{R}^n$ to $\mathbb{R}$:

$$f_i(x) = 1 - \|x - a^i\|_p^p = 1 - \sum_{k=1}^{n} f(x_k - a_k^i),\ i = 1, \ldots, m,$$

$$g_i(x) = 1 - \sum_{k=1}^{n} g(x_k - a_k^i),\ \ i = 1, \ldots, m.$$

By construction of the f_i, $M = [f_i(a^j)] = I$. Let $M' = [g_i(a^j)]$. We have

$$\left|(M' - M)_{ij}\right| \leq \sum_k \left| f(a_k^j - a_k^i) - g(a_k^j - a_k^i) \right|$$

Since $|a_k^j - a_k^i| \leq \| a^j - a^i\|_p \leq 1$ we have

$$\left|(M - M')_{ij}\right| \leq n\|f - g\|_{\sup} \leq nB(p)/d^p.$$

Thus

$$\|M' - I\|_\infty \leq nB(p)/d^p.$$

Take $d = \max\{\lceil p \rceil, \lfloor (nmB(p))^{1/p} \rfloor + 1\}$ so that $\|M' - I\|_\infty \leq nB(p)/d^p < 1/m$. By Lemma 5.9, M' is invertible and hence by Lemma 5.6 , the polynomials g_i, $i = 1, \ldots, m$ are linearly independent. Since the g_i are elements of W, the subspace of $\mathbb{R}[x_1, \ldots, x_n]$ spanned by $\{1, \sum_k x_k^d\} \cup \bigcup_k \{x_k, \ldots, x_k^{d-1}\}$, we have $m \leq \dim(W) = 2 + n(d - 1)$.

We have $m \leq 2 + n(d - 1) \leq 2n(d - 1)$, for $n \geq 2$. Suppose $d = \lfloor (nmB_p)^{1/p} \rfloor + 1$. Then $m \leq 2n(nmB(p))^{1/p}$ and hence

$$m \leq (2^p B(p))^{1/(p-1)} n^{(p+1)/(p-1)}.$$

On the other hand if $d = \lceil p \rceil$ then $m \leq 2 + n(\lceil p \rceil - 1) \leq n\lceil p \rceil$, for $n \geq 2$. Since one can easily show $\lceil p \rceil \leq (B(p))^{1/(p-1)}$, we have the previous inequality in this case as well. Thus the theorem is established with $C(p) = (2^p B(p))^{1/(p-1)}$ as claimed.

∎

References

[1] N. Alon, J. Spencer, and P. Erdős, *The Probabilistic Method*, J. Wiley & Sons, New York, 1992.

[2] László Babai and Péter Frankl, *Linear algebra methods in combinatorics*, unpublished manuscript.

[3] H.J. Bandelt, V. Chepoi, and M. Laurent, *Embedding into rectilinear spaces,* Discrete and Computational Geometry, Volume 19 (1998), 595–604.

[4] J. van den Berg and U. Fiebig, *On a combinatorial conjecture concerning disjoint occurrences of events*, Ann. Probab. Volume 15 (1987), 354–374.

[5] J. van den Berg and H. Kesten, *Inequalities with applications to percolation and reliability*, J. Appl. Probab. Volume 22 (1985), 556–569.

[6] A. Blokhuis, *A new upper bound for the cardinality of 2-distance sets in Euclidean space*, Eindhoven Univ. Technology, Mem. 1981-04 (1981).

[7] Manuel Blum and Russel Impagliazzo, *Generic oracles and oracle classes*, Proceedings of the 28th IEEE Symposium on Foundations of Computer Science, pages 118–126. IEEE, New York, 1987.

[8] C. Borgs, J.T. Chayes, and D. Randall, *The van den Berg - Kesten - Reimer Inequality: a Review*, Microsoft Tech report MSR-TR-98-42, 1998. (see also `www.math.gatech.edu/~randall`)

[9] Peter Brass, *On equilateral simplices in normed spaces,* Contributions to Algebra and Geomtery, Volume 40 (1999), Number 2, 303–307.

[10] J. T. Chayes, A. Puha, and T. Sweet, *Independent and dependent percolation*, IAS/Park City Mathematics Series, Volume 6, AMS, Providence, 1998.

[11] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre, *Correlation inequalities on some partially ordered sets*, Commun. Math. Phys. Volume 22 (1971), 89–103.

[12] P. C. Fishburn and L. A. Shepp, *On the FKB conjecture for disjoint intersections*, Discrete Math. Volume 98 (1991), 105–122.

[13] David Galvin, *personal communication.*

[14] R. Guy ed., *Unsolved Problems: An Olla-Podrida of Open Problems, Often Oddly Posed,* American Mathematical Monthly, Volume 90 (1983), Number 3, 196–200.

[15] Juris Hartmanis and Lane A. Hemachandra, *One-way functions, robustness, and nonisomorphism of NP-complete sets*, Proceedings Structure in Complexity Theory Second Annual Conference, 160–174, Cornell University, Ithaca, NY, June 1987, IEEE.

[16] Roger A. Horn, Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[17] R. Impagliazzo and S. Rudich. *Personal communications.*

[18] Dunham Jackson, *The Theory of Aproximation*, American Mathematical Society, New York, 1930.

[19] Kleitman, Daniel J. *Families of non-disjoint subsets*, J. Combinatorial Theory 1 1966 153–155.

[20] Koornwinder, Thomas H., *A note on the absolute bound for systems of lines*, Nederl. Akad. Wetensch. Proc. Ser. A 79=Indag. Math. Volume 38 (1976), Number 2, 152–153.

[21] Jack Koolen, Monique Laurent, Alexander Schrijver, *Equilateral dimension of the rectilinear space,* Designs, Codes and Cryptography, Volume 21 (2000), Number 1-3, 149–164.

[22] L. Levy, *Sur la possibilité de l'équilibre électrique*, Comptes Rendus, Volume 93 (1881), Number 2, 706-708.

[23] G. Meinardus, *Approximation of Functions: Theory and Numerical Methods*, Springer-Verlag, New York, 1967.

[24] Roger Nussbaum, *personal communication.*

[25] C. M. Petty, *Equilateral sets in Minkowski spaces,* Proceedings of the American Mathematical Society Volume 29 (1971), Number 2, 369–374.

[26] D. Reimer, *Proof of the van den Berg–Kesten Conjecture*, Comb., Prob. & Comput., Volume 9 (2000), no. 1, 27–32.

[27] S. Rudich, *Limits on the provable consequences of one-way functions*, CS Ph.D Thesis, Berkeley (see also `www.cs.cmu.edu/~rudich`)

[28] K. J. Swanepoel, *personal communication.*

[29] M. Talagrand, *Some remarks on the Berg - Kesten inequality, Probab. in Banach Spaces, 9 (Sandjberg, 1993)*, Prog. Probab. Volume 35 (1994), Birkäuser, Boston, 293–297.

[30] Tardos, G. *Query complexity, or why is it difficult to separate* $\mathrm{NP}^A \cap \mathrm{co\text{-}NP}^A$ *from* P^A *by random oracles A?*, Combinatorica, Volume 9 (1989), Number 4, 385–392.

[31] Olga Taussky, *A recurring theorem on determinants*, American Mathematical Monthly, Volume 56 (1949), Number 10, 672–676.

[32] Anthony C. Thompson, *Minkowski Geometry*, Encyclopedia of Mathematics and Its Applications, Volume 63, Cambridge University Press, Cambridge, 1996.

Vita

Clifford D. Smyth

1988 Graduated from Saint Joseph's High School, Metuchen, New Jersey.

1988-93 Attended Stevens Institute of Techonology, Hoboken, New Jersey.

1992 B.S. Mathematics, Stevens Institute of Technology.

1992 Research Assistanship, Brookhaven National Labs, Brookhaven, New York.

1993 M.S. Mathematics, Stevens Institute of Technology.

1993-01 Graduate work in Mathematics, Rutgers, The State University of New Jersey, New Brunswick, New Jersey.

1993-95 GAANN Fellowship, Department of Mathematics.

1995-00 Teaching Assitantship, Department of Mathematics.

2000 Research Assistantship, Department of Mathematics.

2000 J. Kahn, M. Saks, and C. Smyth, A dual version of Reimer's inequality and a proof of Rudich's conjecture , Proceedings Fifteenth Annual IEEE Conference on Computational Complexity, IEEE Computer Society, Los Alamitos, CA, 2000, 98-103.

2001 PhD. Mathematics, Rutgers, The State University of New Jersey.