UNIVERSITY OF HELSINKI
DEPARTMENT OF PHYSICS

BASICS OF MONTE CARLO SIMULATIONS

Exercise 1

Student: Caike Crepaldi

Professor: Flyura Djurabekova

January 2016

Note

Use the make file for the compilation commands. By using the command **make all** you can compile all programs. In problems like problem 4, in which I am asked to do the same thing for all RNG, there are several commands commented in the code with the indication of which command is used for each generator (so I can avoid creating one program for each version of the code). If you compile the code you will get the output for one of the RNG.

Problem 1

The situation described in the problem (where we deal with a needle and a striped floor) is the Buffon's needle problem. Thanks to the 1st lecture (as seen in the slides) we know that in the Buffon's needle problem we have the following relation:

$$P_{hit} = \frac{N_{hit}}{N_{total}} = \frac{2l}{\pi d} \tag{1}$$

Where:

- P_{hit} is the probability that the needle falls across a line;
- N_{hit} is the number of times the needle falls across a line;
- N_{total} is the total number of throws;
- l is the size of the needle;
- d is the distance between the lines;

In order to estimate the value of pi (this estimate value we will call e_π) we can use the values of the measurements l and d, obtained from the size of the needle and the distance between the lines, as seen below:

$$e_\pi = N_{total}\frac{2l}{N_{hit}d} \tag{2}$$

Therefore, by executing, or simulating, this experiment we can estimate the value of pi. For a better estimate value we need to do the experiment several times (in order to get a really big N_{total} and N_{hit}).

Now we need to show that the probability that the needle falls across a line is:

$$P_{hit} = \frac{2l}{\pi d}$$

As discussed initially in class, this probability is the probability of the center of the needle falling within $l/2$ units from either side of the strips AND falling in a certain angle that ensures that the needle ”touches” the line. Thus this probability is the combination (by multiplication) of those two other probabilities.

The first probability P_1 of a needle falling with its center close enogh from the a line is:

$$P_1 = \left(\frac{l}{2} + \frac{l}{2}\right)\frac{1}{d} = \frac{l}{d}$$

Note that we use $\frac{l}{2} + \frac{l}{2}$ because there are 2 lines in considering a single wooden strip, so the needle can fall in one line OR in the other.

The second probability depends on the geometry of the problem. Because of such I will use a image from the wikipedia article to ilustrate the situation.

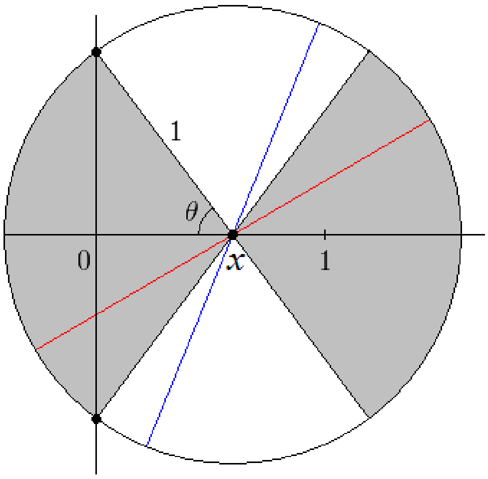

Figure 1: Buffon’s needle position.

Using the image 1, we can see that the angle (θ) that allows the needle to touch the line can be defined using the distance of the center of the needle from the line (called x). We can notice, using the figure, that

$$\theta(x) = \cos^{-1}(x)$$

where $\cos^{-1}$ is the arccosine. Note that this geometry assumes l=2 (see the figure).

Remember that the needle must fall within the range of 2θ out π rad possible configurations. The 2θ range is shown in the figure as the grey area within the circle.

Now we can calculate the second probability by integrating $2\theta(x)/\pi$ from 0 to l/2.

$$P_2 = \int_0^{l/2} \frac{2\theta(x)}{\pi}\, dx = \int_0^{l/2} \frac{2\cos^{-1}(x)}{\pi}\, dx \tag{3}$$

Like before, in order to simplify the calculation, we can assume l=2 (as shown in the figure).

$$P_2 = \frac{2}{\pi}\int_0^1 \cos^{-1}(x)\, dx = \frac{2}{\pi} \tag{4}$$

Therefore the probability that the needle falls across a line is:

$$P_{hit} = P_1 \times P_2 = \frac{l}{d}\frac{2}{\pi} \tag{5}$$

Problem 2

See the output of the code below.

Command 1

```
$ ./ex1p2
LCG - seed: 42   r1:   0.8759805
PM - seed: 123   r2:   0.0009626
MT - seed: 666,  r3:   0.6309263
```

Problem 3

Command 2

```
$ ./ex1p3
 LCG - seed:               10850665   r1:   0.397790879
 Period:
     12386880
 PMG - seed:               10850665   r2:   0.921311021
 Period:
  2147483646
 MTG - seed:     10850665   r3:   0.220242187
 Period:
      2360284
```

We can see that the LCG and PMG random number generators have a period close to the modulus m (the period is m or m-1), as expected by the theory.

Problem 4

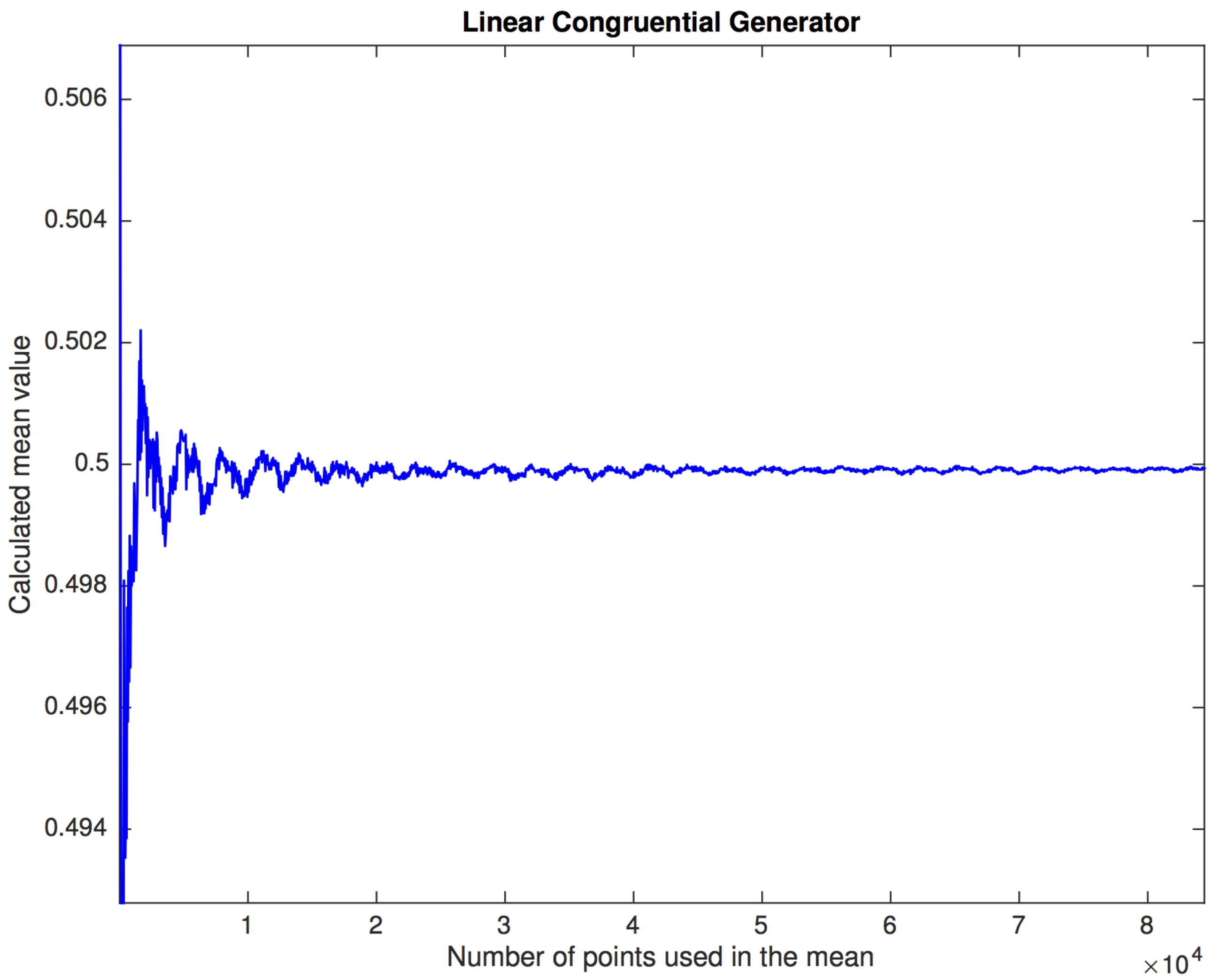

Figure 2: Linear congruential generator mean evolution analysis.

The fluctuations of the mean value seems reach the 0.0001 tolerance after 6×10^4 points in sequence (roughly). See the graph in figure 2.

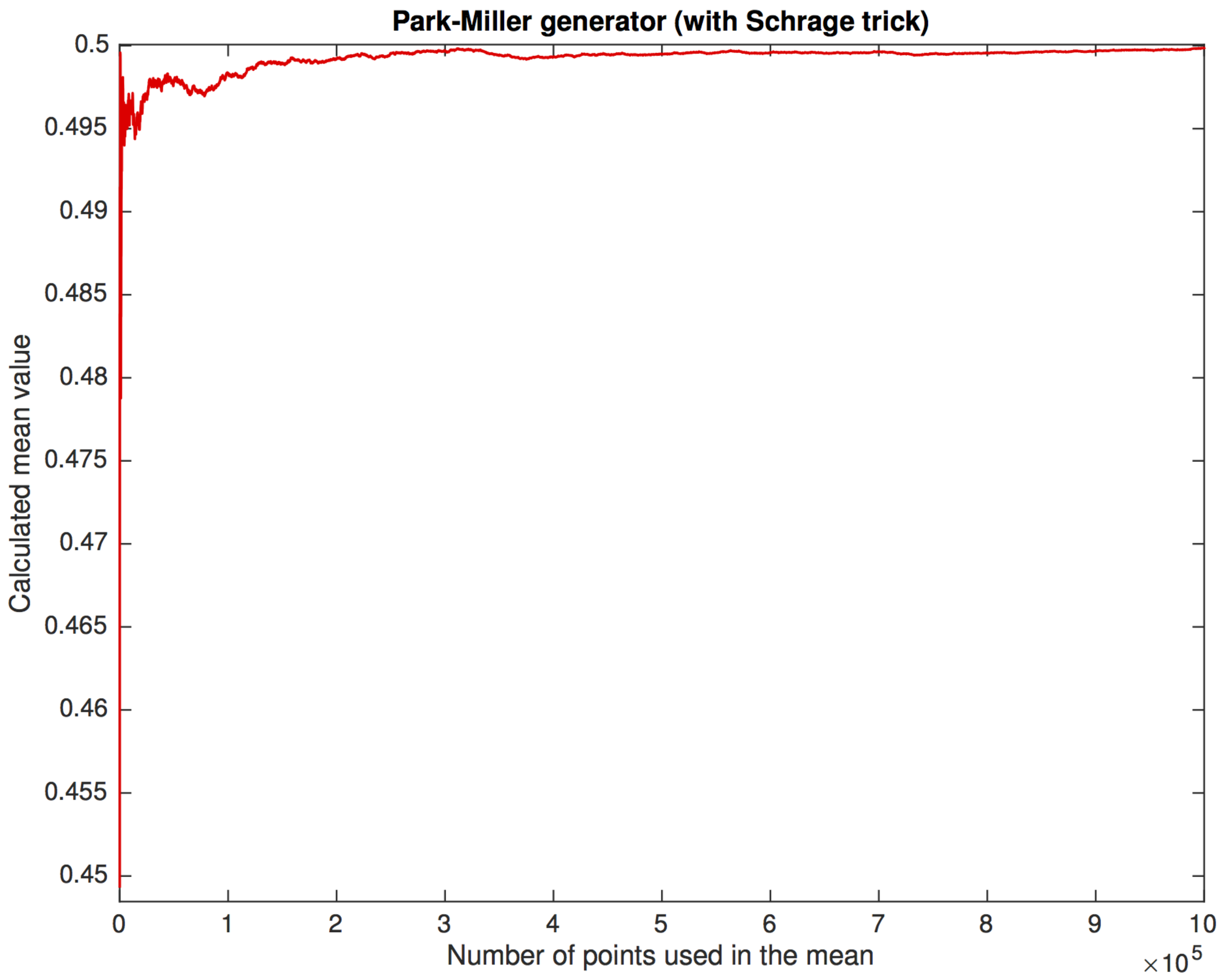

Figure 3: Park-Miller generator mean evolution analysis.

The fluctuations of the mean value seems reach the 0.0001 tolerance after 4×10^5 points in sequence (roughly). See the graph in figure 3.

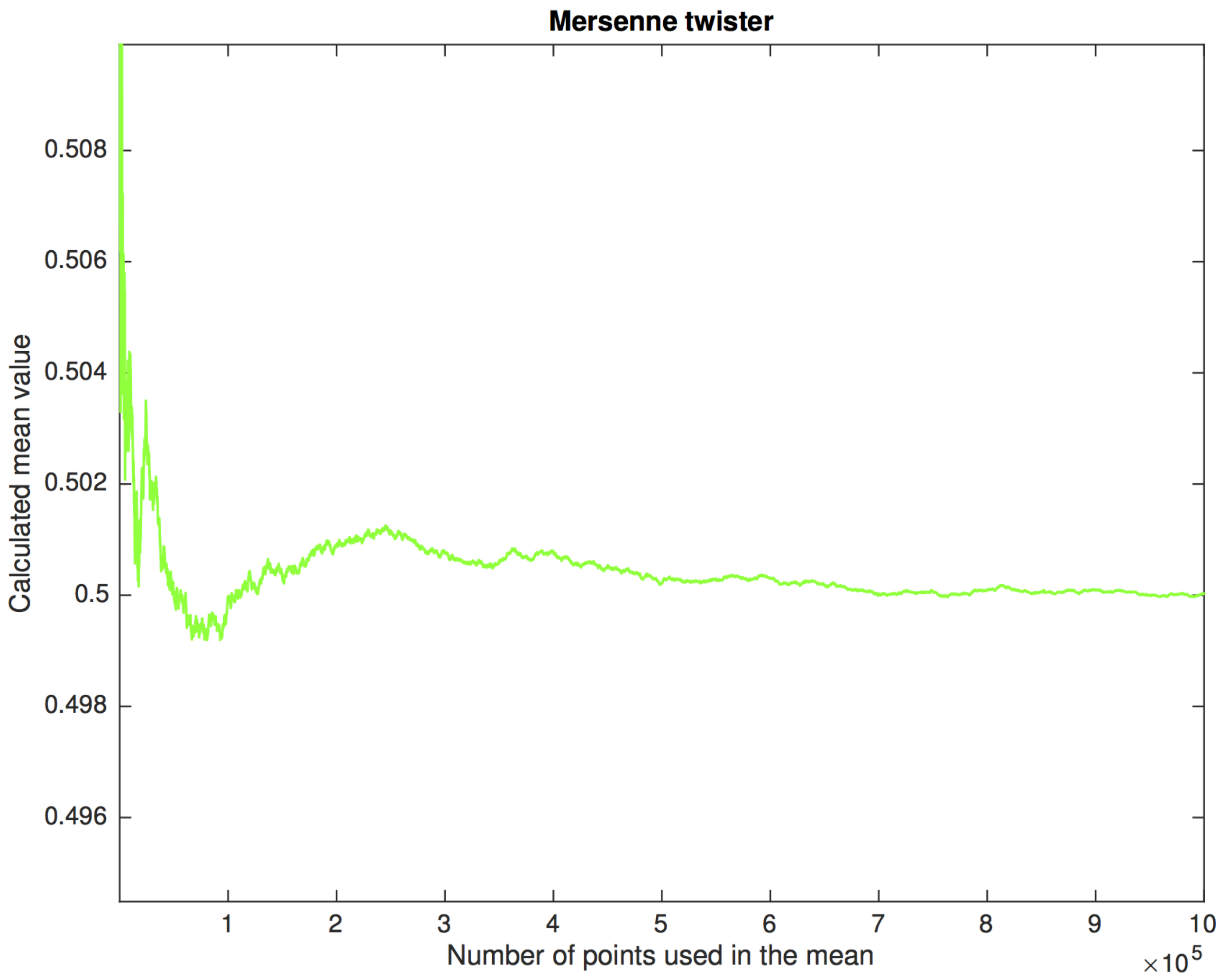

Figure 4: Mersenne twister mean evolution analysis.

The fluctuations of the mean value seems reach the 0.0001 tolerance after 5×10^5 points in sequence (roughly). See the graph in figure 4.

Problem 5

Command 3

```
$ ./ex1p5
 LCG 1 - Passes all the 3 rules
 Period:
     12386880
 LCG 2 - Fails the 1st rule
 Period:
            288
 LCG 3 - Fails the 2nd rule
 Period:
         31680
 LCG 4 - Fails the 3rd rule
 Period:
      3096720
```

We can see that, when all the rules are followed, we get a full cycle as the period (period will be equal to the modulus m). If one of the rules is not respected, the period decreases. In the program's output we see that failing the 3rd rule only is the best case (bigger period), while failing the first rule only is the worst.

Problem 6

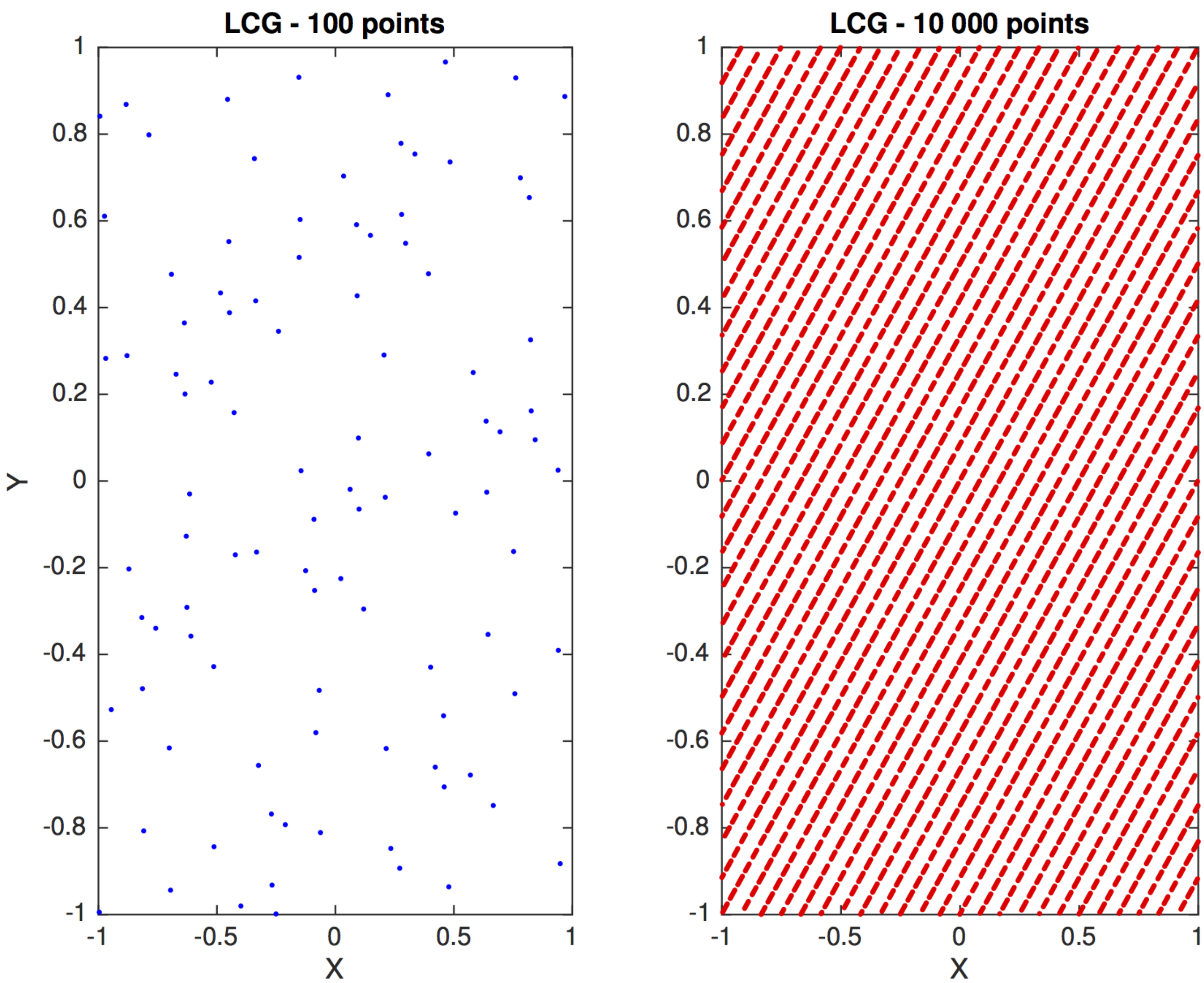

Figure 5: Linear congruential generator 2D plot.

The LCG plot doesn't seem random, it shows some tendencies like strides or lines. This shows that our Linear Congruential Generator is not a good random number generator.

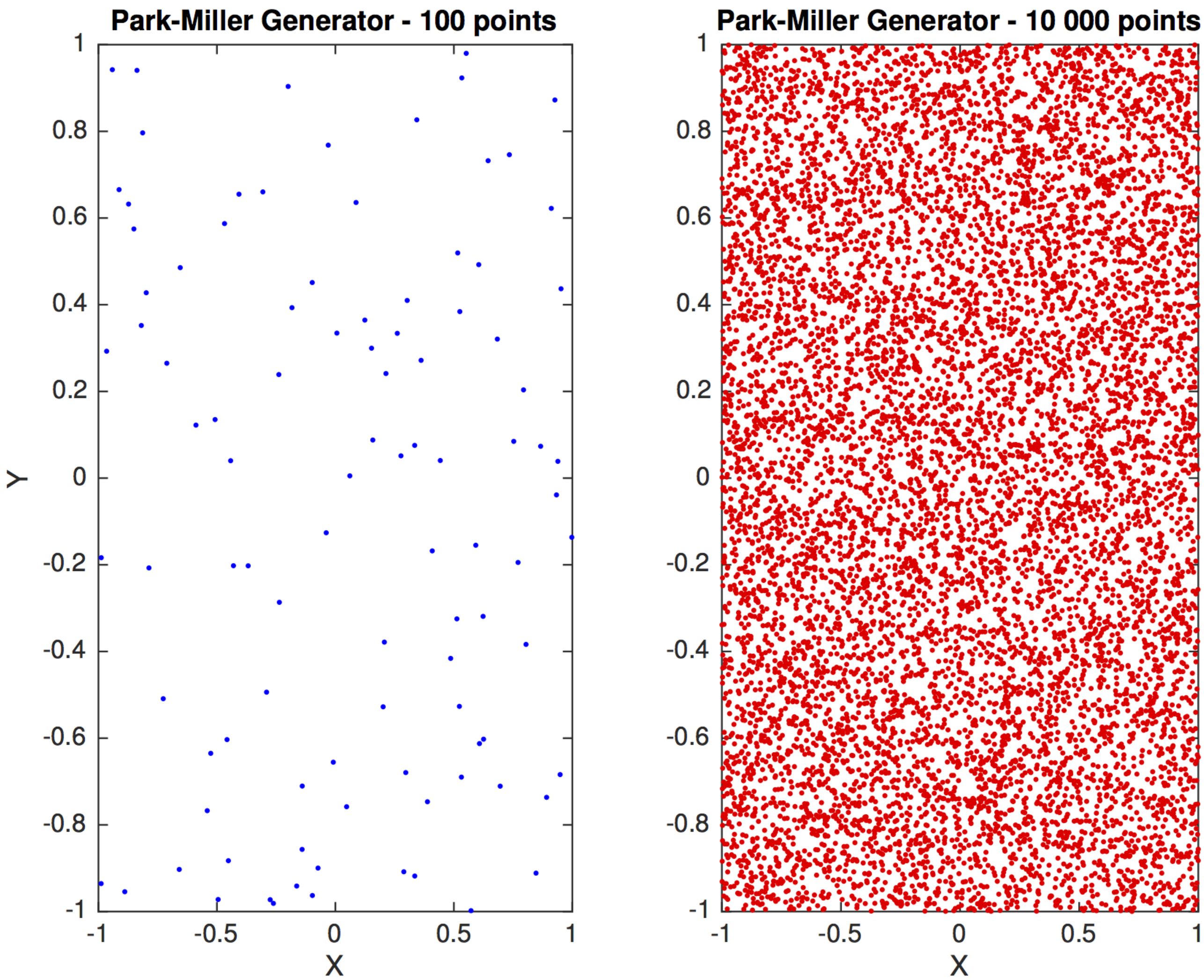

Figure 6: Park-Miller generator 2D plot.

The Park-Miller generator (with the Schrage trick) plot is way better than the LCG plot and its distribution looks more random.

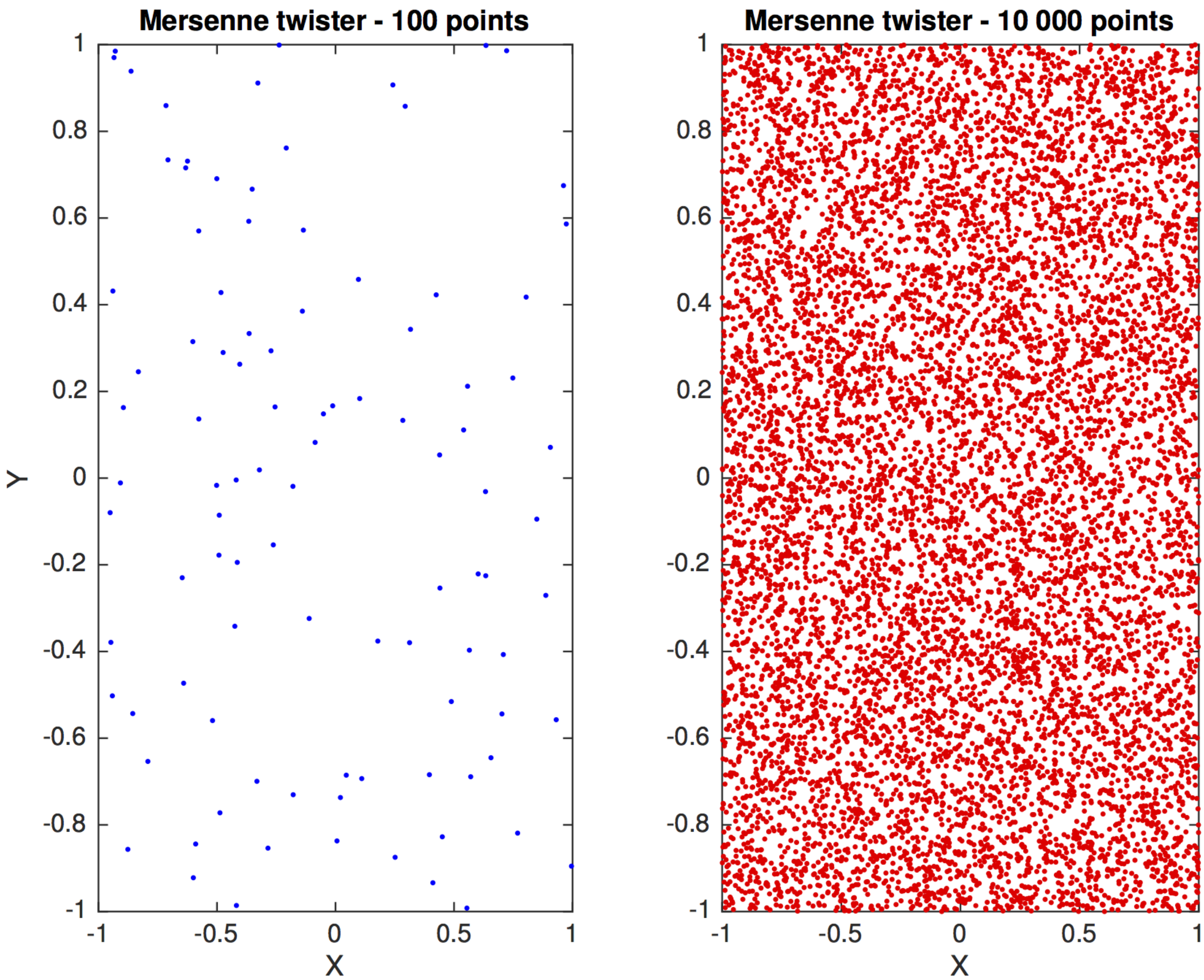

Figure 7: Mersenne twister 2D plot.

The Mersenne twister plot looks very much like the Park-Miller plot. Its distribution looks much more random than the LCG distribution.

Below is the second part of the problem.[1]

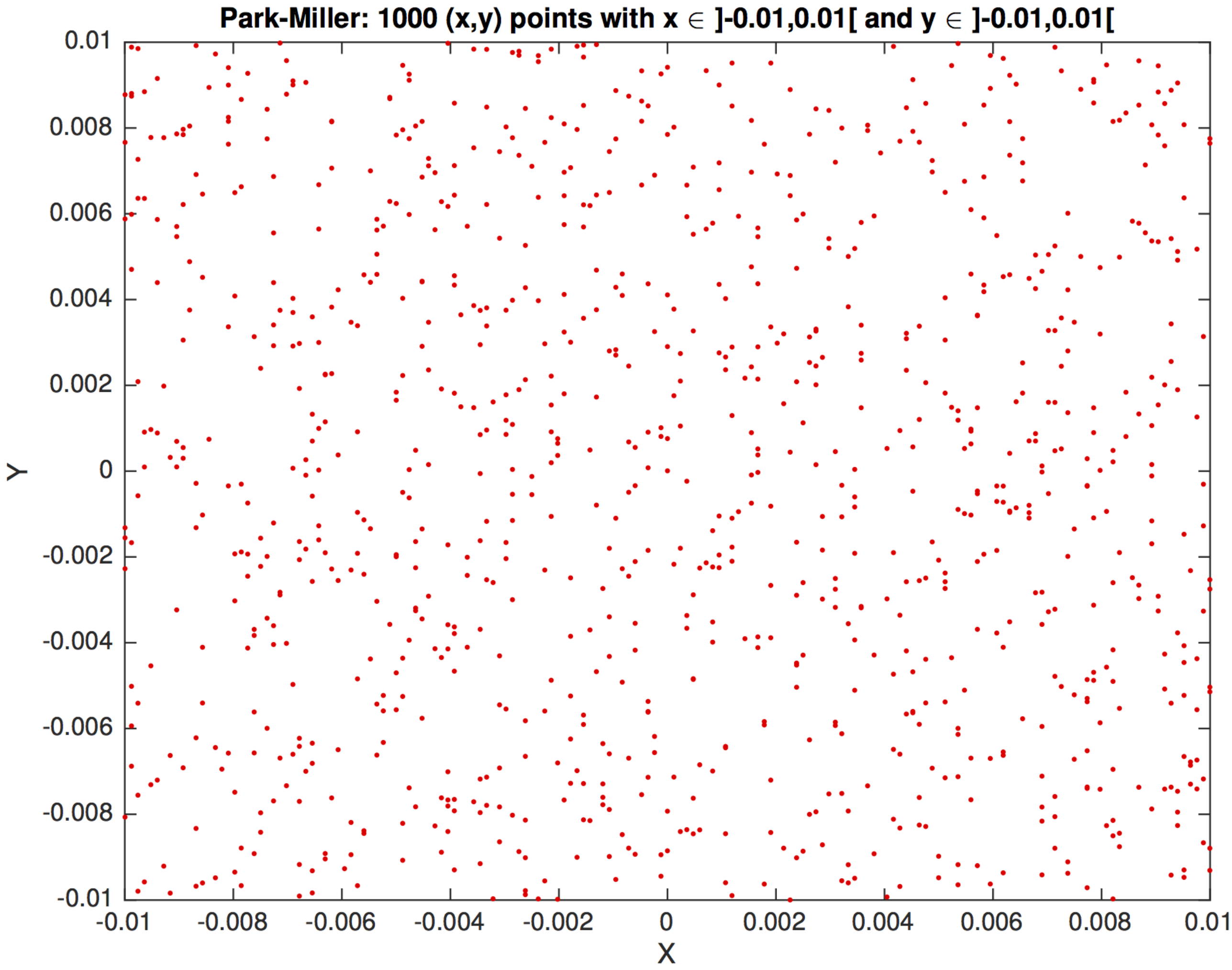

Figure 8: Park-Miller generator 2D plot in the determined interval (with 1000 points).

I can't find a problem in this region. It's the same with the Mersenne twister version. The only one that shows a problematic behavior is the LCG, since in this region it only shows a single diagonal line, but it was expected since the lines (or stripes) design can be easily noticed in the figure 5. This behavior of the LCG didn't change after I changed the seed. What happens when I change the seed is that the line (or strip) change its place and, depending on the seed, may never touch the desired region. In my case, if I use 42 as the seed, the desired region will become a blank space with no points.

See all images in the 'figures' folder inside the zip archive.

Problem 7

I have no idea of how crack the LCG sequence and obtain the (m,a,c) configuration. It should be done by a system of equations but the mod operator does makes the calculation harder.

The system of equations that need to be solved is:

$$\begin{cases} 850666 = (468995a + c) \mod m \\ 865757 = (850666a + c) \mod m \\ 182668 = (865757a + c) \mod m \\ 837799 = (182668a + c) \mod m \end{cases}$$

[1] In the source code the first part is commented. In order to test one part you need to coment the other. This avoids creating a lot of different codes to several similar operations in the same problem.