ISOPERIMETRIC PROBLEMS IN DISCRETE SPACES

Levon Aslanyan

Abstract*: A substantial survey of results, obtained around the discrete isoperimetry problems, is presented in Harper's book of 2004. But a number of important results remain not reflected. These basic results, obtained in period 1980-90, and a number of new results obtained recently are included in this text. Results are presented in terms of geometry of multidimensional unit cube. And of course the considered in this paper issue concerns the vertex-isoperimetry paradigm vs. to the edge-isoperimetry version that is also a research topic in this area. The edge-isoperimetry counterpart of these topics will be surveyed in a separate publication.*

Keywords*: Data structures, discrete optimization, isoperimetry.*

ITHEA Keywords*: G.2 Discrete mathematics, G.2.3 Applications*

Introduction

The classical mathematical problem about isoperimetry is very old. The problem itself is linked to the name of queen Dido of Cartagena, ~800 BC [2]. Mathematical investigation in the area started in ancient time by Archimedes, Zenodot and others.

In modern terms the main isoperimetry result (on a plane) can be formulated as:

Theorem. Let C be a simple closed curve in the plane with length L and bounding a region of area A. Then$\mathrm{L}^2 \geq 4\pi\mathrm{A}$, with equality if and only if C is a circle.

$\mathrm{L}^2 \geq 4\pi\mathrm{A}$ is the so called the basic isoperimetry inequality for the plane. A similar relation holds for the 3 dimensional space. The basic isoperimetry theorem, historically, was proved in 2 stages. Firstly it was proven that if a curve satisfies the isoperimetry requirement, then it must be a circle. The proof is more or less simple and it can be found in Blyashke's book "Circle and ball". The part about the existence of a solution is harder and it was elaborated firstly by Karl Weierstrass, as part of the variation analysis that was appearing that time.

Applications of isoperimetry property could be found in different areas. Our goal in this paper is not the classical/continuous but the so called discrete isoperimetry problem.

Consider a text document and its word count histogram (the so called word vector model). The Armenian language, for example, involves approximately 10000 word roots (H. Adjarian, G. Jahukyan). Thus, such a document may be codded by a 10000-dimensional numerical vector and a collection of n documents corresponds to an n-subset of points of this integer vector space. These n vectors may be arranged as columns of a $10000 \times n$ matrix. Given a subset of vertices, to analyze its structural properties geometrical and algebraic approaches may be applied. The aim is to understand compactness, diversity of types, and other properties. This means, that we are interested to know the isoperimetry like properties, and our example is just to show that multidimensionality is a natural phenomenon of the discrete isoperimetry studies. Classical isoperimetry is two or three dimensional basically.

Consider one more example. When a new web page is created a natural question is to know the closest pages to it, that is the pages that contain a similar set of links, information or keywords. This question translates to the geometric question of finding nearest neighbors in a multidimensional multivalued space. This best match type query needs to be answered quickly. One of the ways to address this issue is to use the so called Random Projection Technique. If each web page is a d-dimensional vector, and d is very large, then instead of spending time d to read the vector in its entirety, once the random projection to a k-dimensional space is done, one needs only read k entries per vector. The use of isoperimetry in problem of best match search is demonstrated in [18,19].

Consider the n-dimensional unite cube E^{n}. A vertex $\alpha \in \mathrm{A} \subseteq \mathrm{E}^{\mathrm{n}}$ is called interior for subset A if $\mathrm{S}_1^{\mathrm{n}}(\alpha) \subseteq \mathrm{A}$, where $\mathrm{S}_{\mathrm{r}}^{\mathrm{n}}(\alpha)$ is the sphere of radius r centered at $\boldsymbol{\alpha}$. Denote by $\mathrm{I}(\mathrm{A})$ and $\mathrm{B}(\mathrm{A}) = \mathrm{A} - \mathrm{I}(\mathrm{A})$ the sets of all interior and boundary vertices of the set $\mathbf{A}$.

Basic Discrete Isoperimetry

Consider the revers lexicographical order and the standard order over the E^{n}. In reverse lexicographic order $1 < 0$. In standard order vertices are arranged layer by layer starting from layer $\mathbf{0}$. Inside the layers vertices are in reverse lexicographic order. L_m^n denotes the initial segment of length m in standard order over the E^{n}.

Theorem (Harper L.) L_m^n is the maximum internal vertex set for size m and dimension n, $1 \le m \le 2^n$.

The simplest proof of this theorem is based on the so called compression technique. We say that set A is i-compressed if partitions of A, $A^0(i)$ and $A^1(i)$ in direction i are in standard order of the $n-1$

dimension space. Arbitrary subset A, step by step can be partitioned and compressed by directions and after finite number of such steps we arrive to a uncompressible set A^c.

Theorem (Aslanyan L., Karakhanyan V., Torosyan B.) Compression does not enlarge the boundary set of vertices, and A^c either equals to L_m^n, or $m = 2^{n-1}$ and $A^c = L_{2^{n-1}-1}^n \cup \alpha_{2^{n-1}+1}$.

I. Leader, [3] "Discrete Isoperimetric Inequalities" (1991) refers to D. J. Kleitman [4] concerning the short proof of Isoperimetry that is factually based on compression. In fact, D. Kleitman presents just a survey in [4]. At that time simple proof was already published in Russian as well as in English [7].

Isoperimetry Inequalities

Isoperimetric inequalities are the key element of proofs in classical isoperimetric analysis.

To consider the discrete isoperimetry inequalities represent $a = |A|$ in the following canonical form

$$a = \sum_{i=0}^{k} \binom{n}{i} + \delta,\ 0 \le \delta < \binom{n}{k+1}.$$

The following isoperimetric inequalities are derived by Nigmatullin R. [5,6]:

I. $B(A) \supseteq B(A^0(i)) \cup B(A^1(i))$ for any $A \subseteq E^n$, and direction i,

II. $|B(A)| \ge |B(A^0(i))| + |A^1(i)| - |A^0(i)|$ for any $A \subseteq E^n$ of reduced form and direction i (A has reduced form, if $|A^1(i)| \ge |A|/2,\ i = 1, \cdots, n$),

III. If $A \subseteq E^n$, $|A| = m = \sum_{t=0}^{k} \binom{n}{t} + \delta$, $0 \le k \le n$ and $0 \le \delta < \binom{n}{k+1}$ then $\exists i \in \overline{1,n}$ such that $|A^1(i)| \ge \sum_{t=0}^{k-1} \binom{n-1}{t} + \delta \frac{k+1}{n}$.

Characterization of Solutions

Theorem (Aslanyan L., Karakhanyan V.) For each isoperimetric subset $A \subseteq E^n$ there exists a vertex $\alpha \in A$ such that $S_k^n(\alpha) \subseteq A$.

Subset A is called critical if $\forall\ \alpha \in A\ S_1^n(\alpha) \cap I(A) = \emptyset$. A number m is critical if L_1^n is a critical set. In this case it is easy to check that all m-element optimal subsets are critical.

Theorem (Aslanyan L., Karakhanyan V.) For each isoperimetric subset $A \subseteq E^n$ of critical cardinality there exists a vertex $\alpha \in A$ such that $S_k^n(\alpha) \subseteq A \subseteq S_{k+2}^n(\alpha)$.

m is a critical cardinality in dimension n iff the n-th entry of $\alpha_{L_m^n}$ equals 1. So the number of critical cardinalities equals 2^{n-1}.

Asymptotic Estimations

This point surveys the asymptotic behavior of discrete compactness when n→ ∞. For this we form sequences a_n of sizes of considered subsets in form $a_n = 2^{n-1}(1+\alpha_n)$. We require $-1 \le \alpha_n \le 1$ to satisfy the plain conditions $0 \le a_n \le 2^n$. Finally we denote $\mathfrak{M}_{a_n} = \{$ the set of sequences $A_n \subseteq E^n, |A_n| = a_n$, by $n = 1,2,\dots\}$ and let $\mathfrak{M} = \cup\, \mathfrak{M}_{a_n}$.

We call set $A \subseteq E^n$ γ-dense if $\gamma(A) = \frac{|I(A)|}{|A|} = \gamma, 0 \le \gamma \le 1$. Further denote by $m(\gamma)$ $(m_{a_n}(\gamma_n))$ the fraction of those subsets $A \subseteq \mathfrak{M}$ $(A_n \subseteq \mathfrak{M}_{a_n})$ which are γ_n dense. In a similar way $\widetilde{m}(\gamma)$ $(\widetilde{m}_{a_n}(\gamma_n))$ denotes the fraction of those subsets $A \subseteq \mathfrak{M}$ $(A_n \subseteq \mathfrak{M}_{a_n})$ with $\gamma(A) \ge \gamma$.

Theorem (Aslanyan L., Arsenyan I., Sahakyan H.)

A) $\lim_{n\to\infty} \gamma(A) = 0$ for almost all subsets $A \subseteq E^n$.For arbitrary integer k $\lim_{n\to\infty} m\left(\frac{k}{2^{n-1}}\right) = \frac{1}{2^k k!\sqrt{e}}$ and $\lim_{n\to\infty} \widetilde{m}(\frac{\eta(n)}{2^n}) = 0$ with $\eta(n) \to \infty$.

B) If $\lim_{n\to\infty} n\alpha(n) \to -\infty$ with $n \to \infty$, then for almost all subsets$A_n \subseteq \mathfrak{M}_{a_n}$ $\gamma(A_{n)} = 0.$

C) If $\lim_{n\to\infty} n\alpha(n) \to \infty$ with $n \to \infty$, then $m_a\left(\frac{1+\alpha(n)}{2}\right)^n \to 1$. If 1- $\alpha(n) \sim \frac{c}{n}$, $c > 0$, then for almost all $A_n \subseteq \mathfrak{M}_{a_n}$ $\gamma(A_{n)} \sim e^{-c/2}$. Moreover, $m_a(1) \to 1$ if 1- $\alpha(n) = o(1/n)$.

D) If $\lim_{n\to\infty} n\alpha(n) \to \infty$ then for almost all $B_n \subseteq \mathfrak{M}_{b_n}$, $|b_n - a_n| = o(\frac{2^n}{n})$ $\gamma(B) \sim \left(\frac{1+\alpha(n)}{2}\right)^n \sim \gamma(A)$.

E) If $n\alpha(n) \to \lambda$ with $n \to \infty$, then $m_a\left(\frac{k}{2^{n-1}}\right) \to \frac{1}{k!}\left(\frac{e^\lambda}{2}\right) e^{-\frac{e^\lambda}{2}}$, and $\widetilde{m}(\frac{\eta(n)}{2^{n-1}}) \to 0$ when $\eta(n) \to \infty$.

Theorem (Aslanyan L., Danoyan H.)

A) For arbitrary subset A the function $f_A(k) = \left|A^{(k)}\right|$ has one or two points of maximum over the domain {1,2,…,n} and in case when the maximum appears in two points, then they are neighbor points

B) Let we have an isoperimetric set S and any $A \subseteq E^n$, such that $|A| \ge |S|$. Then there exists an integer $\gamma \in \{0,1,\dots,n-1\}$ such that $\left|S^{(i)}\right| \le \left|A^{(i)}\right|$ when $i \le \gamma$ and $\left|S^{(i)}\right| \ge \left|A^{(i)}\right|$ when $i > \gamma$.

Conclusion

Studies in the discrete isoperimetry area appeared as issues of chip design and optimization, started in 70's, by a paper of L. Harper. Isoperimetry may vary concerning the real data structure considered and the set of constraint applied. The basic problem is in boundary minimization for the subsets of binary cubes. This paper considered the basic postulation and the aim was to conclude the results about the characterization of the set of all solutions of the problem. These results include postulations such as the existence of a maximal sphere in an arbitrary solution of the problem, the asymptotic description of the subset distribution by the volume of internal vertices, et. Participants of these researches are V. Karakhanyan, B. Torosyan, I. Arsenyan, H. Sahakyan, H. Danoyan. The extended narration of the topic is under preparation for further publication. This text will include theoretical as well as applied area descriptions such as chip design, cross connectivity networking and optimization, bioinformatical applications and others.

References

Harper L. H., Global Methods for Combinatorial Isoperimetric Problems, Volume 90 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2004.

Popular Lectures and Addresses by Sir William Thomson (Baron Kelvin) in Three Volumes, Vol. II, Geology and General Physics, Nature Series, London, MacMillan and Co., 1894.

B. Bollobás, Probabilistic Combinatorics and its Applications, Proceedings of Symposia in Applied Mathematics, Volume 44, 1991.

B. Bollobás, Surveys in Combinatorics, Cambridge University Press, 1979, Chapter 3 “Extremal hypergraph problems” by D. J. Kleitman.

Nigmatullin R. G., Some metric relations in the unite cube (in Russian), in Diskretny Analiz 9, Novosibirsk, 1967, pp. 47-58.

Nigmatullin R. G., Some metric relations in the unite cube (in Russian), in Diskretny Analiz 10, Novosibirsk, 1967, pp. 69-89.

Aslanyan L. H., Karakhanyan V. M. Torosyan B. E. On the compactness of Subsets of Vertices of the n-dimensional unit cube, in Soviet. Math. Dokl., American Mathematical Society, Vol. 19(1978), No. 4, pp. 781-785.

Асланян Л. А., Караханян В. М., Описание множества всех решений дискретной изопериметрической задачи, Доклады Академии Наук, Арм. ССР, 65(4), Ереван, 1977, стр. 211-215.

Асланян Л. А., Караханян В. М., Торосян Б. Е., Решение дискретной изопериметрической задачи, Доклады Академии Наук, Арм. ССР, 65(5), Ереван, 1977, стр. 257-262.

Асланян Л. А., Караханян В. М., Торосян Б. Е., О компактности подмножеств вершин n-мерного единичного куба, Доклады Академии Наук, СССР, 241(1), Москва, 1978 (English translation in Soviet. Math. Dokl., American Mathematical Society, Vol. 19(1978), No. 4, pp. 781-785), стр. 11-14.

Асланян Л. А., Изопериметрическая задача и смежные экстремальные задачи для дискретных пространств, Проблемы Кибернетики, 36, Москва, 1979, стр. 85-127.

Kleitman D. J., Extremal Hypergraph Problems, in London Math. Soc. Lect. Note Ser., Cambridge Univ. Press, No. 38(1979), pp. 44-65.

Bezrukov S. L., Isoperimetric Problems in Discrete Spaces, in: Extremal Problems for Finite Sets., Bolyai Soc. Math. Stud. 3, Budapest 1994, pp. 59–91.

Aslanyan L. A., Akopova I. A., On the distribution of the number of interior points in subsets of the n-dimensional unit cube, Colloquia Mathematika Societatis Yanos Bolyai, 37, Finite and Infinite Sets, (Eger) Hungary, 1981, pp. 47-58.

Асланян Л. А., Дискретная изопериметрическая проблема - асимптотический случай, Доклады Академии наук, Арм. ССР, 74 (3), Ереван, 1982, стр. 99-103.

Асланян Л. А., О единственности компактных булевых функций с точностью до изоморфизма, Working Paper, MTA SzTAKI, E/17 Budapest, 1983, pp. 1-11.

Aslanyan L., Metric decompositions and the Discrete Isoperimetry, IFAC Symposium on Manufacturing, Modelling, Management and Control, July 12-14 2000, Patra, Greece, pp. 433-438.

Aslanyan L. H., Danoyan H. E., On the Optimality of the Hash-Coding Type Nearest Neighbor Search Algorithm, Selected Revised Papers of 9th CSIT conference, submitted to IEEE Xplore, 2013.

Aslanyan L., Danoyan H., Complexity of Hash-Coding Type Search Algorithms with Perfect Codes, Journal of Next Generation Information Technology (JNIT), Volume 5, Number 4, November 30, 2014, ISSN : 2092-8637(Print), 2233-9388(Online), pp. 26-35.

Arsenyan I., Aslanyan L., Sahakyan H., Randomized Set Systems Constrained by Discrete Tomography, Information Theories and Applications, Vol. 23, Number 3, 2016, pp.203-231.

Authors' Information

Levon Aslanyan *– Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia, head of department; 1 P.Sevak str., Yerevan 0014, Armenia; e-mail: lasl@sci.am*

Major Fields of Scientific Research: Discrete analysis – algorithms and optimization, pattern recognition theory, information technologies.