Variations of a combinatorial problem on finite sets

Bas Lemmens

Bas Lemmens is a Warwick Zeeman Lecturer in the Mathematics Department at the University of Warwick, England. He studied mathematics at the University of Amsterdam and received his Ph.D. degree from the Free University in Amsterdam in 2001. His mathematical interests include dynamical systems theory and geometry of Banach spaces.

“Try, try again”, is popular advice. It is good advice. The insect, the mouse, and the man follow it; but if one follows it with more success than the others it is because he varies his problem more intelligently. This strategy for solving mathematical problems was recommended to us by George Pólya [10, p. 209]. In this article we will take Pólya’s suggestion and discuss two variations of a combinatorial problem on finite sets. These variations were found by Koolen, Laurent, and Schrijver [9], and reveal a surprising connection with geometry. At present however, the problem is still unresolved. Perhaps you can find another variation that will lead to its solution.

The problem we are interested in is a generalization of Fisher’s inequality to multisets. We shall see that such a generalization is closely related to sets in \( \mathbb{R}^n \) in which the taxicab (or Manhattan) distance between any two distinct points is the same. On the other hand, it will become clear to us in the third section that the problem can also be reformulated in terms of pairwise touching simplices. In fact, we shall see that our problem is equivalent to determining the maximum number of translated copies of a regular \( n \)-dimensional simplex that can be placed in \( \mathbb{R}^n \) such that any two distinct ones touch but do not overlap. Fig. 1

In seinen Arbeiten zur statistischen Versuchsplanung entdeckte R.A. Fisher im Jahr 1940 die folgende überraschende Ungleichung: Sind \( F_1, \ldots, F_m \) verschiedene Teilmengen der Zahlenmenge \( \{1, \ldots, n\} \) mit der Eigenschaft, dass jede Menge \( F_j \) genau \( r \) Elemente und jeder Durchschnitt \( F_j \cap F_k (j \neq k) \) genau \( \lambda \) Elemente mit \( 0 \leq \lambda \leq r \) besitzt, so gilt \( m \leq n \). Im nachfolgenden Beitrag wird die Fishersche Ungleichung dahingehend verallgemeinert, dass in den zugrunde liegenden Mengen Elemente mehrfach auftreten dürfen. Mit Hilfe dieses Ergebnisses gewinnt der Autor Erkenntnisse zu Fragestellungen aus der kombinatorischen Geometrie, z.B. der Frage nach der maximalen Anzahl paarweise sich berührender \( n \)-dimensionaler Simplices im \( \mathbb{R}^n \).
shows a possible configuration of three 2-dimensional simplices (equilateral triangles) in the plane.

Can you imagine how many translated copies of a regular tetrahedron can be placed in $R^3$ such any two distinct ones touch?

**A generalization of Fisher’s inequality**

Let us begin by recalling Fisher’s inequality.

**Theorem 1** (Fisher). If $F_1, \ldots, F_m$ are $m$ distinct subsets of $\{1, \ldots, n\}$ and there exist integers $r$ and $\lambda$, with $0 \leq \lambda \leq r$, such that the size of each $F_i$ equals $r$ and the intersection of any two distinct sets $F_i$ and $F_j$ contains exactly $\lambda$ elements, then $m \leq n$.

The first ideas for this inequality were found by R.A. Fisher [7] in 1940. Fisher was a statistician, who used a similar inequality in his theory for designing statistical experiments to collect data. R.C. Bose [4] extended the ideas of Fisher in 1949 and proved the inequality as stated in Theorem 1. Bose’s proof is short and elegant; it makes ingenious use of linear algebra. Since then the linear algebra method has been refined and applied to many problems in combinatorics. An extensive collection of examples can be found in an unpublished book by Babai and Frankl [3] and in [1]. We give Bose’s proof in the next paragraph.

First remark that the inequality is trivial if $r = \lambda$, because $m = 1$ in that case. If, on the other hand, $r > \lambda$, we can use linear algebra in the following way: Let $A = (a_{ij})$ be the $n \times m$ incidence matrix given by $a_{ij} = 1$ if $i \in F_j$, and $a_{ij} = 0$ if $i \notin F_j$. Put $M = A^\top A$, where $A^\top$ denotes the transpose of $A$. Now note that $m_{ij} = \sum_{k=1}^{n} a_{ik}a_{kj}$ and hence $m_{ij}$ is equal to the size of the intersection of $F_i$ with $F_j$. Therefore, the $m \times m$ matrix $M$ satisfies:

$$M = \lambda J + (r - \lambda)I, \quad (1)$$

where $J$ is the matrix with all entries equal to one. It suffices to prove that $M$ has rank $m$. Indeed, rank $M = \text{rank } A^\top A \leq \text{rank } A \leq n$, as $A$ has $n$ rows. We first show that $M$ is positive definite, that is, the inner product $\langle x, Mx \rangle > 0$ for all $x \neq 0$. From equation (1) it follows that

$$\langle x, Mx \rangle = \langle x, \lambda Jx \rangle + \langle x, (r - \lambda)Ix \rangle$$

$$= \lambda(x_1 + \ldots + x_m)^2 + (r - \lambda)(x_1^2 + \ldots + x_m^2),$$
so that $\langle x, Mx \rangle > 0$ for all $x \neq 0$, as $r > \lambda$. To complete the proof remark that every positive definite matrix is invertible, as $Mx = 0$ for some $x \neq 0$ implies $\langle x, Mx \rangle = 0$. Thus $M$ is invertible, and hence it has rank $m$.

We wish to replace set with multisets in Fisher’s theorem. Intuitively a multiset is a set with possibly repeated elements. More formally, a multiset on $\{1, \ldots, n\}$ is a function $\mu : \{1, \ldots, n\} \to \mathbb{N} \cup \{0\}$, where $\mu(k)$ is regarded as the number of repetitions of $k$. We will use the notation $F = \{1^{n(1)}, \ldots, n^{n(n)}\}$. The size of $F$ is given by $|F| = \sum_k \mu(k)$, and the intersection of $F$ with another multiset $G = \{1^{v(1)}, \ldots, n^{v(n)}\}$ is defined by $F \cap G = \{1^{\xi(1)}, \ldots, n^{\xi(n)}\}$, where $\xi(k) = \min\{\mu(k), v(k)\}$ for all $k$.

Inspired by Fisher’s inequality we can now ask for the maximum number of distinct multisets $F_1, \ldots, F_m$ on $\{1, \ldots, n\}$ for which there exist integers $r$ and $\lambda$, with $0 \leq \lambda \leq r$, such that each $F_i$ has size $r$ and the intersection of any two distinct multisets $F_i$ and $F_j$ has size $\lambda$. We denote this number by $\psi(n)$. The following example shows that $\psi(n)$ can be larger than $n$: $F_1 = \{1^2, 2^0, 3^1, 4^1\}$, $F_2 = \{1^0, 2^2, 3^1, 4^1\}$, $F_3 = \{1^1, 2^1, 3^2, 4^0\}$, $F_4 = \{1^1, 2^1, 3^0, 4^2\}$, and $F_5 = \{1^0, 2^0, 3^2, 4^2\}$. It is easy to verify that $r = 4$ and $\lambda = 2$ in this example. Thus, our problem is to determine the values of $\psi(n)$.

### How many points at the same distance?

As announced in the introduction our problem can be reformulated in terms of sets in $\mathbb{R}^n$ in which the points are all at the same distance from each other. To do this we need to use the taxicab (Manhattan) distance instead of the usual Euclidean distance. Thus, the distance between two points $x$ and $y$ in $\mathbb{R}^n$ is given by

$$d(x, y) = \sum_k |x_k - y_k|.$$ 

From this formula it is easy to see why this distance is called the taxicab or Manhattan distance, as it resembles the distance that a taxicab has to travel in the rectilinear streets of Manhattan. A more common name in mathematics for the taxicab distance is $\ell_1$-distance and this is the name we shall be using here.

So we wish to study sets in $\mathbb{R}^n$ in which the $\ell_1$-distance between any two distinct points is the same. Such sets are called $\ell_1$-equilateral sets. A simple example in $\mathbb{R}^n$ is the set

$$\{(\pm1, 0, \ldots, 0), (0, \pm1, 0, \ldots, 0), \ldots, (0, \ldots, 0, \pm1)\},$$

with $2n$ points that are all at $\ell_1$-distance 2 from each other.

To obtain the exact variation of our problem we need to put one additional constraint on the $\ell_1$-equilateral sets. Indeed, we shall also require that the $\ell_1$-equilateral set $S$ in $\mathbb{R}^n$ has the property that $\sum_k s_k$ is the same for all $s \in S$. For instance, the set $S$ given by

$$S = \{(2, 0, 1, 1), (0, 2, 1, 1), (1, 1, 2, 0), (1, 1, 0, 2), (2, 2, 0, 0)\}$$

is an $\ell_1$-equilateral set in which the $\ell_1$-distance between any two distinct points is 4 and $\sum_k s_k = 4$ for all $s \in S$. Let us denote the maximum size of such $\ell_1$-equilateral sets in $\mathbb{R}^n$ by $h(n)$. It turns out that the numbers $h(n)$ and $\psi(n)$ are closely related; as matter of fact:
Theorem 2. For all \( n \geq 1 \), \( h(n) \) and \( \varphi(n) \) are equal.

Proof. To show that \( \varphi(n) \leq h(n) \); we let \( \mathcal{F} = \{F_1, \ldots, F_m\} \) be a collection of \( m \) distinct multisets on \( \{1, \ldots, n\} \) such that \( |F_i| = r \) for all \( i \), and \( |F_i \cap F_j| = \lambda \) for all \( i \neq j \). For each \( i \), write \( F_i = \{1^{\mu_i(1)}, \ldots, n^{\mu_i(n)}\} \) and define the vector \( s^i \in \mathbb{R}^n \) by \( s^i = (\mu_i(1), \ldots, \mu_i(n)) \). Put \( S = \{s^1, \ldots, s^m\} \) and remark that \( |S| = m \), as \( \mathcal{F} \) has size \( m \).

Note also that \( \sum_k s^i_k = \sum_k \mu_i(k) = |F_i| = r \) for each \( i \). Moreover,

\[
d(s^i, s^j) = \sum_k |s^i_k - s^j_k| \\
= \sum_k (s^i_k + s^j_k - 2 \min\{s^i_k, s^j_k\}) \\
= \sum_k \mu_i(k) + \sum_k \mu_j(k) - 2 \sum_k \min\{\mu_i(k), \mu_j(k)\} \\
= |F_i| + |F_j| - 2|F_i \cap F_j| \\
= 2r - 2\lambda
\]

for each \( i \neq j \). This implies that \( S \) is an \( \ell_1 \)-equilateral set in \( \mathbb{R}^n \) of size \( m \) and \( \sum_k s_k \) is the same for all \( s \in S \). Therefore, \( \varphi(n) \leq h(n) \) for all \( n \geq 1 \).

To prove the opposite inequality; we assume that the equilateral set is contained in the integer lattice \( \mathbb{Z}^n \). It is known that we may do this without loss of generality. In fact, this was proved by Koolen, Laurent, and Schrijver [9] by using so called cut metrics [6]. So let \( S \) be an \( \ell_1 \)-equilateral set in \( \mathbb{Z}^n \) with \( h(n) \) elements, and suppose that \( \sum_k s_k = r \) for all \( s \in S \). By translating \( S \) we may also assume that each element of \( S \) has nonnegative integer coordinates. For each \( s \in S \) we define a multiset \( F_s = \{1^{s_1}, \ldots, n^{s_n}\} \). Clearly, the multisets \( F_s \) are distinct, and \( |F_s| = \sum_k s_k = r \) for all \( s \in S \). Let \( D \) denote the distance between points in \( S \) and remark that

\[
|F_s \cap F_t| = \sum_k \min\{s_k, t_k\} = 1/2 \sum_k (s_k + t_k - |s_k - t_k|) = (2r - D)/2
\]

for all \( s \neq t \) in \( S \). Hence, \( \varphi(n) \geq h(n) \) for all \( n \geq 1 \), which completes the proof of Theorem 2. \( \square \)

There are some known results for \( h(n) \). To begin, Koolen, Laurent, and Schrijver [9] have proved that \( h(n) = n \) for \( 1 \leq n \leq 3 \), \( h(4) = 5 \), and \( h(n) \geq n + 1 \) for \( n \geq 4 \). Alternatively, if \( n \geq 4 \) and \( n \) is even, we can derive the lower bound, \( n + 1 \), from the following \( \ell_1 \)-equilateral set in \( \mathbb{R}^n \):

\[
\begin{align*}
s^1 &= (a, 0, 1, 1, \ldots, 1, 1) \\
s^2 &= (0, a, 1, 1, \ldots, 1, 1) \\
s^3 &= (1, 1, a, 0, \ldots, 1, 1) \\
s^4 &= (1, 1, 0, a, \ldots, 1, 1) \\
\vdots \\
s^{n-1} &= (1, 1, 1, 1, \ldots, a, 0) \\
s^n &= (1, 1, 1, 1, \ldots, 0, a) \\
s^{n+1} &= (0, 0, 2, 2, \ldots, 2, 2),
\end{align*}
\]

where \( a \) is a positive integer.
where \( a = n - 2 \) and \( \sum_{i} s_i = 2n - 4 \). For odd \( n \) it is not so easy to write down an explicit example of size \( n + 1 \). Other interesting examples are listed in Table 1.

Table 1 Equilateral sets

<table>
<thead>
<tr>
<th>( n = 5 )</th>
<th>( n = 6 )</th>
<th>( n = 8 )</th>
<th>( n = 12 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4, 0, 1, 1, 2)</td>
<td>(4, 0, 1, 1, 1)</td>
<td>(2, 0, 1, 1, 2, 0, 1, 1)</td>
<td>(2, 0, 1, 1, 2, 0, 1, 1, 2, 0, 1, 1)</td>
</tr>
<tr>
<td>(0, 4, 1, 1, 2)</td>
<td>(0, 4, 1, 1, 1)</td>
<td>(2, 0, 1, 1, 0, 2, 1, 1)</td>
<td>(2, 0, 1, 1, 0, 2, 1, 1, 2, 0, 1, 1)</td>
</tr>
<tr>
<td>(1, 1, 4, 0, 2)</td>
<td>(1, 1, 4, 0, 1, 1)</td>
<td>(1, 1, 2, 0, 1, 1, 2, 0)</td>
<td>(1, 1, 2, 0, 1, 1, 2, 0, 1, 1, 2, 0)</td>
</tr>
<tr>
<td>(1, 1, 0, 4, 2)</td>
<td>(1, 1, 0, 4, 1, 1)</td>
<td>(1, 1, 0, 2, 1, 1, 0, 2)</td>
<td>(1, 1, 0, 2, 1, 1, 0, 2, 1, 1, 0, 2)</td>
</tr>
<tr>
<td>(2, 2, 0, 0, 4)</td>
<td>(1, 1, 1, 1, 0, 4)</td>
<td>(0, 0, 0, 0, 2, 2, 2, 2)</td>
<td>(0, 0, 0, 0, 2, 2, 2, 2, 2, 2, 2)</td>
</tr>
<tr>
<td>(2, 2, 0, 2, 4)</td>
<td>(1, 1, 1, 0, 2, 4)</td>
<td>(0, 0, 2, 2, 2, 0, 0, 0)</td>
<td>(0, 0, 2, 2, 2, 0, 0, 0, 0, 0, 0)</td>
</tr>
<tr>
<td>(2, 2, 0, 0, 2)</td>
<td>(1, 1, 2, 0, 2, 0, 2)</td>
<td>(0, 2, 0, 2, 0, 2, 0, 2)</td>
<td>(0, 2, 0, 2, 0, 2, 0, 2, 0, 0, 0, 0)</td>
</tr>
<tr>
<td>(2, 2, 0, 0, 2)</td>
<td>(0, 0, 0, 0, 2, 2, 0, 2)</td>
<td>(0, 0, 0, 0, 2, 2, 0, 2, 0, 2)</td>
<td>(0, 0, 0, 0, 2, 2, 0, 2, 0, 2, 0, 0)</td>
</tr>
</tbody>
</table>

An upper bound for \( h(n) \) has been obtained by Alon and Pudlák [2]. They proved that the maximum size of an \( \ell_1 \)-equilateral set in \( \mathbb{R}^n \) (without the restriction on the sums of the coordinates) is at most \( cn \log n \) for all \( n > 1 \). Closely related to this result is a conjecture of Kusner [8] which says that the maximum size of an \( \ell_1 \)-equilateral set in \( \mathbb{R}^n \) is \( 2n \). We have already seen, at the beginning of this section, the obvious example of an \( \ell_1 \)-equilateral set in \( \mathbb{R}^n \) that attains the upper bound of \( 2n \). In fact, this is the only example known of size \( 2n \). At present Kusner’s conjecture is proved for \( n = 1, 2 \) and 3 in [5], and for \( n = 4 \) in [9]. In connection with these results it also worth mentioning that it can be shown that the size of every \( \ell_1 \)-equilateral set in \( \mathbb{R}^n \) does not exceed \( h(2n - 1) + 1 \) [9]. Therefore it is interesting to investigate if \( h(n) \) is linear in \( n \).

Pairwise touching regular simplices

The second variation of our problem concerns regular simplices. Regular simplices are generalizations of equilateral triangles in the plane and regular tetrahedrons in 3-dimensional space, to general \( n \)-dimensional spaces. A number of them are depicted in Fig. 2. For our problem we wish to place translated copies of a regular \( n \)-dimensional simplex in \( \mathbb{R}^n \) such that any two distinct ones touch but do not overlap. Fig. 1 shows a possible configuration of three translated copies of a regular 2-dimensional simplex in \( \mathbb{R}^2 \) that are pairwise touching. The problem is to determine how many regular \( n \)-dimensional simplices we can place in \( \mathbb{R}^n \) in this way. Let us denote this number by \( t(n) \). You might well wonder how \( t(n) \) could be related to \( \varphi(n) \). Surprisingly, \( t(n - 1) \) and \( \varphi(n) \) are equal for all \( n \geq 1 \). A convenient way to prove this equality is to first show the following result and subsequently to apply Theorem 2.
Theorem 3. For all $n \geq 1$, $t(n-1)$ and $h(n)$ are equal.

As it is difficult to visualize regular simplices in dimension four or more, it is useful to give a formal definition of a regular $n$-dimensional simplex. This definition is quite complicated, but, as we shall see, very practical for proving things. To begin we introduce the following notion. A set $L$ in $\mathbb{R}^N$ is called an affine subspace if it is a translation of a linear subspace $V$ in $\mathbb{R}^N$. It is said to be $n$-dimensional if $V$ has dimension $n$. For instance, any straight line in the plane is a 1-dimensional affine subspace. Now, consider $n+1$ points $a_0, a_1, \ldots, a_n$ in $\mathbb{R}^N$, where $N \geq n$. If $a_0, a_1, \ldots, a_n$ are not contained in an $(n-1)$-dimensional affine subspace in $\mathbb{R}^N$, then the convex hull of $a_0, a_1, \ldots, a_n$ given by

$$\left\{ \sum_i \lambda_i a_i : \lambda_i \geq 0 \text{ for all } i, \text{ and } \sum_i \lambda_i = 1 \right\},$$

is called an $n$-dimensional simplex. It is said to be regular if the standard Euclidean distance between every pair of points in $\{a_0, a_1, \ldots, a_n\}$ is the same.

A simple way to construct a regular $n$-dimensional simplex is to take the convex hull of the $n+1$ standard basis vectors in $\mathbb{R}^{n+1}$ given by

$$\Delta_n = \left\{ x \in \mathbb{R}^{n+1} : x_k \geq 0 \text{ for all } k, \text{ and } \sum_k x_k = 1 \right\}.$$

Fig. 3 illustrates the construction of $\Delta_2$. The simplex $\Delta_n$ is called the standard regular $n$-dimensional simplex $\mathbb{R}^{n+1}$. Equipped with the formal definition of a regular simplex we can now prove Theorem 3.

Proof. To see that $t(n-1) \geq h(n)$ we let $S$ be an $\ell_1$-equilateral set in $\mathbb{R}^n$ such that $\sum_k s_k = r$ for all $s \in S$. If $D$ is the distance between the elements of $S$, then for every distinct $s \neq t$ in $S$ we have that

$$D = \sum_k |s_k - t_k| = \sum_k (2 \max\{s_k, t_k\} - s_k - t_k) = -2r + 2 \sum_k \max\{s_k, t_k\},$$

and hence $\sum_k \max\{s_k, t_k\} = (2r + D)/2$ for all $s \neq t$. Now, define for each $s \in S$ the set $\alpha_s$ by

$$\alpha_s = \left\{ x \in \mathbb{R}^n : x_k \geq s_k \text{ for all } k, \text{ and } \sum_k x_k = (2r + D)/2 \right\}.$$
Note that $\alpha_s$ is a regular $(n - 1)$-dimensional simplex in the affine subspace $\{ x \in \mathbb{R}^n : \sum_k x_k = (2r + D)/2 \}$ (Fig. 4 (A)). Moreover, the simplices $\alpha_s$ and $\alpha_t$ are translates of each other, and they touch each other at the point $s \vee t$ with coordinates $(s \vee t)_k = \max\{ s_k, t_k \}$ for $1 \leq k \leq n$, whenever $s \neq t$. Thus, $\{ \alpha_s : s \in S \}$ is a collection of pairwise touching translates of a regular $(n - 1)$-dimensional simplex, and hence $t(n - 1) \geq h(n)$ for all $n \geq 1$.

To show the opposite inequality we let $\{ \alpha_i : i \in I \}$ be a collection of pairwise touching translations of a regular $(n - 1)$-dimensional simplex. We may assume that each $\alpha_i$ is a translate of the standard regular simplex $\Delta_{n-1}$ and that each $\alpha_i$ is contained in the affine subspace $\{ x \in \mathbb{R}^n : \sum_k x_k = 1 \}$. Note that for each $i \in I$, there exists a vector $x^i \in \mathbb{R}^n$, with $\sum_k x^i_k = 0$, such that

$$
\alpha_i = \left\{ x \in \mathbb{R}^n : x_k \geq x^i_k \text{ for all } k, \text{ and } \sum_k x_k = 1 \right\}
$$

(Fig. 4 (B)). The intersection of $\alpha_i$ and $\alpha_j$ satisfies

$$
\alpha_i \cap \alpha_j = \left\{ x \in \mathbb{R}^n : x_k \geq \max\{ x^i_k, x^j_k \} \text{ for all } k, \text{ and } \sum_k x_k = 1 \right\}.
$$

(2)
As the simplices $\alpha_i$ and $\alpha_j$ are touching, it follows from (2) that they touch at the point $x^i \lor x^j$ with coordinates $(x^i \lor x^j)_k = \max\{x^i_k, x^j_k\}$ for $1 \leq k \leq n$. This implies that $
\sum_k \max\{x^i_k, x^j_k\} = 1$. Therefore, we can use the equality

$$d(x^i, x^j) = \sum_k |x^i_k - x^j_k| = \sum_k (2 \max\{x^i_k, x^j_k\} - x^i_k - x^j_k)$$

to deduce that $d(x^i, x^j) = 2$ for all $i \neq j$. Thus, $\{x^i : i \in I\}$ is an $\ell_1$-equilateral set in $\mathbb{R}^n$ such that $\sum_k x^i_k = 0$ for all $i \in I$ and hence $t(n - 1) \leq h(n)$ for every $n \geq 1$, which completes the proof.

Note that $h(4) = 5$, so that by Theorem 3 we can place five translated copies of a regular tetrahedron in $\mathbb{R}^3$ such that any two distinct ones touch. At first glance this might seem impossible, but it can be done. Fig. 5 shows the way to do it. A simple way to think about this configuration is to see it as a part of the regular lattice packing pictured in Fig. 6.

![Fig. 5 Touching tetrahedrons](image)

![Fig. 6 Lattice packing of tetrahedrons](image)
We conclude by noting that despite the variations we have presented, the problem has not given away its secrets; but, it certainly has become more intriguing.

Acknowledgment

The author is grateful to Sandjai Bhulai for helping him with Figure 5, and to Pyrrhos Stathis for taking the picture for Figure 6. The author also wishes to thank Jan Lemmens for making the sculpture shown in Figure 6.

References


Bas Lemmens
Mathematics Institute
University of Warwick
CV47AL Coventry, U.K.
e-mail: lemmens@maths.warwick.ac.uk