

# Cryptology, isoperimetric problems and shadows

Gyula O.H. Katona

Alfréd Rényi Institute of Mathematics, Budapest

The practical problem is the following. Objects should be labeled with some geometric pictures. To avoid easy falsification, the pictures are chosen randomly. That is, a space  $S$  with a distance  $d$  and a measure  $\mu$  is given. The label of one object (picture) is a randomly chosen element of  $S$ . More precisely we will mark out some subsets  $A_i$  of  $S$  and if the random point falls in  $A_i$  then it can be used as a label of an object numbered  $i$ . The sets  $A_i$  must satisfy certain properties as described below.

If  $A \subset S, 0 < \varepsilon$  then define  $n(A, \varepsilon) = \{x \in S : d(A, x) \leq \varepsilon\}$ . The family of subsets  $A_1, \dots, A_m$  is called a *geometric code* with parameters  $\varepsilon, \rho$  and  $\alpha$  if (1)  $\mu(A_i) \leq \rho$  holds for every  $i$  ( $1 \leq i \leq m$ ), (2) the sets  $n(A_i, \varepsilon)$  ( $1 \leq i \leq m$ ) are pairwise disjoint, (3)  $\alpha \leq \mu(\cup A_i)/\mu(S)$ .

A geometric code can be applied in the following way. Choose random elements of  $S$  according to  $\mu$ . If  $x \in A_i$  then let its code  $c(x)$  be the binary form of  $i$ . On the other hand, if  $x \in A_i$  holds for no  $i$  then  $x$  is a waste. (1) ensures that a random choice of  $c(x)$  (not knowing  $x$ ) reproduce it with a small probability. (2) implies that reading  $x$  with an error at most  $\varepsilon$   $c(x)$  still can be recovered. Finally assumption (3) is needed to lowerbound the probability of the waste. The problem is to find the maximum of  $m$ , given  $S, \varepsilon, \rho$ , and  $\alpha$ . We give an inequality what has to be satisfied among these parameters, supposing that (1) the measure of a ball is not changed by moving its center, (2) the space satisfies the Brunn-Minkowski inequality.

In our implementation a label is a rectangle containing many small circles (they have a three-dimensional nature, this is why it is hard to copy them), what can be represented by their centers. Because of the computational approximation, it can be supposed that these centers are elements of a grid in the rectangle. Therefore an element of the space  $S$  is a subset of the set

of the points of the grid where the sizes of the subsets are between a lower and an upper bound. From practical experiences we know that some of the points can be "lost" during the control, therefore the distance  $d$  should be defined accordingly. One  $A_i$  is therefore a family of subsets of the grid points. Roughly speaking the largest number of such families should be found in such a way that deleting a small number of points from one member of a family is different from a subset obtained by deleting the points from the member of another family. This leads to the usage of the theory of extremal problems of finite sets, especially the "shadow theory".

Let us illustrate the problem in a very-very special case. Choose two families,  $\mathcal{A}$  and  $\mathcal{B}$  of 3-element subsets of an  $n$ -element set in such a way, that deleting one element from an  $A \in \mathcal{A}$  and one element from  $B \in \mathcal{B}$ , the so obtained two-element sets are different. Determine  $\max \min\{|\mathcal{A}|, |\mathcal{B}|\}$ . The complete asymptotic solution of this "easy-looking" problem will be presented.