SIF40AH/DIF4997
Nano-particle and polymer physics I
SOLUTION of EXERCISE 6

A) The d-dimensional vectors $\vec{X} = (X_1, X_2, \ldots, X_d), X_i \in [0,1]$ are generated from a uniform random-number generator. We intend to study the distribution of $\vec{X}$ within a d-dimensional sphere-shell. The generality of d dimensions (line, circle, sphere, ...) complicates things a bit, but dont't give up.

The volume of a d-dimensional sphere of radius $X = |\vec{X}|$ equals

$$V_d(X) = \Omega_d X^d, \tag{1}$$

where Ω_d is the volume of a sphere in d dimensions with radius equal to 1.[1]

The volume of a d-dimensional shell of sphere with thickness $\mathrm{d}X$ and radius X equals

$$\mathrm{d}V_d = \Omega_d \cdot d \cdot X^{d-1}\mathrm{d}X. \tag{2}$$

Eq. (2) may be seen from the fact that $\mathrm{d}V_d = A_d \mathrm{d}X$, so $A_d = \frac{\mathrm{d}V_d(X)}{\mathrm{d}X} = \Omega_d \cdot d \cdot X^{d-1}$, implying Eq. (2). Alternatively, visualize it by integration:

$$V_d(X) = \int_0^X A_d \, \mathrm{d}X = \int_0^X \Omega_d \, d \, X^{d-1}\mathrm{d}X = \Omega_d X^d \tag{3}$$

The number of vectors, n, within a shell of sphere at radius X, relative to the number N within the whole sphere of radius R is

$$\frac{n}{N} = \frac{\mathrm{d}V_d}{V_d(R)} = \frac{\Omega_d \, d \, X^{d-1}\mathrm{d}X}{\Omega_d R^d} = \frac{d \cdot X^{d-1}\mathrm{d}X}{R^d}. \tag{4}$$

So far for infinitesimal $\mathrm{d}X$. For finite $\mathrm{d}X = \Delta X$ we evaluate X in $\bar{X}$ within the interval $(X, X + \Delta X)$:

$$n = N \cdot \frac{d}{R^d} \cdot \bar{X}^{d-1}\Delta X. \tag{5}$$

An estimate of $\bar{X}$ is the arithmetic middle in the interval:

$$\bar{X}_1 = \frac{\Delta X}{2}, \quad \bar{X}_2 = \frac{3\Delta X}{2}, \quad \bar{X}_i = \frac{(2i-1)\Delta X}{2} = (i - 1/2)\Delta X. \tag{6}$$

The number of vectors within the interval ΔX is thus

$$n = \frac{Nd}{R^d} \cdot \left(i - \frac{1}{2}\right)^{d-1} (\Delta X)^d. \tag{7}$$

Simulation: We have chosen: $d = 2, \Delta X = 1/10, R = 1, N = 100000$

With these parameters the estimated numbers of vectors is according to Eq. (7):

$$n = \frac{100000 \cdot 2}{1} \cdot \left(i - \frac{1}{2}\right)^1 \left(\frac{1}{10}\right)^2 = 2000 \cdot \left(i - \frac{1}{2}\right)^1 \tag{8}$$

Estimated and simulated result in the following table. (Numbers from P.Skjetne using Turbo Pascal ver 5.5).

[1] $\Omega_1 = 2, \Omega_2 = \pi, \Omega_3 = 4\pi/3, \Omega_4 = \pi^2/2, \Omega_5 = 8\pi^2/15, \Omega_6 = \pi^3/6$, generally: $\Omega_d = \dfrac{2\pi^{d/2}}{d \cdot \Gamma(d/2)}$

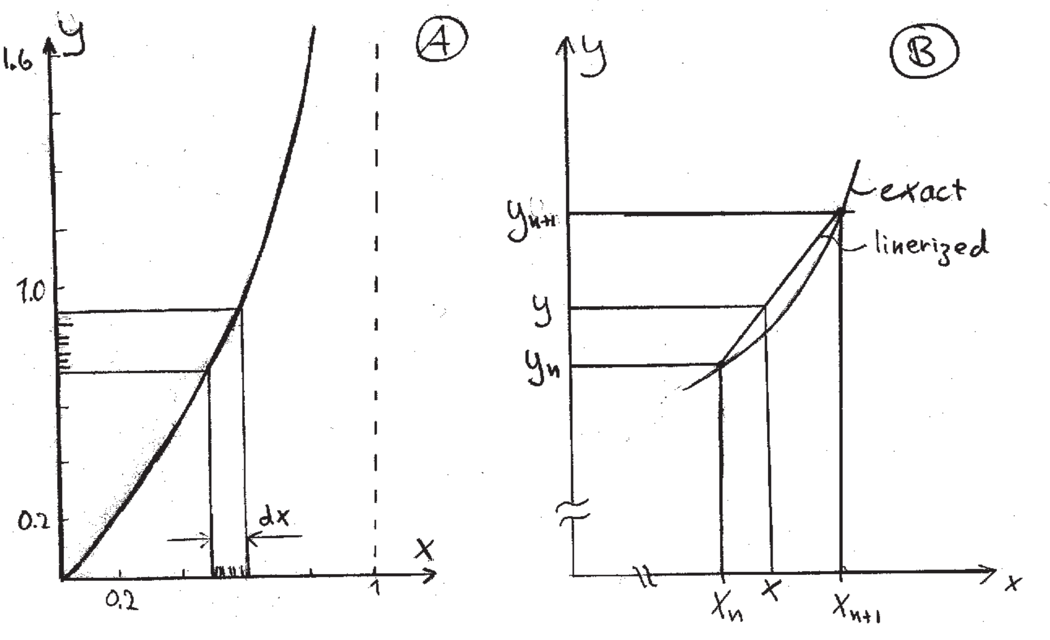
Interval	$\bar{X}$	Theoretical	Simulated
1	0.05	1000	1024
2	0.15	3000	3044
3	0.25	5000	5051
4	0.35	7000	6881
5	0.45	9000	9122
6	0.55	11000	11072
7	0.65	13000	13014
8	0.75	15000	14893
9	0.85	17000	16904
10	0.95	19000	18995
Sum		101000	100000

The theoretical values do not summarize to $N = 100000$ because of the approximation of $\bar{X}$.

B) Available is the uniform distribution $p(x) = 1 \; \forall \; x \in [0, 1]$, and we want to obtain a distribution $p(y) = \exp\{-y\} = e^{-y}$. Note that $p(y)$ is normalized because $\int_0^\infty p(y)\mathrm{d}y = \left[-e^{-y}\right]_0^\infty = 1$.

Because $p(x)$ is uniform the hits on x is uniformly distributed along the x-axis. The distribution along y-axis should be according to $p(y) = e^{-y}$, that is highest density of hits at $y = 0$ and decreasing constantly to 0 (figure A below). In the numerical transformation the numbers of hits $\mathrm{d}N_x$ within $\mathrm{d}x$ is mapped to exactly the same number of hits $\mathrm{d}N_y$ within (a wider) $\mathrm{d}y$. As the density of hits is $p(x)$ and $p(y)$, respectively, we obtain:

$$\mathrm{d}N_x = \mathrm{d}N_y \quad \Rightarrow \quad p(x)\mathrm{d}x = p(y)\mathrm{d}y. \tag{9}$$



To determine the formulae of transformation we integrate Eq. (9) from $(0, 0)$ to (x, y):

$$\int_0^x p(x)\mathrm{d}x = \int_0^y p(y)\mathrm{d}y \quad \Rightarrow \quad \int_0^x 1 \,\mathrm{d}x = \int_0^y e^{-y}\mathrm{d}y \quad \Rightarrow \quad x = 1 - e^{-y} \tag{10}$$

The inverse function is

$$\underline{y(x) = -\ln(1 - x)}, \tag{11}$$

and with x uniformly distributed on $x \in [0, 1]$ we obtain the required distribution $p(y)$.

We may also argument for this distribution by an approximate numerical method:

We divide the interval $x \in [0, 1]$ in N equal intervals and approximates the tranformation graph to a straight line between two neighbouring points (figure B above). The point (x_n, y_n) is given by

$$\begin{aligned} x_n &= 1 - e^{-y_n}, \quad \text{where } x_n = \frac{n}{N} \\ \Rightarrow \quad y_n &= -\ln\left(1 - \frac{n}{N}\right) \end{aligned} \tag{12}$$

Inbetween the neighbouring points we approximate to a straight line:

$$\frac{y(x) - y_n}{x - x_n} = \frac{\Delta y}{\Delta x} = \frac{y_{n+1} - y_n}{x_{n+1} - x_n} = \frac{y_{n+1} - y_n}{1/N} \tag{13}$$

The n to be used for the actual x is the one which makes x belong to the interval $(\frac{n}{N}, \frac{n+1}{N})$. $y(x)$ is found to be:

$$\begin{aligned} y(x) &= y_n + N \cdot [y_{n+1} - y_n] \cdot \left(x - \frac{n}{N}\right) \\ &\overset{(12)}{=} -\ln\left(1 - \frac{n}{N}\right) - N\left[\ln\left(1 - \frac{n+1}{N}\right) - \ln\left(1 - \frac{n}{N}\right)\right]\left(x - \frac{n}{N}\right) \\ &= -\ln\left(1 - \frac{n}{N}\right) - N\ln\left(1 - \frac{1}{N} \cdot \left(1 - \frac{n}{N}\right)^{-1}\right)\left(x - \frac{n}{N}\right) \\ &= -\ln\left(1 - \frac{n}{N}\right) + N\frac{1}{N} \cdot \left(1 - \frac{n}{N}\right)^{-1}\left(x - \frac{n}{N}\right) \\ &\approx -\ln\left(1 - \frac{n}{N}\right) + \left(1 + \frac{n}{N}\right)\left(x - \frac{n}{N}\right) \end{aligned} \tag{14}$$

where we have utilized that for large N (small ϵ) is $\ln(1+\epsilon) \approx \epsilon$. Further, $\frac{n}{N} \to x$ for large N, so the result is:

$$y(x) \approx -\ln(1 - x), \tag{15}$$

as equals the result from the analytical method above.

C) The Box-Muller algorithm to generate random Gaussian distributed numbers is given in text.

Simulation:
The result of drawing x_1 and x_2 randomly in $[0, 1]$ and using the Box-Muller algorithm is plotted below. In the simulation we have used $N = 100000$ and normalized $y(x)$. $y \in [-5, 5]$ is divided in 20 intervals and the number of hits within each interval is plotted. (Data from P. Skjetne, theoretical and numerical curve:)