

Monte Carlo simulation methods, homework 2

1

$$a := x - y, \quad b := x + y$$

$$H(x, y) = x^2 + y^2 + 5(x - y)^2 = \frac{11}{2}a^2 + \frac{1}{2}b^2$$

$$p(x, y) \propto e^{-H(x,y)}$$

What is $\langle a^2 \rangle$ and $\langle b^2 \rangle$?

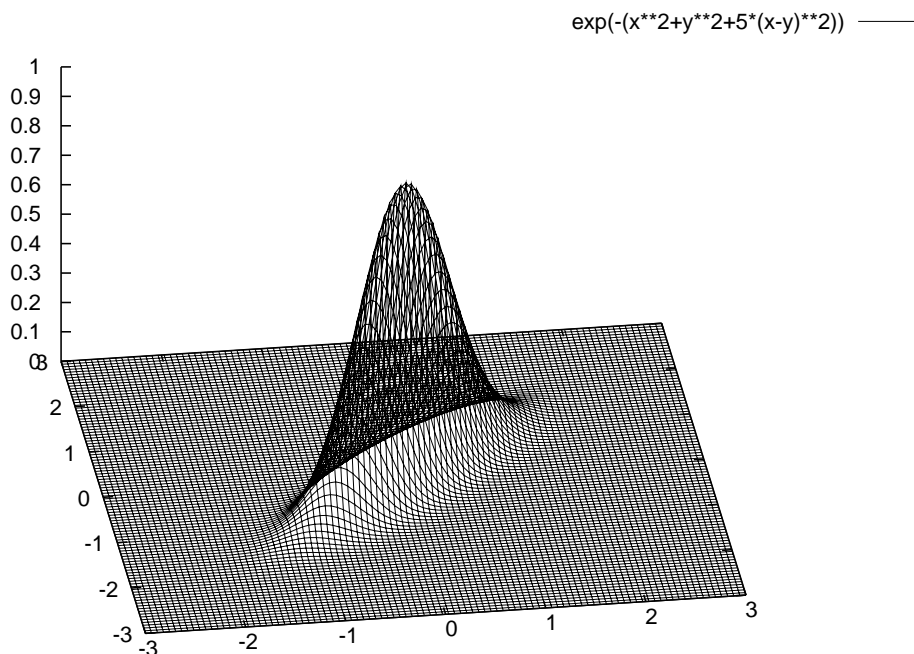


Figure 1: $e^{-H(x,y)}$

1.1 Monte Carlo with Metropolis update

1. Choose initial values for x and y . Since $(0,0)$ is the maximum point of the distribution, choosing it as the initial point is rather good choice, because then the algorithm starts to sample the 'important' points right away. Would the results be reasonable if the start point was, say, $(10000,10000)$?
2. Generate new 'trial' point:
There is some freedom how to do this, and this case choosing it so that $x_t = x_i + C * a$
 $y_t = y_i + C * b$
where (x_i, y_i) is the current point, C is some tunable constant, and a and b are random numbers from uniform distribution $[-1,1]$.
3. Calculate $H(x_t, y_t)$, and the next element in the Markov chain is chosen with rule:
if $(e^{-\Delta H} \equiv e^{-(H_t - H_i)} > \text{random number from uniform distribution } [0,1])$ $x_{i+1} = x_t, y_{i+1} = y_t$, else $x_{i+1} = x_i, y_{i+1} = y_i$.

In other words, if the trial state has lower energy than old state (is more probable), then always set the trial state as the next Markov chain element. If this is not the case, the next element will be the trial state with probability $e^{-\Delta H}$ and the old state with probability $1 - e^{-\Delta H}$.

4. 'Measure' relevant quantities and do something to them (note that when using metropolis update, rejecting the trial point does not mean that the measurement data should be rejected as well). In this case 'measuring' can simply mean that calculate quantities a^2 and b^2 with current x and y . These values can then be used to calculate the sum of these values ($\sum a^2$, $\sum b^2$). Note that usually it is reasonable to discard some of the first measurements (details depends on the problem) because of the initial configuration may be 'bad', but for example in this case with start point $(0,0)$ there is no need for that.
5. Repeat stages 2-4 as long as needed.
6. Now the more or less accurate values for the expectation values are simply:

$$\langle a^2 \rangle = \frac{1}{N} \sum a^2$$

$$\langle b^2 \rangle = \frac{1}{N} \sum b^2$$

where N is the number of iterations done.

1.2 Analytic

The task is to calculate

$$\langle (x - y)^2 \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - y)^2 p(x, y) dx dy$$

and

$$\langle (x + y)^2 \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y)^2 p(x, y) dx dy$$

Using a and b , equations can be written (denominator comes from normalisation)

$$\langle a^2 \rangle = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a^2 e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} dx dy}$$

$$\langle b^2 \rangle = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b^2 e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} dx dy}$$

Doing variable change in the integral from x, y to a, b

$$(a, b) = (x - y, x + y) \Leftrightarrow (x, y) = \left(\frac{a}{2} + \frac{b}{2}, -\frac{a}{2} + \frac{b}{2} \right)$$

and using $\int_{-\infty}^{\infty} e^{-x^2} = \sqrt{\pi}$, the denominator can be calculated:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} \left| \det \frac{\partial(x, y)}{\partial(a, b)} \right| da db$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} \left| \det \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \right| da db = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{11}{2}a^2 - \frac{1}{2}b^2} da db$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} e^{-\frac{11}{2}a^2} da \int_{-\infty}^{\infty} e^{-\frac{1}{2}b^2} db = \frac{1}{2} * \frac{\sqrt{2}}{\sqrt{11}} \sqrt{\pi} * \sqrt{2} \sqrt{\pi} = \frac{\pi}{\sqrt{11}}$$

Using integration by parts, numerator can be calculated similarly and the results are

$$\langle a^2 \rangle = \frac{1}{11} \approx 0.0909$$

$$\langle b^2 \rangle = 1$$

1.3 Monte Carlo results

The calculation was done with several settings, and from an example run with 10000 samples the best settings gave results

$$\langle a^2 \rangle = 0.0915688$$

$$\langle b^2 \rangle = 1.0404$$

Relative error from exact result for $\langle a^2 \rangle$ with different 'trial constant' C and measurement count starting from point (0,0):

C	1000	10000	50000	100000	1000000	Acceptrate
0.01	0.70	0.32	0.17	0.16	0.083	0.995
0.1	0.047	0.036	0.084	0.062	0.024	0.95
1	0.028	0.0073	0.0038	0.0072	0.0040	0.56
10	0.26	0.29	0.027	0.0019	6e - 5	0.019
100	1	0.99	0.046	0.20	0.24	0.00017

From these values, constant value 1 seems to be the best, as indicated also by its good accept rate. It's worth noting that the constant has rather notable effect on the results, and it may always not be that easy to 'know' reasonable constant beforehand.

What kind of results does the error estimate(lecture notes page 51) assuming statistically independent measurement give in this case?

For $\langle b^2 \rangle$ with C=0.01

Error	1σ-error
0.998	0.00011
0.96	0.00035
0.76	0.00092
0.17	0.0026
0.36	0.00095

For $\langle b^2 \rangle$ with C=1

Error	1σ-error
0.30	0.036
0.040	0.015
0.0095	0.0064
0.0058	0.0045
0.00044	0.0014

In case C=0.01 the estimate fails badly, which is not a surprise given the small update step, while in case C=1 it seems rather good indicating smaller autocorrelation time.

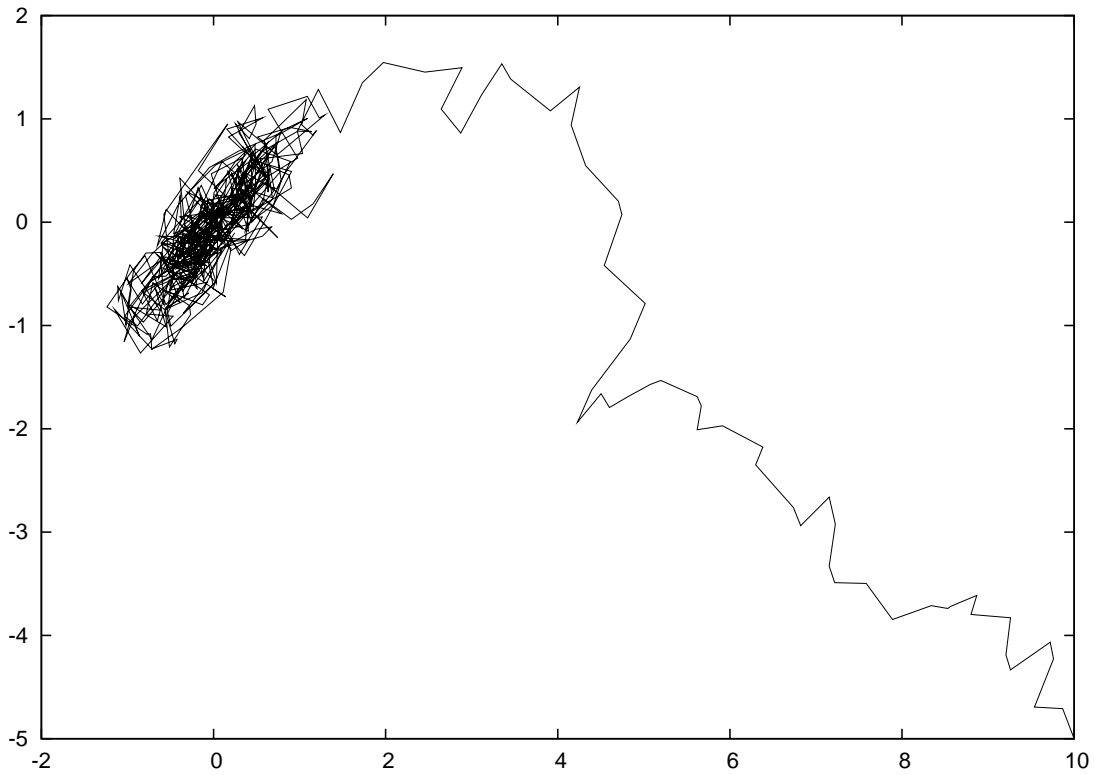


Figure 2: First 1000 points (connected with lines) generated by the algorithm with start point (10,-5) and $C = 1$.

2

The cycle length can be determined simply by comparing new 'irnd' number to the seed (given that seed is $< 2^{31}$). This is because 'irnd' variable fully determines the next value, so when the generator returns number equal to the seed, the generator is back to the beginning.

Seed	Cycle
1	536870912
4	134217728
32	16777216
58804	134217728
32768	16384
65536	8192
574042	268435456
5444066	268435456
268435456	2
1846665012	134217728

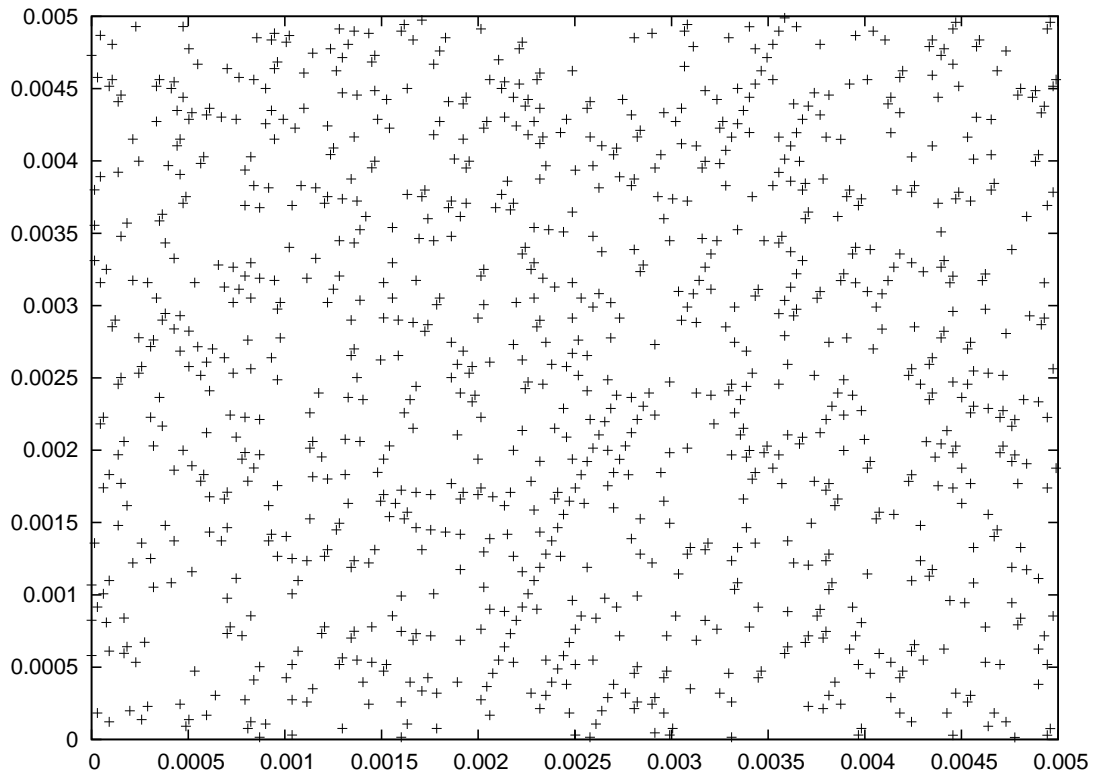


Figure 3: Seed 1, 1000 points

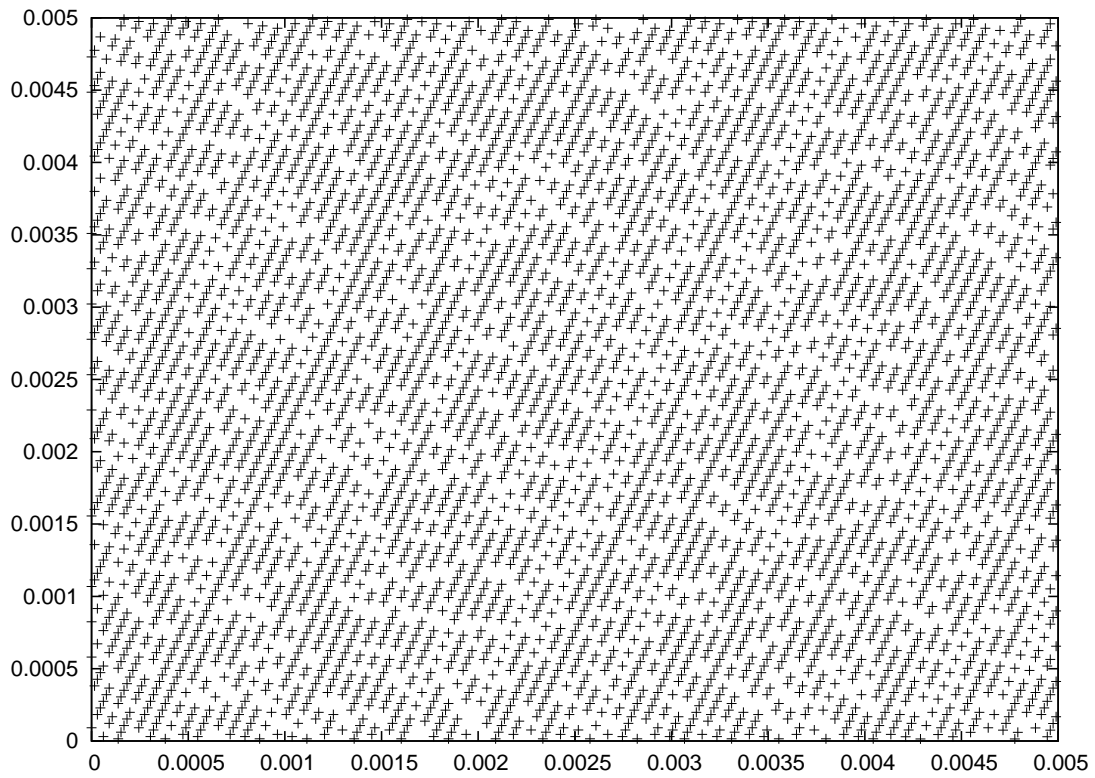


Figure 4: Seed 1, 5000 points

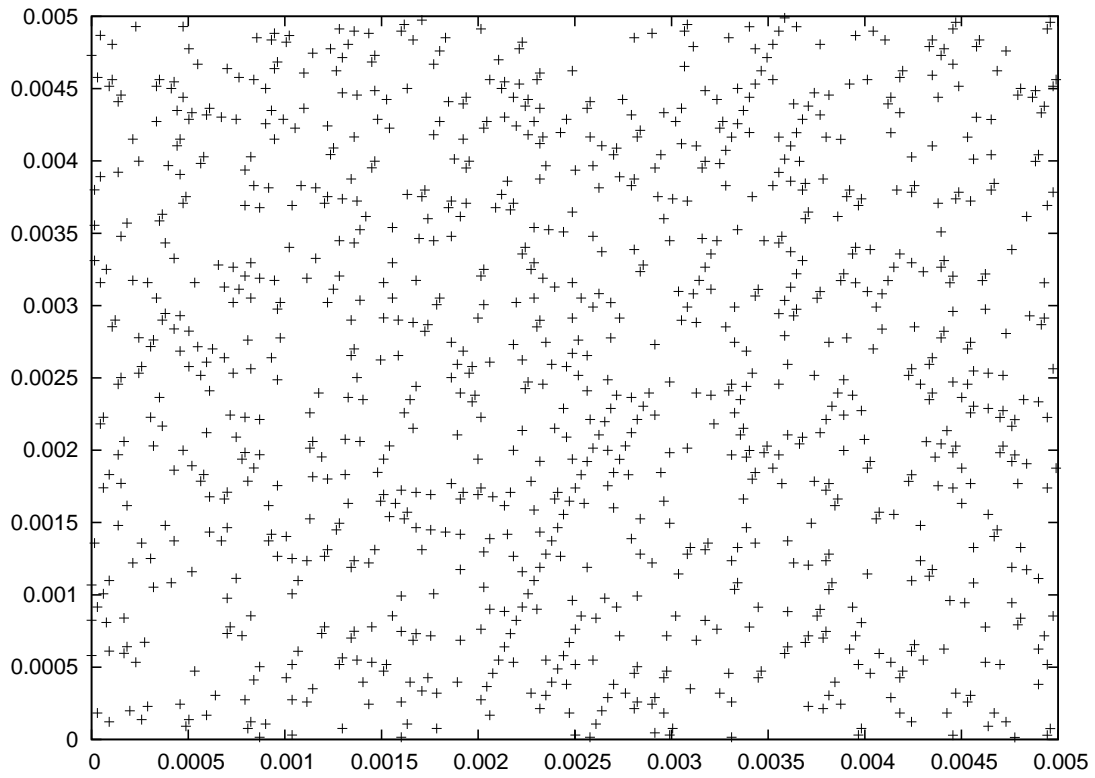


Figure 5: Seed 1, 1000 points

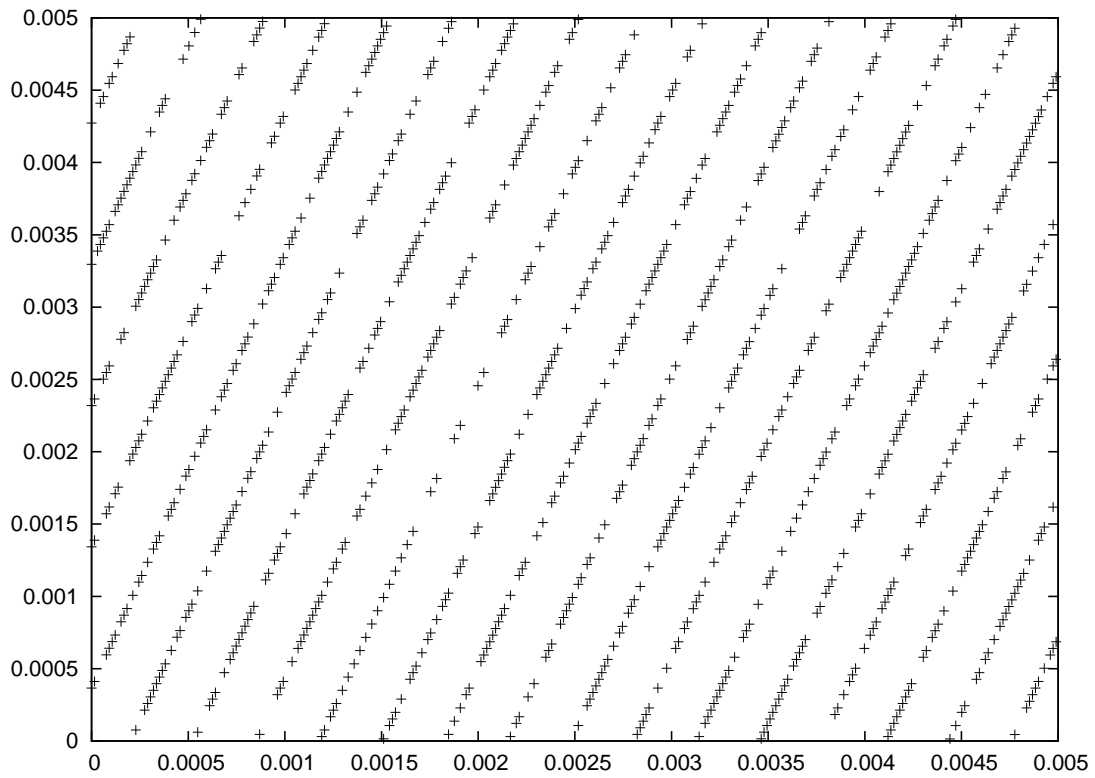


Figure 6: Seed 4, 1000 points

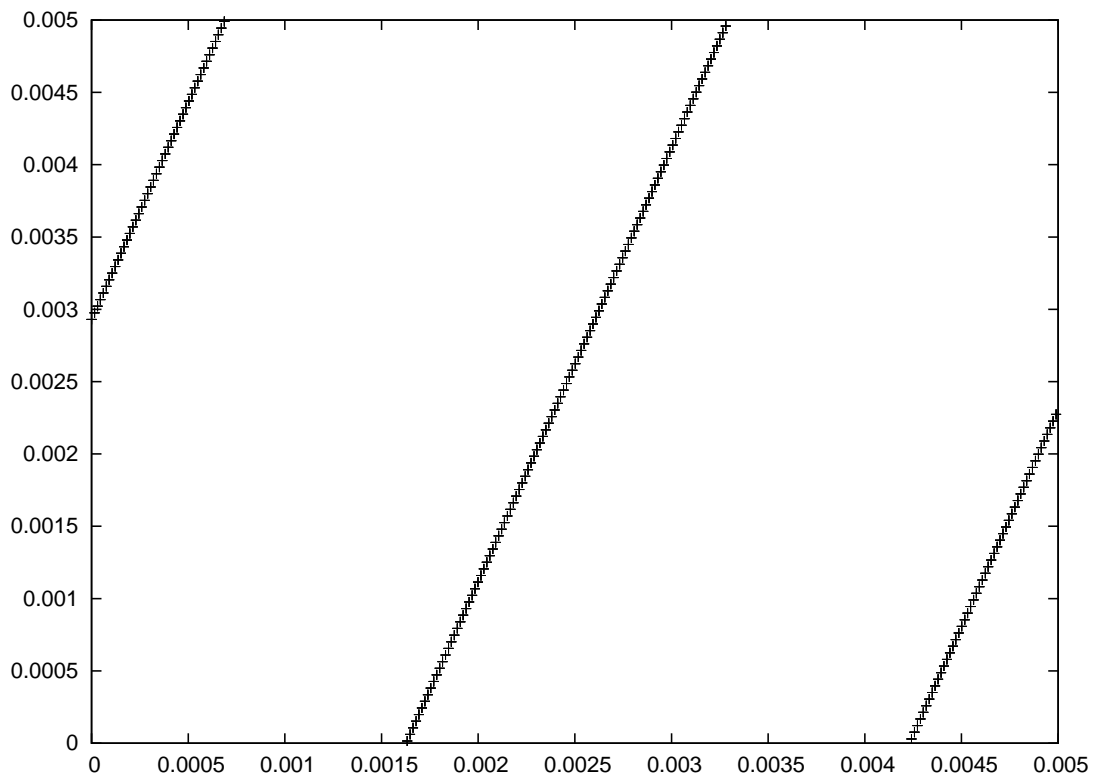


Figure 7: Seed 32, 1000 points