COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA 13 (4) (1993) 483–495

k-SETS AND RANDOM HULLS*

MICHA SHARIR

Received July 29, 1990

We re-examine the probabilistic analysis of Clarkson and Shor [5] involving *k-sets* of point sets and related structures. By studying more carefully the equations that they derive, we are able to obtain refined analysis of these quantities, which lead to a collection of interesting relationships involving k-sets, convex hulls of random samples, and generalizations of these constructs.

1. Introduction

Let S denote a set of n points in the plane in general position (in particular, no three points lie on a line). For each ordered pair, (a,b), of points in S, we define its *weight* $w(a,b)$ to be the number of points of S lying to the right of the line passing through a and b and directed from a to b. Note that $w(a,b)=n-2-w(b,a)$. Let us denote by f_k the number of pairs (a,b) whose weight is k, for $k=0,\dots,n-2$. f_k is closely related to the number of *k-sets* of S, i.e. subsets of S of size k that can be separated from their complement by a line, in the sense that for each pair (a,b) of weight k the subset of points of S lying to the right of the line $\vec{ab}$ is a k-set, and for each k-set S' of S there exists at least one pair (possibly several pairs) (a,b) of weight k (or $k-1$) so that the subset of points of S lying to the right of $\vec{ab}$ is S' (or S' with the point a or b removed). See [1,2,3,5,6,7,8,11,13,14] for more information concerning k-sets.

Obtaining sharp upper bounds on the maximum possible number of k-sets for a set of n points in the plane is one of the most difficult and tantalizing open problems in combinatorial geometry. An easy upper bound of $O(n\sqrt{k})$ has been obtained several times in the past 20 years (see [7], [11], [6]), and a tiny improvement, to

AMS subject classification code (1991): 52 A 22, 52 A 37, 05 C 99

* Work on this paper has been supported by Office of Naval Research Grant N00014-89-J-3042 and N00014-90-J-1284, by National Science Foundation Grant CCR-89-01484, and by grants from the U.S.-Israeli Binational Science Foundation, the Fund for Basic Research administered by the Israeli Academy of Sciences, and the G.I.F., the German-Israeli Foundation for Scientific Research and Development.

$O(n\sqrt{k}/\log^\star k)$, has been recently obtained by Pach, Steiger and Szemerédi [13]. But these bounds are still a far cry from the known lower bound of $\Omega(n \log k)$.

In this paper we analyze in great detail the relationships, noted by Clarkson and Shor [5], between the quantities f_k and the quantities h_r — the expected number of vertices of the convex hull of a random sample of size r from S. These relationships form a linear system of equations, whose matrix is easy to invert, and possesses various properties that enable us to derive numerous interesting formulas involving these quantities. Among our results we have formulas that express each f_k as a linear combination of the h_r's, a system of linear relationships between the h_r's that hold for any set of n points in the plane, formulas for certain sums involving the f_k's, and several other 'universal' properties of the sequences $\{h_r\}$.

Since the analysis of [5] applies in a much more general context, so does our machinery. We give a few extensions of our methods to sets of points in higher dimensions, to Delaunay triangulations in the plane, and more.

Unfortunately, so far our machinery does not seem to imply upper bounds for the quantities f_k; we do not know whether our technique can even imply the weak bound of $O(n\sqrt{k})$ mentioned above. We believe however that our technique gains additional insights into the structure of k-sets, and hope that it will eventually be useful for cracking the 20 years old k-set problem.

The paper is organized as follows. Section 2 presents the basic results concerning k-sets and random hulls in the plane. Section 3 derives a few additional formulas, and Section 4 discusses extensions to higher dimensions and more general contexts. We conclude the paper in Section 5 with a discussion of our results and some open problems.

2. k-sets and random hulls

Before we begin our analysis, we note the *symmetry property* of the f_j's (which is immediate from the definition):

$$f_k = f_{n-2-k}\,, \qquad k = 0, 1, \ldots n-2\,. \tag{1}$$

For each $r = 2, 3, \ldots, n$ let h_r denote the expected number of vertices of the convex hull of a random sample of r points of S (where each subset of r elements has equal probability to be chosen). We always have $3 \le h_r \le r$. Also $h_r = r$ for every $r = 2, \ldots, n$ if and only if the points of S are in convex position.

Clarkson and Shor [5] have established a linear relationship between the quantities f_k and h_r. Specifically, they have shown, using straightforward probabilistic analysis, that

$$\sum_{j=0}^{n-r} \frac{\binom{n-j-2}{r-2}}{\binom{n}{r}} f_j = h_r \tag{2}$$

for $r = 2, 3, \ldots, n$. This system of equations is non-singular; in fact it is triangular with 1's in the lower-left to upper-right diagonal and 0's below that diagonal. It is therefore quite easy to solve it explicitly. Before doing so, let us first simplify the system, by rescaling the variables f_j and h_r as follows

$$f_k = \frac{n!}{k!(n-k-2)!} x_k , \qquad k = 0, \ldots, n-2 \tag{3}$$

$$g_r = \frac{h_r}{r(r-1)} , \qquad r = 2, \ldots, n . \tag{4}$$

We then have

Lemma 2.1. *The variables x_k and g_r satisfy the following system of equations:*

$$\sum_{j=0}^{m} \binom{m}{j} x_j = g_{n-m} , \tag{5}$$

for $m=0,1,\ldots,n-2$.

Proof. Let us replace in (2) the h_r's by the g_r's, and the f_j's by the x_j's. We thus obtain

$$\sum_{j=0}^{n-r} A_{rj} x_j = g_r ,$$

where

$$A_{rj} = \frac{\binom{n-j-2}{r-2}}{r(r-1)\binom{n}{r}} \cdot \frac{n!}{j!(n-j-2)!} = \binom{n-r}{j} .$$

The assertion of the lemma is now immediate. ∎

Lemma 2.2. *For each $k=0,\ldots,n-2$ we have*

$$x_k = g_{n-k} - \binom{k}{1} g_{n-k+1} + \binom{k}{2} g_{n-k+2} \cdots (-1)^k g_n , \tag{6}$$

or, in other words,

$$f_k = \frac{n!}{k!(n-k-2)!} \left[g_{n-k} - \binom{k}{1} g_{n-k+1} + \binom{k}{2} g_{n-k+2} \cdots (-1)^k g_n \right] . \tag{7}$$

Proof. The assertion of the lemma is an easy consequence of the *inversion formula* (5.58) of [9], applied to the system (5). ∎

Remark: The right-hand side of (6) is the last element in the k-th order difference sequence of the sequence $\mathbf{G}=(g_2,g_3,\ldots,g_n)$.

We note that the system (5) has many nice properties. To state them, let Δ denote the difference operator, defined so that if $U=(u_1,u_2,\ldots,u_n)$ then

$$\Delta U = (u_1 - u_2, u_2 - u_3, \ldots, u_{n-1} - u_n) .$$

Let Δ^k denote the k-fold iteration of Δ; since in our case the sequences start with index 2, we will use the same convention for the differenced sequences; thus, for example, $(\Delta\mathbf{G})_2 = g_2 - g_3$, etc.

Lemma 2.3. *For each $k,m \geq 0$, $k+m \leq n-2$, we have*

$$\sum_{j=0}^{m} \binom{m}{j} x_{j+k} = (\Delta^k \mathbf{G})_{n-m-k} \,. \tag{8}$$

Proof. This is easily verified, using standard properties of binomial coefficients. ∎

Note that equation (6) is a special case of (8) with $m=0$.

Corollary 2.4. *The sequence* **G** *is monotonically decreasing and convex.*

Proof. Put $k=1,2$ in (8). ∎

Next we observe that the system (5) can be solved in closed form (in other words, the sums (6) and (7) can be simplified) for certain values of the vector **G**. We have:

Lemma 2.5. *If $g_r = \frac{1}{(r-a)(r-a-1)\dots(r-q)}$, for two fixed integers $a \leq q$ and for $r = 2,\dots,n$, then*

$$x_k = \frac{(k+q-a)!(n-k-q-1)!}{(q-a)!(n-a)!}\,, \qquad k = 0,\dots,n-2\,. \tag{9}$$

Proof. We substitute these values of the g_r's into (6), and thus need to compute the sum

$$\frac{1}{(n-k-a)\dots(n-k-q)} - \binom{k}{1}\frac{1}{(n-k-a+1)\dots(n-k-q+1)} +$$
$$\binom{k}{2}\frac{1}{(n-k-a+2)\dots(n-k-q+2)} \dots + (-1)^k \frac{1}{(n-a)\dots(n-q)}\,.$$

Consider the function

$$F(x) = x^{n-k-q-1} - \binom{k}{1}x^{n-k-q} + \binom{k}{2}x^{n-k-q+1} + \dots$$
$$\dots + (-1)^k x^{n-q-1} = x^{n-k-q-1}(1-x)^k\,.$$

Then the sum above is the $(q-a+1)$-fold indefinite integral of $F(x)$, evaluated at 1. As is well known, this is equal to

$$\frac{1}{(q-a)!}\int_0^1 F(x)(1-x)^{q-a}dx = \frac{1}{(q-a)!}\int_0^1 x^{n-k-q-1}(1-x)^{k+q-a}dx =$$
$$\frac{(k+q-a)!(n-k-q-1)!}{(q-a)!(n-a)!}\,,$$

as asserted. ∎

Let us check these equations for the case when all points of S lie in convex position, that is when $h_r = r$ and $g_r = \frac{1}{r-1}$, for $r = 2, \ldots, n$. In this case $a = q = 1$, so

$$x_k = \frac{k!(n-k-2)!}{(n-1)!} ,$$

which in turn implies that $f_k = n$ for each k, a result that can also be immediately established by direct geometric reasoning.

Next we study the implications of the symmetry property (1) of the f_j's, which is easily seen to translate to a similar property of the x_j's; that is, $x_j = x_{n-2-j}$ for $j = 0, \ldots, n-2$. This property implies that the sequences (or, rather, vectors) $\mathbf{G}$ have only $\lfloor \frac{n}{2} \rfloor$ degrees of freedom. Thus there should exist $\lfloor \frac{n-1}{2} \rfloor$ independent linear dependencies among the values $g_2, \ldots, g_n$ that must hold for any set S of n points in the plane. There are several ways to obtain such a system of dependencies, but the most elegant that we have found are the following two systems:

Proposition 2.6. *For each $k = 0, 1, \ldots, \lfloor \frac{n-3}{2} \rfloor$ we have*

$$(\Delta^k \mathbf{G})_{k+2} = 2(\Delta^k \mathbf{G})_{k+3} \, . \tag{10}$$

Proof. It follows from basic properties of the binomial coefficients that (10) can be rewritten as

$$\sum_{j=0}^{n-2k-2} \binom{n-2k-2}{j} x_{j+k} = 2 \sum_{j=0}^{n-2k-3} \binom{n-2k-3}{j} x_{j+k} \, .$$

By the symmetry property (1) of the f_j's we have

$$\sum_{j=0}^{n-2k-3} \binom{n-2k-3}{j} x_{j+k} = \sum_{j=0}^{n-2k-3} \binom{n-2k-3}{j} x_{n-k-2-j} \, .$$

Putting $j = n-2k-2-\ell$, the sum becomes

$$\sum_{\ell=1}^{n-2k-2} \binom{n-2k-3}{n-2k-2-\ell} x_{\ell+k} = \sum_{\ell=1}^{n-2k-2} \binom{n-2k-3}{\ell-1} x_{\ell+k} \, .$$

Hence

$$2 \sum_{j=0}^{n-2k-3} \binom{n-2k-3}{j} x_{j+k} =$$

$$\sum_{j=0}^{n-2k-3} \binom{n-2k-3}{j} x_{j+k} + \sum_{j=1}^{n-2k-2} \binom{n-2k-3}{j-1} x_{j+k} =$$

$$\sum_{j=0}^{n-2k-2} \binom{n-2k-2}{j} x_{j+k} ,$$

as asserted. ∎

As a corollary, we obtain an even more elegant system of dependencies:

Proposition 2.7. *Define* $b_r = h_r/r = (r-1)g_r$, *for* $r = 2,\dots,n$. *Then, for* $k = 1,2,\dots,\lfloor\frac{n-1}{2}\rfloor$ *we have*

$$(\Delta^k \mathbf{B})_{k+1} = 0 \,, \tag{11}$$

where $\mathbf{B}$ *is the vector* $(b_2, b_3, \dots, b_n)$.

Proof. This is just a rewriting of equations (10). Indeed, we can write (10) as

$$\sum_{\ell=0}^{k}(-1)^\ell \binom{k}{\ell} g_{k+\ell+2} - 2\sum_{\ell=0}^{k}(-1)^\ell \binom{k}{\ell} g_{k+\ell+3} = 0$$

or

$$\sum_{\ell=0}^{k+1}(-1)^\ell \left[\binom{k}{\ell} + 2\binom{k}{\ell-1}\right] g_{k+\ell+2} = 0 \,.$$

But, as is easily checked,

$$\left[\binom{k}{\ell} + 2\binom{k}{\ell-1}\right] g_{k+\ell+2} = \frac{1}{k+1}\binom{k+1}{\ell}(k+\ell+1)g_{k+\ell+2} =$$

$$\frac{1}{k+1}\binom{k+1}{\ell} b_{k+\ell+2} \,.$$

Hence, multiplying by $k+1$ and replacing $k+1$ by k, we obtain the asserted equality.∎

One interesting property of the equations (10) and (11) is that only their number depends on n, but their coefficients do not; that is, as n grows we get the same relationships as for smaller values of n, plus additional ones.

Let us write down explicitly the first few relationships (11). For $k=1$ we have

$$b_2 - b_3 = 0 \,.$$

This follows trivially from the definition of the b_r's, since we always have $b_2 = h_2/2 = 1$, and $b_3 = h_3/3 = 1$.

For $k=2$ we have $b_3 - 2b_4 + b_5 = 0$, or

$$\frac{1}{3}h_3 - \frac{1}{2}h_4 + \frac{1}{5}h_5 = 0 \,,$$

or

$$h_5 = 2.5h_4 - 5 \,.$$

In geometric terms, this says that if we draw a random sample of 5 points out of an arbitrary n-point set S and then draw a random sample of 4 points from S, then the expected sizes of the hulls of the 5-sample and of the 4-sample satisfy the above relationship. But a random sample of 4 points can also be obtained by first choosing a random sample of 5 points and then choose a random subsample of 4 points from it. It follows that the above relationship has actually nothing to do with S. It states that if we take any set of 5 points in the plane (in general position), whose hull has h_5 vertices, and we draw a random sample of 4 points from it, then the expected size of the hull of the sample (over all possible draws) and h_5 satisfy the above relationship; that is, the expected size of the hull of the 4-sample is $\frac{2}{5}h_5+2$. The reader is invited to check this directly for each of the three essentially different ways to arrange 5 points in the plane — a convex pentagon, a convex quadrilateral with one point inside, and a triangle with two points inside.

For $k=3$ we obtain

$$b_4 - 3b_5 + 3b_6 - b_7 = 0 .$$

Again, we can rewrite this in terms of the h_r's. With some manipulations, we obtain

$$h_7 = \frac{7}{2}h_6 - \frac{35}{4}h_4 + 21 ,$$

which gives a "universal" linear dependence between the size of the hull of an arbitrary 7-point set in the plane and the average sizes of the hulls of 6-subsets and of 4-subsets of the set.

The dependencies (10) have another interesting consequence. Consider for example the case $k=3$, with $n=2k+3=9$. The relationship (10) states that

$$g_5 - 3g_6 + 3g_7 - g_8 = 2(g_6 - 3g_7 + 3g_8 - g_9) .$$

But $g_6-3g_7+3g_8-g_9$ is x_3 for the given 9-point set S, by (6). On the other hand, $g_5-3g_6+3g_7-g_8$ is the expected value of x_3 in a random sample of 8 points from S. Translating back into the f_k's, we obtain that the expected value of f_3 in a random 8-sample of S is $\frac{8}{9}f_3(S)$. In general, we obtain

Proposition 2.8. *Let S be an arbitrary set of $2k+3$ points in general position in the plane. Then the expected value of f_k in a random sample of $2k+2$ points from S is $\frac{2k+2}{2k+3}f_k(S)$.*

Remark: If S has $n>2k+3$ points, the above argument shows that the expected value of f_k in a $(2k+2)$-sample of S is $\frac{2k+2}{2k+3}$ times that value for a $(2k+3)$-sample.

3. Sums involving the f_k's

In this section we derive several interesting formulas involving sums of the f_k's. We first recall the following formula (see equation (5.26) of [9]):

$$\sum_{m=j}^{p-q} \binom{p-m}{q}\binom{m}{j} = \binom{p+1}{q+j+1} , \tag{12}$$

for integers $j\geq 0$, $p\geq q\geq 0$.

Proposition 3.1. *For each integer $k\geq 0$ we have*

$$\sum_{j=0}^{n-2} \frac{f_j}{(j+1)(j+2)\dots(j+k+1)} = \frac{n!}{(n+k-1)!}\sum_{r=2}^{n}\binom{r+k-2}{k}g_r . \tag{13}$$

Proof. We expand the sum on the right hand side of (13), replacing g_r by the left hand side of (5) and applying (12):

$$\sum_{r=2}^{n}\binom{r+k-2}{k}g_r = \sum_{r=2}^{n}\sum_{j=0}^{n-r}\binom{r+k-2}{k}\binom{n-r}{j}x_j =$$

$$\sum_{j=0}^{n-2}\left[\sum_{r=2}^{n-j}\binom{r+k-2}{k}\binom{n-r}{j}\right]x_j =$$

$$\sum_{j=0}^{n-2}\binom{n+k-1}{j+k+1}x_j = \sum_{j=0}^{n-2}\frac{(n+k-1)!j!(n-j-2)!}{(j+k+1)!(n-j-2)!n!}f_j =$$

$$\frac{(n+k-1)!}{n!}\cdot\sum_{j=0}^{n-2}\frac{j!}{(j+k+1)!}f_j ,$$

from which (13) readily follows. ∎

Again, let us look at a few special cases of (13). In doing so, we will also derive upper bounds on these sums, taking advantage of the facts that the right hand side of (13) has only positive terms, and that $g_r \le \frac{1}{r-1}$:

For $k=0$ we have

$$\sum_{j=0}^{n-2}\frac{f_j}{j+1} = n\sum_{r=2}^{n}g_r \le nH_{n-1} ,$$

where H_m is the m-th harmonic number.

For $k=1$ we have

$$\sum_{j=0}^{n-2}\frac{f_j}{(j+1)(j+2)} = \sum_{r=2}^{n}(r-1)g_r \le n-1 .$$

In general, we have

Lemma 3.2. *For each $k \ge 1$ we have*

$$\text{(14)}\qquad \sum_{j=0}^{n-2}\frac{f_j}{(j+1)(j+2)\dots(j+k+1)} \le \frac{n-\frac{k!}{(n+1)(n+2)\dots(n+k-1)}}{k\cdot k!} .$$

Proof. Straightforward, and is left to the reader. ∎

Remarks: (1) The preceding analysis is general, and does not require the symmetry property for the f_k's. It therefore also applies in the more general context discussed in Section 4.

(2) There are applications of (13) and (14) to complexity analysis of randomized incremental algorithms, in which the points of S are inserted one by one in a random order. For instance, in the case $k=1$, the fraction $\frac{1}{(j+1)(j+2)}$ is half the probability that a particular pair of points of S will ever arise as an edge of the convex hull during the process, so the sum (14) implies that the expected number of edges that ever appear on the hull during the random insertion process is $O(n)$; in fact, the precise worst-case upper bound is $2(n-1)$, which is attained if and only if the points of S are in convex position. Similarly, the case $k=0$ yields bounds on the expected time complexity of certain steps of such an incremental randomized algorithm. We refer the reader to [5,10,12,14,15] for related analyses, but note that our analysis (as well as that in [15]) gives the best (worst-case) constants of proportionality in the resulting bounds. So far, we do not have applications of this sort for the sums with $k \geq 2$.

(3) The sums (13) can be interpreted in terms of the generating function

$$F(z) = \sum_{j=0}^{n-2} f_j z^j .$$

Indeed, they give the values $F^{(-k)}(1)$, where $F^{(-k)}$ is the k-fold indefinite integral of F. Can this be used to obtain upper bounds on the f_j's?

4. Extensions

An interesting feature of the system (5) is its "universality" in the sense that it covers a wide class of similar geometric situations. The context is as follows (see also [5]). We have a set S of n objects (points in the plane in the previous section) and with every (ordered or unordered) b-tuple of them ($b=2$ in the previous section) we associate an integer weight by mapping the b-tuple into some region in space (an open halfspace in the previous section) and counting the number of objects of S that meet this region (excluding the objects in the b-tuple itself). We denote by F_k the number of b-tuples with weight k, and by H_r the expected number of 0-weight b-tuples in a random sample of r objects from S. We then have (see [5]):

$$\sum_{j=0}^{n-r} \frac{\binom{n-j-b}{r-b}}{\binom{n}{r}} F_j = H_r$$

for $r=b,b+1,\ldots,n$. We now rescale the variables F_j and H_r by substituting

$$F_k = \frac{n!}{k!(n-k-b)!} X_k , \qquad k = 0,\ldots,n-b$$

$$G_r = \frac{H_r}{r(r-1)\ldots(r-b+1)} = \frac{(r-b)!}{r!} H_r , \qquad r = b,\ldots,n .$$

It is easily verified that the substitution leads again to (5); more precisely, we obtain

$$\sum_{j=0}^{m} \binom{m}{j} X_j = G_{n-m} , \tag{15}$$

for $m=0,1,\ldots,n-b$.

Suppose, in addition, that the F_k's satisfy a symmetry property of the form

$$F_k = F_{n-b-k} , \qquad k = 0,1,\ldots n-b . \tag{16}$$

Then we can extend Propositions 2.6 and 2.7 to obtain

Proposition 4.1. *For each* $k=0,1,\ldots,\lfloor\frac{n-b-1}{2}\rfloor$ *we have*

$$(\Delta^k \mathbf{G})_{k+b} = 2(\Delta^k \mathbf{G})_{k+b+1} \; . \tag{17}$$

Proposition 4.2. *Define* $B_r = (r-b+1)G_r$, *for* $r = b,\ldots,n$. *Then, for* $k = 1,2,\ldots,\lfloor\frac{n-b+1}{2}\rfloor$ *we have*

$$(\Delta^k \mathbf{B})_{k+b-1} = 0 , \tag{18}$$

where $\mathbf{B}$ *is the vector* $(B_b, B_{b+1}, \ldots, B_n)$.

The proofs of these propositions are straightforward modifications of the proofs of Propositions 2.6 and 2.7, and we leave them to the reader.

Here is one example of such an extension. Let S be a set of n points in general position in 3-space, and let $b=3$. With every oriented (but otherwise unordered) triple of points of S we associate the open halfspace whose bounding plane passes through the three points and whose inward-drawn normal is positively oriented with the triple. The weight of the triple is the number of points of S in that halfspace. This fits into the above context, provided H_r is defined as the expected number of *faces* of the convex hull of a random sample of r points from S. The symmetry property (16) clearly also holds.

As a corollary, we obtain

Corollary 4.3. *In the context of point sets in 3-space, let* $G_r = \frac{H_r}{r(r-1)(r-2)}$ *and* $B_r = \frac{H_r}{r(r-1)}$, *for* $r=3,\ldots,n$. *Then*

(a) *The sequence* G_r *is monotonically decreasing and convex.*

(b)

$$(\Delta^k \mathbf{B})_{k+2} = 0$$

for $k=1,2,\ldots,\lfloor\frac{n}{2}\rfloor-1$.

(c)

$$(\Delta^k \mathbf{G})_{k+3} = 2(\Delta^k \mathbf{G})_{k+4}$$

for $k=0,1,\ldots,\lfloor\frac{n}{2}\rfloor-2$.

(d) *For* $k=0,1,\ldots,n-3$ *we have*

$$F_k=\frac{n!}{k!(n-k-3)!}\left[G_{n-k}-\binom{k}{1}G_{n-k+1}+\binom{k}{2}G_{n-k+2}\cdots(-1)^kG_n\right].$$

(e) *Let* V_r *denote the expected number of vertices of the hull of an* r*-sample of* S, *and let* $W_r=\frac{V_r}{r(r-1)}$. *Then we have*

$$(\Delta^kW)_{k+2}=\frac{2k!(k+1)!}{(2k+2)!}$$

for $k=1,2,\ldots,\lfloor\frac{n}{2}\rfloor-1$.

Proof. The first four parts of the lemma follow from the general arguments given above. For the fifth assertion, we note that Euler's formula implies (since the points of S are in general position) $H_r=2V_r-4$, so that $B_r=2W_r-\frac{4}{r(r-1)}$. We thus need to apply the operator $(\Delta^k\cdot)_{k+2}$ to the sequence $(\frac{1}{3\cdot2},\frac{1}{4\cdot3},\ldots)$. Applying equation (5.51) of [9], we readily obtain the asserted equality. ∎

Again, it is interesting to consider a few special cases of these formulas. For instance, consider (b) with $k=1$. We have $B_3=B_4$ or $H_3=\frac{1}{2}H_4$. This is clearly correct, because $H_3=2$ (the convex hull of 3 points consists of two oppositely-oriented faces) and $H_4=4$.

For $k=2$ we have $B_4-2B_5+B_6=0$, or, after some manipulations,

$$H_6=3H_5-10\,;\qquad V_6=3V_5-9\,.$$

As above, these are general relationships between the hull size of a set of 6 points in 3-space and the average size of the hull of subsets of 5 points of that set. The reader is invited to check the validity of these relationships for a few configurations.

The preceding analysis can easily be extended to obtain relationships between k-sets and random hulls in any dimension.

Let us also apply equations (9) in the case when all points of S lie in convex position in 3-space. In this case we have $H_r=2r-4$ for each r, so $G_r=\frac{2}{r(r-1)}$. This gives

$$X_k=\frac{2(k+1)!(n-k-2)!}{n!},$$

or

$$F_k=2(k+1)(n-k-2)$$

for $k=0,\ldots,n-3$. We have thus shown that

Lemma 4.4. *The number of* k*-sets of a set* S *of* n *points in convex position in 3-space (or rather the number of* k*-triangles, namely triangles spanned by* S *and containing* k *points of* S *on one side) is* $2(k+1)(n-k-2)$, *for* $k=0,\ldots,n-3$.

Is there a more direct proof of the lemma?

We conclude this section with another example where our general system (15) arises. Let us go back to the case of k-sets of planar point sets, and let $F_k = \sum_{j=0}^{k} f_j$, for $k=0,\ldots,n-2$. Substitute $f_j = F_j - F_{j-1}$ in (2) to obtain

$$\sum_{j=0}^{n-r} \frac{\binom{n-j-2}{r-2}}{\binom{n}{r}} (F_j - F_{j-1}) = h_r ,$$

for $r=2,\ldots,n$. We ignore the equation for $r=2$, since it merely states that $F_{n-2} = 2\binom{n}{2}$. removing this equation also eliminates the variable F_{n-2}, and we are left with $n-2$ equations in $n-2$ variables, namely

$$\sum_{j=0}^{n-r} \left[\frac{\binom{n-j-2}{r-2}}{\binom{n}{r}} - \frac{\binom{n-j-3}{r-2}}{\binom{n}{r}} \right] F_j = h_r ,$$

or

$$\sum_{j=0}^{n-r} \frac{\binom{n-j-3}{r-3}}{\binom{n}{r}} F_j = h_r ,$$

for $r=3,\ldots,n$. Hence we have again our universal system, with $b=3$, so we can apply our whole battery of results as described above. (Note however that the F_k's do not satisfy the symmetry property.) Note also that this transformation is fairly general, and can be used to obtain equations involving the prefix sums of the f_j's in general.

Remark: There is one more case of the general context that we would like to mention here. It concerns Delaunay triangulations in the plane. More precisely, let S be a set of n points in general position in the plane, and for each triple of points of S assign weight equal to the number of points of S in the open disc whose bounding circle passes through the triple. It is easily checked that this fits into our general context, with $b=3$ and with H_r defined as the expected number of Delaunay triangles in the triangulation of a radnom sample of r points of S; note that we have $H_r = 2h_r - 5$, where h_r is the expected hull size as in Section 2. This particular case arises in the analysis of the recent randomized incremental algorithm of Guibas et al. [10], and is studied in detail in the companion paper [15].

5. Conclusion

In this paper we have derived a potpourri of formulas and other relationships that involve k-sets, random hulls and their generalizations. We are not aware of other, perhaps more direct, proofs of our results. Our analysis is a refinement of that in [5]. We hope that, besides the general combinatorial flavor of our results, they will also turn out to be useful in the analysis of various randomized algorithms, and of course in deriving upper bounds for the number of k-sets.

It is also interesting to draw an analogy between the analysis given above and the theory of f-vectors and h-vectors for convex polytopes [4]. The reader familiar with that theory can spot several similarities (albeit on a rather superficial level) between the two setups. Is there a concrete geometric transformation that maps the setup that we have considered to that of convex polytopes, simplicial complexes, etc.?

References

[1] N. Alon, and E. Győri: The number of small semispaces of a finite set of points in the plane, *J. Combin. Theory Ser. A* **41** (1986), 154–157.

[2] B. Aronov, B. Chazelle, H. Edelsbrunner, L. Guibas, M. Sharir, and R. Wenger: Points and triangles in the plane and halving planes in space, *Discrete Comput. Geom.* **6** (1991), 435–442.

[3] I. Bárány, Z. Füredi, and L. Lovász: On the number of halving planes, *Proc. 5th ACM Symp. on Computational Geometry,* 1989, 140–144.

[4] Brønstad: *An introduction to Convex Polytopes*, Springer-Verlag, Heidelberg 1983.

[5] K. Clarkson, and P. Shor: Applications of random sampling in computational geometry, II, *Discrete Comput. Geom.* **4** (1989), 387–421.

[6] H. Edelsbrunner, and E. Welzl: On the number of line separations of a finite set in the plane, *J. Combin. Theory Ser. A* **38** (1985), 15–29.

[7] P. Erdős, L. Lovász, A. Simmons, and E. G. Strauss: Dissection graphs of planar point sets, In *A Survey of Combinatorial Theory,* J. N. Srivastava et al., eds., North-Holland, Amsterdam, 1973, 139–149.

[8] J. E. Goodman, and R. Pollack: On the number of k-subsets of a set of n points in the plane, *J. Combin. Theory, Ser. A* **36** (1984), 101–104.

[9] R. Graham, D. Knuth, and O. Patashnik: *Concrete Mathematics,* Addison–Wesley, Reading, MA, 1989.

[10] L. Guibas, D. Knuth, and M. Sharir: Randomized incremental construction of Delaunay and Voronoi Diagrams, *Algorithmica* **7** (1992), 381–413.

[11] L. Lovász: On the number of halving lines, *Ann. Univ. Sci. Budapest, Eötvös, Sect. Math.* **14** (1971), 107–108.

[12] C. Ó'Dúnlaing, K. Mehlhorn, and S. Meiser: Abstract Voronoi diagrams, manuscript, 1989.

[13] J. Pach, W. Steiger, and E. Szemerédi: An upper bound on the number of planar k-sets, *Proc. 30th IEEE Symp. on Foundations of Computer Science,* 1989, 72–79.

[14] M. Sharir: On k-sets in arrangements of curves and surfaces, *Discrete Comput. Geom.* **6** (1991), 593–613.

[15] E. Yaniv: Randomized incremental construction of Delaunay triangulations: Theory and practice, M. Sc. thesis, Tel Aviv University, Tel Aviv, Israel, 1991.

Micha Sharir

School of Mathematical Sciences
Tel Aviv University
`sharir@math.tau.ac.il`

and

Courant Institute of Mathematical Sciences
New York University