

Segment Counting versus Prime Counting in the Ulam Square

Leszek J Chmielewski, Arkadiusz Orłowski, and Grzegorz Gawdzik

Faculty of Applied Informatics and Mathematics – WZIM,

Warsaw University of Life Sciences – SGGW

ul. Nowoursynowska 159, 02-775 Warsaw, Poland

{leszek_chmielewski,arkadiusz_orlowski}@sggw.pl

<http://www.wzim.sggw.pl>

Abstract. Points that correspond to prime numbers in the Ulam square form straight line segments of various lengths. It is shown that the relation between the number of the segments and the number of primes present in a given square steadily grows together with the growing values of the prime numbers in the studied range up to 25 009 991 and is close to double log. These observations were tested also on random dot images to see if the findings were not a result of merely the abundance of data. In random images the densities of the longer segments and the lengths of the longest ones are considerably smaller than those in the Ulam square, while for the shorter segments it is the opposite. This could lead to a cautious presumption that the structure of the set of primes might contain long-range relations which do not depend on scale.

Keywords: prime counting function, pi function, Ulam spiral, Ulam square, line segments, segment counting, Hough transform

1 Introduction

The Ulam spiral has been the object of interest as one of the tools of study on the fascinating set of prime numbers since its first presentation in [7]. It presents the set of the prime numbers as a two-dimensional pattern [9], so it has made it possible to investigate some of the properties of this set with the use of the image processing methods. The Ulam spiral will be considered as the part of the square in which it is embedded, which will be called the Ulam square. This square will be considered as an image having its pixels, sometimes called points.

In our previous papers [3,4,6] we have presented an image processing approach to the analysis of the regular line patterns present in the Ulam spiral, with use of the Hough transform. In [3] we have proposed a variant of the Hough transform capable of detecting the sequences of points which form regular segments of straight lines. As it is in the Ulam square, the slope of these lines is defined by a ratio of co-prime numbers. In [4] we have shown in more detail which segments are actually found and what is the distribution of their lengths.

In [6] we have investigated the directional structure of the Ulam spiral. We have shown that the image of the spiral indeed has some strongly preferred directions and that the general directionality is far from symmetrical with respect to the vertical and horizontal axes. This structure differs much from that of a random dot image of similar dot density, which has a different relation of dominating directions and has the symmetries which the Ulam spiral does not have.

Long segments necessitate for a more ordered structure than the short ones. It is expected that the longer segments will be less numerous than the shorter ones. In [4] we have shown that in the squares considered the longest segment had 16 points. Interestingly, there were no segments of lengths 15 and 14 points. There is only one segment of length 16, one of length 13, two of 12, and the number of shorter segments grows rapidly with the decrease of their length. The question arises what is the relation of the number of segments of various lengths formed in the Ulam spiral containing numbers less than n to the function $\pi(n)$, that is, the number of primes less than or equal to n (cf. for example [1,8]). Other questions also arise, for example, will the just stated limits for segment length be exceeded, and if so, for how large numbers? In the present paper we concentrate on the relation of the numbers of segments to the number $\pi(n)$.

The next parts of this paper are organized as follows. In Sect. 2 the method of finding line segments, described in detail elsewhere, is briefly reminded. In Sect. 3 the results obtained for the Ulam square are presented. In Sect. 4 they are compared to those obtained for dot images created by randomizing the Ulam square, to see if these results are specific for the Ulam square or are typical for any unstructured dot images. The paper is concluded in the last Section.

2 Finding line segments

The method of finding line segments described in [3] is illustrated in Fig. 1 where the central part of the Ulam spiral of dimensions 21×21 is shown. This spiral contains 85 primes, of which the largest one is 439. Grey and color pixels correspond to primes written in black. Black pixels correspond to composite numbers written in grey. The origin of the coordinate system O_pq is located in the center of the spiral corresponding to the unit 1. The image presents three segments of length 5 and one of length 6. There are no more segments of these and larger lengths in this image, but there are 21 segments of length 4, 110 segments of length 3 and 515 segments of length 2, not displayed in this image.

The slope of a line is defined by numbers $\Delta p, \Delta q$ denoted here $\Delta p = i, \Delta q = j$, limited to $i \in [0 : N]$ and $j \in [-N, N] \setminus \{0\}$. In the current implementation, for practical reasons, it is $N = 10$. This defines the dimensions of the accumulator array D_{ij} in the Hough transform, called the *direction array* (see [3] for the detailed description of the method). During the evidence accumulation stage of the line detection algorithm, which can be viewed as the process of voting for straight lines, the elemental voting subset is a pair of pixels representing a pair of primes. If the slope represented by the pair corresponds to an element of D , the pair is stored in D_{ij} , where i/j is the reduced fraction $\Delta p/\Delta q$; otherwise it

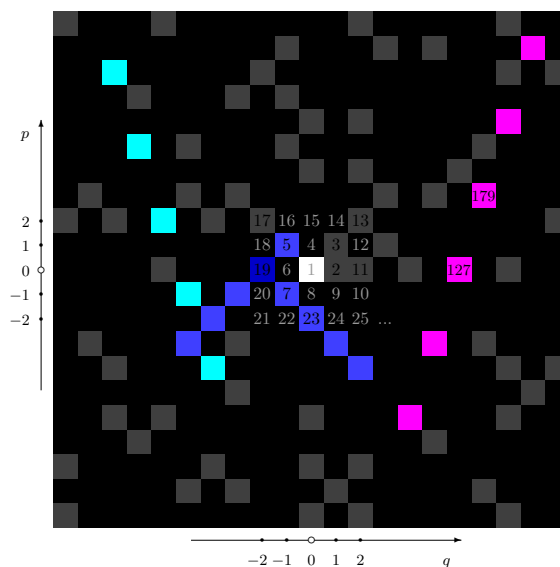


Fig. 1. Central part of the Ulam spiral of dimensions 21×21 and with segments of lengths 5 and 6 marked in colors. Notations: (p, q) – coordinates with origin in the first number of the spiral; primes: black on grey or color background; other numbers: grey on black or white. **Blue:** two segments of length 5, slopes $(\Delta p, \Delta q) = (1, 1)$ and $(1, -1)$, respectively, with one common point corresponding to prime number 19 marked with **darker blue**. **Cyan:** segment of length 5, slope $(\Delta p, \Delta q) = (3, -1)$. **Magenta:** segment of length 6, slope $(3, 1)$ (color images in the electronic edition).

is neglected. To store the votes, in each D_{ij} a one-dimensional data structure is formed (a vector or a list, whichever is more convenient). For each vote the line offset, defined as the intercept with the axis Op for horizontal lines and with Oq for the remaining ones, is stored. Additionally, the locations of voting points are stored and the pairs and their primes are counted.

After the accumulation process the accumulator can be analyzed in various ways. In [4] we were interested in finding the numbers of contiguous segments of different lengths. In [6] we have analyzed the directional properties of the segments in the Ulam spiral. The accumulated results make it possible to look for other regularities in the set of segments present in the Ulam spiral.

Assume the square is $S \times S$, so it contains $P = \pi(S^2)$ primes. The complexity of the accumulation is $O(PN^2)$ as it needs the search in neighborhood proportional to $N \times N$ around each prime. Complexity of the analysis depends on a number proportional to the number of directions N^2 , number of offsets S and needs sorting up to S points with quick sort, giving $O(N^2 S^2 \log(S))$. Time requirements are not crucial, as the analysis is made only once. The results presented here required several minutes to complete on a general-purpose PC. Memory requirements are $O(N^2 S^2)$ and can be considerable.

In the present analysis and also in the previous papers we have looked for contiguous segments, in the sense corresponding to the way the slope is represented in the direction array D . Let us consider an example of pairs of pixels $P_k(p, q)$: $P_1(-1, -1)$, $P_2(-2, 0)$ representing the primes 7, 23, and $P_3(0, 6)$, $P_4(3, 7)$ representing the primes 127, 179 (see Fig. 1). The pixels P_1, P_2 have $(\Delta p, \Delta q) = (i, j) = (1, 1)$ and are neighbors in the classical sense as well as in the sense of fitting to a direction conforming with D . The pixels P_3, P_4 have $(i, j) = (3, 1)$ and are neighbors in the sense of D , but not in the classical sense. However, they will be considered neighbors, as for this direction their distance is minimum. This explains in what sense the contiguity is considered here.

The fact that only the contiguous segments are considered here explains, among others, why in Fig. 1 there are only 515 segments of length 2, and not all the possible pairs of 85 primes, which would be $85 \times 84 / 2 = 3570$.

It is due to practical reasons that at present we have restricted our investigations to the central part of the Ulam squares of size up to 5001×5001 points, which contains 1 566 540 primes, the largest of which is 25 009 991. This is very little in comparison to the largest primes known at present, with the number of digits exceeding 20 millions [2]. However, it is only the question of time that this purely secondary limit will be overcome.

3 Results for the Ulam square

Let us now come back to the question of our interest in this paper, that is, what is the relation between the number of segments of various lengths and the number of primes in the Ulam square? First, let us compare the number of segments to the prime counting function $\pi(n)$. This relation is shown in Fig. 2.

The graph of the relation of the number of segments to the largest prime number has a similar shape as the prime counting function $\pi(n)$. In the double log scale it is very close to linear, although some discrepancies from the linear shape can be seen with an unarmored eye for all segments having more than four points. For the longest segments, the number of which is small, the relation is less regular. It is striking that for the smallest square considered, that of size 51×51 , the number of three-point straight segments is larger than the number of primes on which these segments are spanned. Actually, there are 378 primes and 411 three-point segments. The fact that the number of two-point segments is larger than the number of primes is not disappointing.

From now on we shall focus on the relations of the number of segments and the number of primes. This relation is shown in Fig. 3 for all the lengths of segments present in the images. It can be seen that this relation is always close to linear in the double log scale, provided that the number of segments involved is large. The longest segments are rare and for them the values differ not only quantitatively, but also qualitatively with the growth of the number of primes, so a steady relation can not be expected in the range of their large variability.

It is interesting to see that the relation is steady throughout the whole investigated range of the prime numbers, and its nature is the same for small as

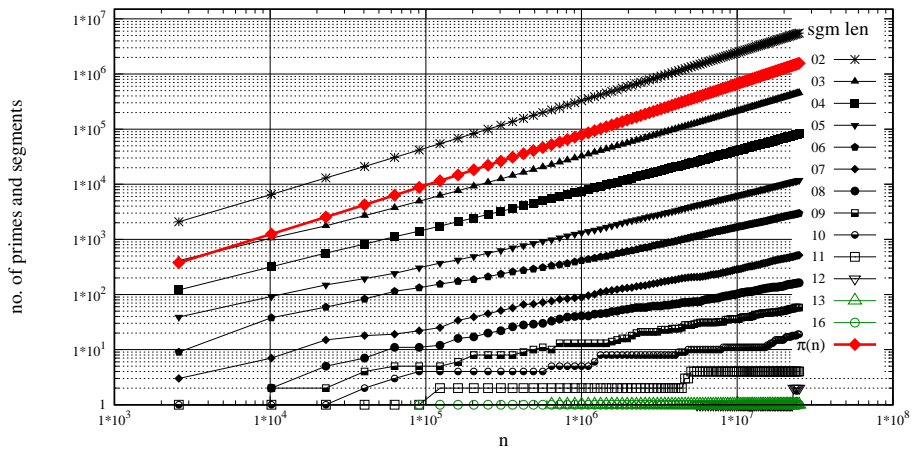


Fig. 2. Number of primes less than or equal to n , that is, the $\pi(n)$ function, and the numbers of segments of lengths encountered in the images considered, versus n . If the number of segments is zero the data point is absent. Color used to enhance discernibility.

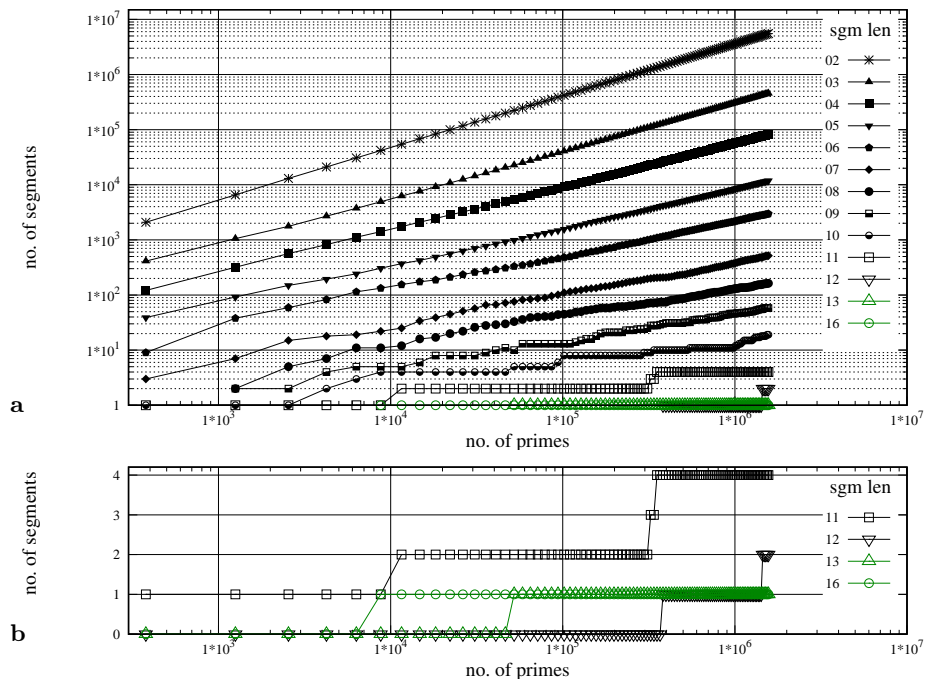


Fig. 3. Number of segments versus number of primes in the square. (a) Results for all the segments; (b) results for the longest segments.

well as for large values of the largest prime. The primes become rare together with the growth of their value n , in the sense that the distance between them increases with n . Such structures as regular linear patterns necessitate for not only the presence of the primes, but also the existence of some constant relations between them, which is a more restrictive or demanding condition than the presence itself. So, it might well be expected that at some point the number of regular structures could fall down. No such phenomenon is observed in the range of primes investigated until now. A modest believe can be formulated that the structure of the set of the prime numbers might contain the regularities or relation which do not depend on the scale and span the set of the prime numbers in its entirety.

4 Comparison with a random square

It is important to find out to what extent the regularities presented above are specific for the Ulam square itself, or are related to that in nearly every image, even that formed randomly, some regularities can be found, as a consequence of the abundance of data. We shall consider the dot images having the density of dots very close to that observed in the Ulam square, but displaced randomly.

The way of randomizing the Ulam square used in this study will be more sophisticated than that applied in [6], where in a square the dots were displaced with uniform probability density, while their number was kept equal to that in the Ulam square of equal size. However, actually the density of primes decreases with the distance from the center, as fixed by the random counting function $\pi(n)$. The randomization method should account for this. In other words, in the randomization process the prime numbers should be changed locally.

Here it has been done in the following way. Let us form a vector V_i of integers from 1 to the number M of pixels in the square. It represents the new set of numbers which will replace the set of primes. The elements of V are initially empty. Let us take three subsequent primes, p_l , p_m and p_r . Now, let us draw a random number p'_m belonging to the closed interval $[p_l, p_r]$, with uniform probability density. This number is saved in V in the element with index p'_m , so $V_{p'_m} := p'_m$; but if $V_{p'_m}$ was not empty, another number is drawn. In this way, p'_m has replaced p_m in the set. This process is performed for all triplets of subsequent primes, from the smallest to the largest ones. The boundary elements of V require attention. In the first triplet of primes, p_l is replaced by 1, so $p_m = 2$ and $p_r = 3$. Therefore, p'_m can equally probably become 1, 2 or 3. In the last triplet, p_r is replaced by M . Finally, the set represented by the nonempty elements of V is used instead of prime numbers to fill the now randomized Ulam square. The proposed randomization method makes the primes move only in the direction according to the way the numbers are written down into the Ulam spiral. One realization of the random image is shown in Fig. 4 and two others, as a side effect of other visualization, in Fig. 6.

The calculations in this study are repeated for the squares of growing sizes. To keep the segments in their place during the calculations, when the size of

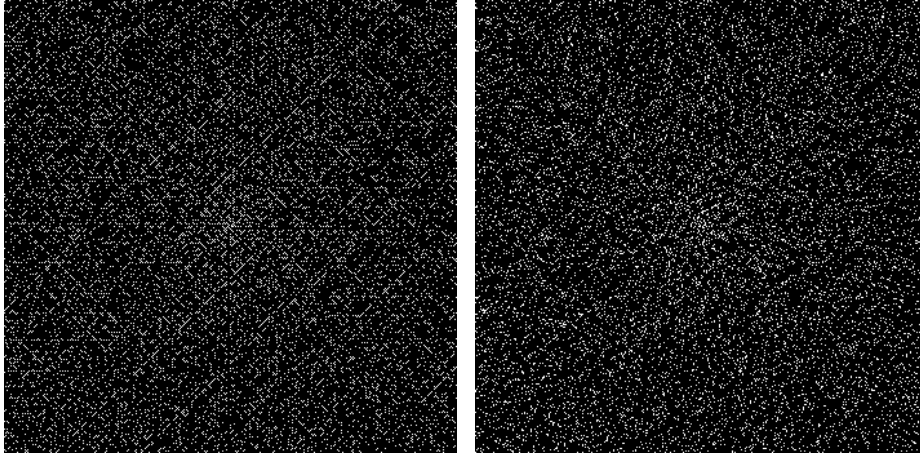


Fig. 4. The actual Ulam square of size 351×351 (left) and one of its randomized realizations (right). In each image there are 123 201 pixels, of which 11 578 represent primes. The largest prime is 123 176. The structure of the randomized square differs from that of the actual one even in visual inspection.

the square is increased, only the newly added pixels are drawn in this random process, and the middle ones remain.

The result for randomized data can be seen in Fig. 5a. To make it easier to compare the graphs to those of Fig. 3, they are mixed together in one plot.

The graphs received for the Ulam square and those for the corresponding random square are different. Among the most conspicuous differences, the following two can be pointed out.

1. In the Ulam square the longest segment contains 16 points, while in the considered realization of the random square it has 8 points. This suggests that in the Ulam square the regularity is stronger than that in a random square.
2. In the random square the numbers of the shorter segments are larger than those in the Ulam square – this holds for 2 and 3-point segments, while the numbers of the longer segments is smaller – this holds for the 4, 5, 6-point and longer segments. These relations do not depend on the size of the squares. This seems to suggest the existence of some factor which promotes the appearance of longer structures in the Ulam square.

It should be stressed that the graphs in Fig. 5a were found for a single realization of the randomization process of the square. To check how the results change from one realization to another, we have run the program for the same square of a specified size ten times each. The numbers of points were chosen to place the data in a possibly even way on the horizontal axis of the graph. These were the smallest sizes for which the number of primes exceeded 200, 1 000, 2 000, 10 000, 20 000, 100 000, 200 000, 1 000 000, which occurred for squares of size 51, 101, 151, 351, 501, 1 151, 1 701 and 3 951. The results are shown in Fig. 5b. It

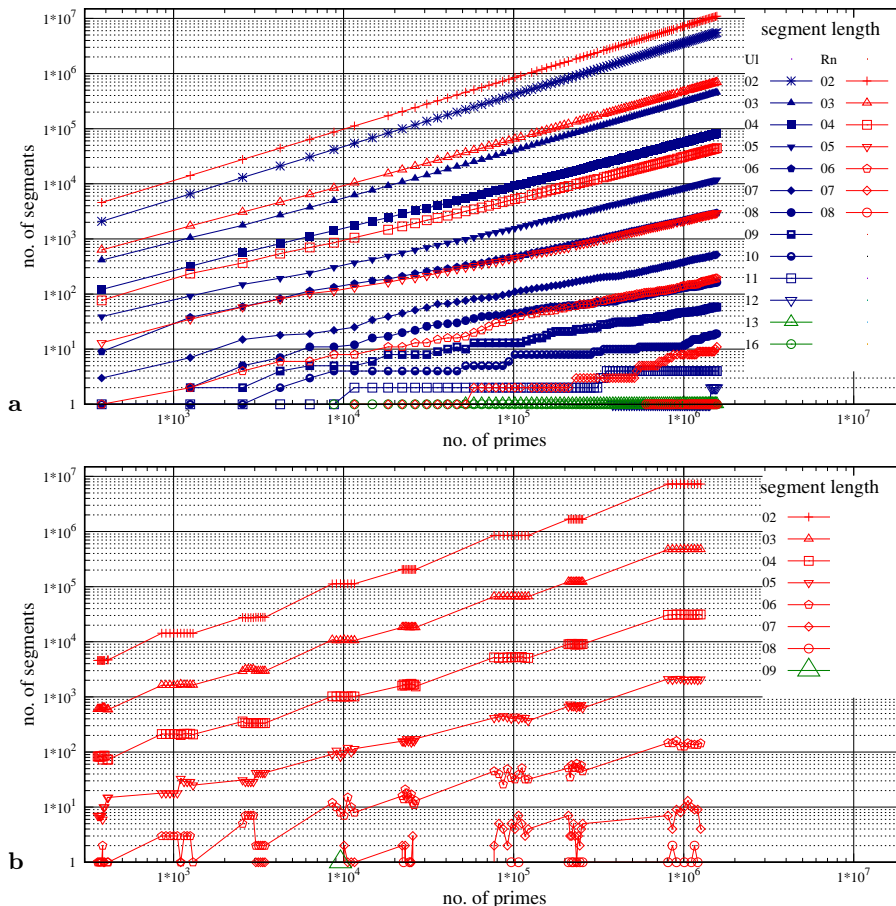


Fig. 5. (a) Number of segments vs. number of primes in the Ulam square (UI: thick blue or green lines, full symbols for lengths present also in random squares) and in random square (Rn: thin red lines, empty symbols). (b) Fragments of graphs of Fig. a for random data: results for 10 realizations of the random process (dithered horizontally for better visibility). Note a single 9-point segment near no. of primes = 10^4 .

can be seen that the longer and hence the less numerous the segments, the more variable the results. This could be expected, due to that the variabilities of the detailed results for shorter segments are muted in the summing processes. For shorter segments the differences can be of the same order of magnitude as the values themselves (segments of lengths 6 and more), but the location of their graphs is much lower than of those of the actual Ulam square. Therefore, this variability does not change the relations between the graphs in a qualitative way.

It should be noted that the results are stable for shorter segments in randomized squares, besides that the locations of the segments in the realizations are entirely different. This has been shown in Fig. 6 using the example of two random calculations for the square 351×351 which are small enough to keep the

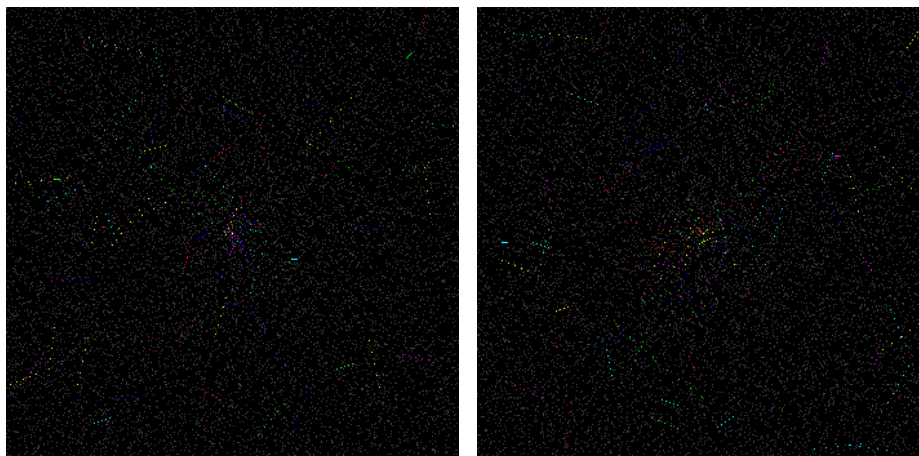


Fig. 6. Segments of length 5 points in two randomized calculations for the square 351×351 . Between 1000 and 1040 (approx.) such segments are present in an image. Colors were altered in a loop: {R, G, B, C, M, Y}. The realizations shown here are different from that in the right part of Fig. 4.

details visible. Segments of length 5 form different patterns in these realizations. For other lengths the results are similar.

In one of the realizations a single segment of length 9 points appeared in the square of size 1151 and its existence is shown in Fig. 5b with a large triangle. This indicates that the limit of 8 for the segment length which appeared in the calculations summarized in Fig. 5a is not general.

5 Summary and prospects

The points which correspond to prime numbers in the Ulam square form numerous straight line segments of various lengths. The question of what is the relation between the number of such segments and the number of primes has been studied. It has been demonstrated that the relation between the number of segments of a specified length and the value of the largest prime involved in forming these segments is double *log*. This has been demonstrated within the range of numbers up to 25 010 001, which form the square of 5001×5001 . The increase of the number of line segments with the growing values of the prime numbers has been observed, irrespective of that the density of primes in the set of integers goes down. It can be cautiously presumed that the structure of the set of the prime numbers might contain the regularities which do not depend on the scale and span the whole set of these numbers.

Very long segments which are rare in the Ulam square, like single segments of 16 and 13 points, two segments of 12 points, and a relatively small number of other longer segments do not follow the same tendency as the shorter segments, due to their small number.

It has been checked if the regularities found are actually specific for the Ulam square itself, or are typical for dot images. It is known that even in random dot images some regularities can be found due to the abundance of data. As a result, it has been found out that in the Ulam square the number of long segments is considerably larger than that in random dot images, and that the longest segments observed in the Ulam square are absent in them. However, the shorter segments, namely the ones of lengths 2 and 3, are more numerous in the random images than in the Ulam square. This suggests that the Ulam square is an object in which the relations between the points are more long-range than those in random dot images.

This research was carried out in hope that the application of the image processing techniques to the Ulam spiral will make it possible to broaden the set of tools and methods used in the analysis of the properties of the fascinating set of prime numbers.

The data used in this and our other papers related to the prime numbers and the detailed results in a downloadable form can be found in the web page [5].

References

1. C. K. Caldwell. How many primes are there? In *The Prime Pages* [2]. primes.utm.edu/howmany.html [Online; accessed 15 Oct 2016].
2. C. K. Caldwell, editor. *The Prime Pages*, 2016. primes.utm.edu [Online; accessed 15 Oct 2016].
3. L. J. Chmielewski and A. Orłowski. Hough transform for lines with slope defined by a pair of co-primes. *Machine Graphics & Vision*, 22(1/4):17–25, 2013. [Open access: mgv.wzim.sggw.pl/MGV22.html].
4. L. J. Chmielewski and A. Orłowski. Finding line segments in the Ulam square with the Hough transform. In L. J. Chmielewski et al., editors, *Computer Vision and Graphics: Proc. Int. Conf. ICCVG 2016*, volume 9972 of *Lecture Notes in Computer Science*, pages 617–626, Warsaw, Poland, 19–21 Sep 2016. Springer. doi:10.1007/978-3-319-46418-3_55.
5. L. J. Chmielewski and A. Orłowski. Prime numbers in the Ulam square, 2016. www.lchmiel.pl/primes [Online; accessed 14 Oct 2016].
6. L. J. Chmielewski, A. Orłowski, and M. Janowicz. A study on directionality in the Ulam square with the use of the Hough transform. In S. Kobayashi, A. Piegat, et al., editors, *Hard and Soft Computing for Artificial Intelligence, Multimedia and Security: Proc. Conf. ACS 2016*, volume 534 of *Advances in Intelligent Systems and Computing*, pages 81–90, Międzyzdroje, Poland, 19–21 Oct 2016. Springer. doi:10.1007/978-3-319-48429-7_8.
7. M. L. Stein, S. M. Ulam, and M. B. Wells. A visual display of some properties of the distribution of primes. *The American Mathematical Monthly*, 71(5):516–520, 1964. doi:10.2307/2312588.
8. E. W. Weisstein. Prime counting function. From MathWorld—A Wolfram Web Resource, 2016. mathworld.wolfram.com/PrimeCountingFunction.html [Online; accessed 15 Oct 2016].
9. E. W. Weisstein. Prime spiral. From MathWorld—A Wolfram Web Resource, 2016. mathworld.wolfram.com/PrimeSpiral.html [Online; accessed 15 Oct 2016].