ICM 2002 · Vol. III · 1–3

Random points, convex bodies, lattices

Imre Bárány*

Abstract

Assume K is a convex body in R^d, and X is a (large) finite subset of K. How many convex polytopes are there whose vertices come from X? What is the typical shape of such a polytope? How well the largest such polytope (which is actually $\operatorname{conv} X$) approximates K? We are interested in these questions mainly in two cases. The first is when X is a random sample of n uniform, independent points from K and is motivated by Sylvester's four-point problem, and by the theory of random polytopes. The second case is when $X = K \cap Z^d$ where Z^d is the lattice of integer points in R^d. Motivation comes from integer programming and geometry of numbers. As we will see, the two cases behave quite similarly.

2000 Mathematics Subject Classification: 52A22, 05A16, 52C07.
Keywords and Phrases: Convex bodies, Lattices, Random samples, Convex polytopes, Limit shape.

1. Sylvester's question

In the 1864 April issue of the Educational times J.J. Sylvester [26] posed the innocent looking question that read: *"Show that the chance of four points forming the apices of a reentrant quadrilateral is 1/4 if they be taken at random in an indefinite plane."* It was understood within a year that the question is ill-posed. (The culprit is, as we all know by now, the "indefinite plane" since there is no natural probability measure on it.) So Sylvester modified the question: let K be a convex body and choose four random, independent points uniformly from K, and write $P(K)$ for the probability that the four points form the apices of a reentrant quadrilateral, or, in more modern terminology, that their convex hull is a triangle. How large is $P(K)$?, and for what K is $P(K)$ the largest and the smallest? This question became known as Sylvester's four-point problem. It took fifty years to find the answer: Blaschke [16] showed that for all convex bodies $K \subset R^2$

$$P(\text{disk}) \leq P(K) \leq P(\text{triangle}).$$

*Rényi Institute of Mathematics, Hungarian Academy of Sciences, PoB 127 Budapest, 1364 Hungary, and Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK. E-mail: barany@renyi.hu and barany@math.ucl.ac.uk

Assume now, more generally, that $X_n = \{x_1, \dots, x_n\}$ is a random sample of n uniform, independent points from the convex body K and write $p(n,K)$ for the probability that X_n is in "convex position", that is, no x_i is in the convex hull of the others. Sylvester's question is just the complenetary problem for $n = 4$: $P(K) = 1 - p(4,K)$. The probability $p(n,K)$ has been determined in various special cases [22],[17],[9],[27]. The following result from [7] describes the asymptotic behaviour of $p(n,K)$.

Theorem 1.1. *For every convex set $K \subset R^2$ of area one*

$$\lim n^2 \sqrt[n]{p(n,K)} = \frac{e^2}{4} A^3(K)$$

where $A(K)$ is the supremum of the affine perimeter of all convex sets $S \subset K$.

The affine perimeter, $AP(K)$ can be defined in many ways (see[23]), for instance $AP(K) = \int_{\partial K} \kappa^{1/3} ds$ where κ is the curvature and integration goes by arc-length. (This definition works for smooth convex bodies, the extension for all convex bodies can be found in [23].) Theorem 1 of [6] says that there is a unique convex compact set $K_0 \subset K$ with $AP(K_0) = A(K)$. The proof of Theorem 1.1 gives more than just the asymptotic behaviour of $p(n,K)$, namely, if the random points $x_1, \dots, x_n$ are in convex position, then their convex hull is, with high probability, very close to K_0. For the precise formulation see [7].

Next, define $Q(X_n)$ as the random collection of all convex polygons spanned by the points of X_n, that is, $P \in Q(X_n)$ iff $P = \text{conv}\{x_{i_1}, \dots, x_{i_k}\}$ for some k-tuple of points from X_n that is in convex position ($k \geq 3$). How many polygons are there in $Q(X_n)$? The answer is in [7]. Write $E|Q(X_n)|$ for the expectation of the size of $Q(X_n)$.

Theorem 1.2. *For every convex set $K \subset R^2$ of area one*

$$\lim n^{-1/3} \log E|Q(X_n)| = 3 \cdot 2^{-2/3} A(K).$$

Further, there is a limit shape to the polygonss in $Q(X_n)$, meaning that all but a small fraction of the polygons in $Q(X_n)$ are very close to K_0. We use $\delta(S,T)$ to denote the Haussdorf distance of $S, T \subset R^2$.

Theorem 1.3. *For every convex body $K \subset R^2$ and for every $\varepsilon > 0$*

$$\lim \frac{E|\{P \in Q(X_n) \colon \delta(P, K_0) > \varepsilon\}|}{E|Q(X_n)|} = 0.$$

The proofs of these results are too long to be presented in this survey, and the interested reader is adviced to consult [7]. We will see in section 3 that very similar phenomena hold for the lattice case.

2. Higher dimensions

Much less is known in higher dimensions. One reason is that the unicity of the convex subset of K with maximal affine surface area is not known. It is not known, for instance, which convex subset of the unit cube in R^3 has maximal affine surface arae. But there are other reasons as well, connected to the lack of the multiplicative rule (5.3) from [7]. Yet one can prove the following asymptotic formula [8]. Here $\mathcal{C}^d$ denotes the set of all convex bodies in R^d, and $p(n, K)$ stands for the same probability as before (just $K \in \mathcal{C}^d$), and $n \geq d + 1$.

Theorem 2.1. *For every $K \in \mathcal{C}^d$, and for all $n \geq n_0$*

$$c_1 < n^{2/d-1} \sqrt[n]{p(n,K)} < c_2$$

where n_0, c_1, c_2 are positive constants that depend only on d.

With Vinogradov's convenient $\ll$ or $\ll_d$ notation this says that

$$1 \ll_d n^{2/d-1} \sqrt[n]{p}(n,K) \ll_d 1$$

From this one can estimate the size of $E|Q(X_n)|$ when $K \in \mathcal{C}^d$:

$$n^{(d-1)/(d+1)} \ll_d \log E|Q(X_n)| \ll_d n^{(d-1)/(d+1)}.$$

At this point it is high time to look at lattice polyopes contained in some fixed $K \subset \mathcal{C}^d$. So let Z^d be the lattice of the integers in R^d and consider, for a large integer m, the lattice $\frac{1}{m}Z^d$. Assume K contains n points from this lattice. As m is large, $n = (1 + o(1))m^d \operatorname{Vol} K$. Write $\mathcal{P}_m(K)$ for the collection of all $\frac{1}{m}Z^d$-lattice polytopes contained in K. The next theorem, which follows easily from the results of [12], shows a very strong analogy between $\mathcal{P}_m(K)$ and $Q(X_n)$.

Theorem 2.2. *For every $K \in \mathcal{C}^d$*

$$n^{(d-1)/(d+1)} \ll_d \log|\mathcal{P}_m(K)| \ll_d n^{(d-1)/(d+1)}.$$

3. Lattice polygons and limit shape

In the plane Theorem 2.2. can be proved in stronger form (see [5] and [28]):

Theorem 3.1. *For every $K \in \mathcal{C}^2$*

$$\lim n^{-2/3} \log|\mathcal{P}_n(K)| = 3\sqrt[3]{\frac{\zeta(3)}{4\zeta(2)}A(K)}.$$

(Here $\zeta(.)$ stands for Riemann ζ function.) Note that this result is in complete analogy with Theorem 1.2: just the constant is different. (Also, n is in power $-2/3$ instead of $-1/3$ as K contains $(1 + o(1))n^2 \operatorname{Area} K$ lattice points.) The analogy carries over to Theorem 1.3 as well:

Theorem 3.2. *For every convex body* $K\mathcal{C}^2$ *and for every* $\varepsilon > 0$

$$\lim \frac{|\{P \in \mathcal{P}_n(K)) \colon \delta(P, K_0) > \varepsilon\}|}{|\mathcal{P}_n(K)|} = 0.$$

This shows again that all but a tiny fraction of the polygons in $\mathcal{P}_n(K)$ are very close to K_0. In other words, these polygons have a limit shape. Theorems of this type were first proved by Bárany [5], Vershik [28], and later by Sinai [25] (for the case when K is the unit square). Sinai's proof uses probability theory and gives a central limit theorems about how small that tiny fraction of polygons is. This has been generalized by Vershik and Zeitouni [29] for all convex bodies in R^2. A cental limit theorem of this type holds for the random sample case as well, see [14] for the precise statement.

4. The integer convex hull

The "integer convex hull", $I(K)$, of a convex body $K \in \mathcal{C}^d$ is, by definition, the convex hull of the lattice points contained in K:

$$I(K) = \operatorname{conv}(Z^d \cap K).$$

$I(K)$ is clearly a convex polytope. How many vertices does it have? Motivation for the question comes from different sources: integer programming, classical enumeration questions (like the circle problem), and from the theory of random polytopes. In integer programming one wants to know that $I(K)$ does not have too many vertices, assuming, say, that K is a nice rational polytope. The latter means that K can be given by m inequalities with integral coefficients; the size of such an inequality is the number of bits necessary to encode it as a binary string. Then the size of the rational polytope is the sum of the sizes of the defining inequalities. Strengthening earlier results by Shevchenko [24], and Hayes and Larman [21], Cook, Hartman, Kannan, and McDiarmid [18] showed that for a rational polytope K of size ϕ

$$f_0(I(K)) \le 2m^d(12d^2\phi)^{d-1}.$$

Here, as usual, $f_i(P)$ stands for the number of i-dimensional faces of the polytope P. Thus, in particular, $f_0(P)$ is the number of vertices of P. Most likely, the same inequality holds for all $i = 0, 1, \ldots, d-1$:

$$f_i(I(K)) \ll \phi^{d-1}$$

where the implied constant depends on d and m as well.

The above inequality for $f_0(I(K))$ is best possible, as is shown Bárány, Howe, and Lovász in [11]:

Theorem 4.1. *For fixed* $d \ge 2$ *and for every* $\phi > 0$ *there exists a rational simplex* $P \subset R^d$ *of size at most* ϕ *such that* $I(P)$ *has* $\gg_d \phi^{d-1}$ *vertices.*

The construction uses algebraic number theory. It shows further that the estimate $f_i(I(K)) \ll \phi^{d-1}$ for all i is best possible, if true.

What about the integer convex hull of other convex bodies? Balog and Bárány [3] considered case $K = rB^2$ where B^2 is the Euclidean unit ball centered at the origin and r is large and showed that

$$0.3r^{2/3} < f_0(I(rB^2)) < 5.5r^{2/3}.$$

Later Balog and Deshoullier [4] determined the average of $f_0(I(rB^2))$ on an interval $[R, R+H]$ which turned out to be very close to $3.453R^{2/3}$ as R goes to infinity (H has to be large). Bárány and Larman [13] determined the order of magnitude of $f_i(I(rB^d))$. (The method, and the result, apply not only for the unit ball but for smooth enough convex bodies as well.)

Theorem 4.2. *For every $d \geq 2$ and every $i = 0, 1, \ldots, d-1$*

$$r^{d(d-1)/(d+1)} \ll_d f_i(I(rB^d)) \ll_d r^{d(d-1)/(d+1)}.$$

This result is related to a beautiful theorem of G. E. Andrews [1] stating that a lattice polytope P in R^d with volume $V > 0$ has $\ll_d V^{(d-1)/(d+1)}$ vertices. The above theorem shows that Andrews' estimate is best possible (apart from the constant implied by $\ll_d$). A similar (perhaps less compact) example was given earlier V. I. Arnol'd [2].

This kind of question can be considered in a more general setting. Let G be the group of all isometries of R^d with translations by elements of Z^d factored out. G is a compact topological group with a Haar measure which is a unique probability measure when normalized properly. So we can talk about the expectation of the random variable $f_0(I(gK))$.

For the next result we assume $K \in \mathcal{C}^d$ and define the function $u\colon K \to R$ by

$$u(x) = \mathrm{Vol}(K \cap (2x - K)),$$

that is, $u(x)$ is the volume of the inetrsection of K with K reflected around x. Set, finally, $K(u < t) = \{x \in K\colon u(x) < t\}$. The following is an unpublished result of Bárány and Matoušek:

Theorem 4.3 *Consider all $K \in \mathcal{C}^d$ with the ratio of the radii of the smallest circumscribed and the largest inscribed balls to K buonded by D. Then, as* $\mathrm{Vol}\, K$ *goes to infinity,*

$$\mathrm{Vol}\, K(u < 1) \ll E f_0(I(gK)) \ll \mathrm{Vol}\, K(u < 1)$$

where the constants implied by $\ll$ depend only on d and D.

It follows easily from Minkowski's classical theorem that all vertices of $I(K)$ belong to $K(u < 2^d)$. (This is the first step in proving the upper bound.) It is not hard to see that $\mathrm{Vol}\, K(u < 2^d) \ll \mathrm{Vol}\, K(u < 1)$. So the meaning of the theorem is that the average number of vertices of $I(gK)$ is essentially the volume of $K(u < 2^d)$. Probably the same is true for the expected number of i-dimensional faces of $I(gK)$ but there is no proof in sight.

We mention further that Theorem 4.3 is quite analogous to a result in the theory of random polytopes. Given $K \in \mathcal{C}^d$, and a random sample of n points, X_n, from K, $K_n = \operatorname{conv} X_n$ is called a *random polytope* on n points. It is shown in [10] that, assuming $\operatorname{Vol} K = n$ (which is the proper scaling for comparison with Theorem 4.3), for all $i = 0, 1, \ldots, d-1$

$$\operatorname{Vol} K(u < 1) \ll Ef_i(K_n) \ll \operatorname{Vol} K(u < 1)$$

where the implied constants depend only on dimension.

The behaviour of $\operatorname{Vol} K(u < 1)$ is more or less known (from [10]): it is $(\operatorname{Vol} K)^{(d-1)/(d+1)}$ for smooth enough convex bodies and $(\log \operatorname{Vol} K)^{d-1}$ for polytopes, and it is between these bounds for all convex bodies.

5. Random 0-1 polytopes

Finally we mention a recent development, prompted by a question of K. Fukuda and G. M. Ziegler [30]. They asked how many facets a 0-1 polytope in R^d can have; a 0-1 polytope is a polytope whose vertices only have 0 or 1 coordinates. So such a polytope is the convex hull of a subset of the vertices of the unit cube, Q^d, in R^d. Write $G(d)$ for the maximal number of facets a 0-1 polytope can have. It is not hard to see that $2^d \le G(d) \le 2d!$. The upper and lower bounds have been improved slightly: the lower bound by a construction of Christoff (see [30]), and the upper bound by Fleiner, Kaibel, and Rote [20].

The vertices of every 0-1 polytope are on a sphere (centered at $(1/2, \ldots, 1/2)$). There is a formula for the expected number of facets of a random polytope with n uniform independent points from the (unit) sphere in R^d. It says that, in the range when $2d < n < 2^d$, the expected number of facets is of order $(\log n/d)^{d/2}$. So if the analogy between random points and lattice points carries over the 0-1 case one should expect $G(d)$ to be of order $d^{d/2}$. This is too much to ask for at the moment, yet the following is true (see [15]).

Theorem 5.1. *There is a constant $c > 0$ such that for all $d \ge 2$*

$$G(d) \gg \left(c \frac{d}{\log d}\right)^{d/4}.$$

The construction giving this estimate is random. Write K_n for the convex hull of n random, uniform, and independent 0-1 vectors. Assume x is a point from Q^d, and define

$$p(x, n) = \mathbb{P}[x \in K_n].$$

General principles would tell that, for most $x \in Q^d$, $p(x,n)$ is either close to one or close to zero. To be more specific, set

$$P(t) = \{x \in Q^d : p(x,n) \geq t\}.$$

The proof of Theorem 5.1 is based on the fact that for all small $\varepsilon > 0$ and large enough d $P(1-\varepsilon) \subset P(\varepsilon)$, of course, but the drop from $1-\varepsilon$ to ε is very abrupt: $P(\varepsilon)$ is in a small neighbourhood of $P(1-\varepsilon)$. This shows that $P(1-\varepsilon) \subset K_n$ with high probability. But only a tiny fraction of K_n lies outside $P(\varepsilon)$: most of the boundary of $P(\varepsilon)$ is outside K_n. Thus most of the boundary of $P(\varepsilon)$ is cut off by facets of K_n. These facets lie outside $P(1-\varepsilon)$. Comparing the surface area of $P(\varepsilon)$ with the amount a facet can cut off from it gives the lower bound.

The actual proof is technical, difficult, and makes extensive use a beautiful result of Dyer, Füredi, and McDiarmid [19]. Their target was to determine the threshold $n - n(d)$ such that K_n contains most of the volume of Q^d. As they prove, this happens at $n = (\frac{2}{\sqrt{e}})^d$. Their method describes where $p(x,n)$ drops from one to zero as $d \to \infty$. The analysis carries over for other values of n. In our case higher precision is required as we need a good estimate on how fast $p(x,n)$ drops from one to zero. We were able to control this only where the curvature of the boundary of $P(\varepsilon)$ behaves nicely. This is perhaps the spot where the exponent $d/2$ for the random spherical polytope) is lost and we only get $d/4$ for K_n.

References

[1] G. E. Andrews, A lower bound for the volume of strictly convex bodies with many boundary points, *Trans. Amer. Math. Society,* 106 (1963), 270-279.

[2] V. I. Arnol'd, Statistics of integral lattice polytopes (in Russian), *Funk. Anal. Pril,* 14 (1980), 1-3.

[3] A. Balog, I. Bárány, On the convex hull of the integer points in a disc, *Discrete Comp. Geometry,* 6 (1992), 39-44.

[4] A. Balog, J-M. Deshouliers,, On some convex lattice polytopes, in: *Number theory in progress,* (K. Győri, ed.), de Gruyter, 1999, pp. 591-606.

[5] I. Bárány, The limit shape of convex lattice polygons, *Discrete Comp. Geometry,* 13 (1995), 270-295.

[6] I. Bárány, Affine perimeter and limit shape, *J. Reine Ang. Mathematik,* 484 (1997), 71-84.

[7] I. Bárány, Sylvester's question: the probability that n points are in convex position, *Annals of Probability,* 27 (1999), 2020-2034.

[8] I. Bárány, A note on Sylvester's four point problem, *Studia Math. Hungarica,* 38 (2001), 73-77.

[9] I. Bárány, Z. Füredi, On the shape of the convex hull of random points, *Prob. Theory Rel. Fields,* 77 (1988), 231-240.

[10] I. Bárány, D. G. Larman, Convex bodies, economic cap coverings, random polytopes, *Mathematika,* 35 (1988), 274-291.

[11]I. Bárány, R. Howe. L. Lovász, On integer points in polyhedra: a lower bound, *Combinatorica,* 12 (1992), 135-142.
[12]I. Bárány, A. M. Vershik, On the number of convex lattice polytopes, *GAFA Journal,* 2 (1992), 381-393.
[13]I. Bárány, D. G. Larman, The convex hull of the integer points in a large ball, *Math. Annalen,* 312 (1998), 167-181.
[14]I. Bárány, G. Rote, W. Steiger, C. Zhang,, A central limit theorem for random convex chains, *Discrete Comp. Geometry,* 23 (2000), 35-50.
[15]I. Bárány, A. Pór, On 0-1 polytopes with many facets, *Advances in Math.,* (2000), 1-28.
[16]W.Blaschke, *Vorlesungen über Differenzialgeometrie II. Affine Differenzialgeometrie*, Springer, 1923.
[17]C. Buchta, On a conjecture of R.E. Miles about the convex hull of random points, *Monatsh. Math.,* 102 (1986), 91-102.
[18]W. Cook, M. Hartman, R. Kannan, C. McDiarmid, On integer points in polyhedra, *Combinatorica,* 12 (1992), 27-37.
[19]M. E. Dyer, Z. Füredi, C. McDiarmid, Volumes spanned by random points in the hypercube, *Random Structures and Algorithms,* **3** *(1992), 91–106.* 3 (1992), 91-106.
[20]T. Fleiner, V. Kaibel, G. Rote:, Upper bounds on the maximal number of facets of 0/1-polytopes, *European J. Combinatorics,* **21** *(2000), 121–130.* 21 (2000), 211-130.
[21]A. C. Hayes and D. G. Larman, The vertices of the knapsack polytope, *Discrete App. Math.,* 6 (1983), 135-138.
[22]B. Hostinsky, Sur les probabilités géométriques, *Publ. Fac. Sci. Univ. Brno,* (1925).
[23]E. Lutwak, Extended affine surface area, *Adv. Math.,* 85 (1991), 39-68.
[24]V. N. Shevchenko, On the number of extreme points in linear programming (in Russian), *Kibernetika,* 2 (1981), 133-134.
[25]Ya. G. Sinai, Probabilistic approach to analyse the statistics of convex polygonal curves (in Russian), *Funk. Anal. Pril.,* 28 (1994), 41-48.
[26]J. J. Sylvester, Problem 1491, *The Educational Times, (London)* (April 1864), 1-28.
[27]P. Valtr, The probability that n random points in a triangle are in convex position, *Combinatorica,* 16 (1996), 567-574.
[28]A. M. Vershik, The limit shape for convex lattice polygons and related topics (in Russian), *Funk. Anal. Appl.,* 28 (1994), 16-25.
[29]A. M. Vershik, O. Zeitouni, Large deviations in the geometry of convex lattice polygons, *Isreal J. Math.,* 109 (1999), 13-28.
[30]G. M. Ziegler, Lectures on 0/1 polytopes, in: *Polytopes — Combinatorics and Computation* (G. Kalai and G. M. Ziegler, eds.), DMV-Seminars,, Birkhäuser-Verlag, 2000, pp. 1-44.