# Using Random Numbers

Modeling and Simulation of Biological Systems 21-366B

Lecture 2-3

A textbook on probability:

G.R. Grimmett and D.R. Stirzaker

Probability and Random Processes

OXFORD

# Motivation

Why randomness?

It seems to be present everywhere.
It may be a useful approach to simplify models
It may address our limited knowledge about a system
It may be useful to express variability in a population
…..

Examples:

stock market behavior, individual stocks and averages
motion of microscopic particles in water
bacterium trajectory in a fluid
Affinity maturation of B cell antibodies
traffic behavior
….

# Cont.

Randomness may be in the structure (spatial) or dynamics

Even deterministic models may exhibit an apparently random behavior

- Atomistic models of fluids and solids (movies later)
- Insufficient sampling of a phenomenon (movies later)
- Grain boundaries in polycrystalline materials (pictures)

# Cont.

Large networks of biochemical reactions seem to have randomness in its connectivity

metabolic networks

protein-protein interactions

Internet connectivity

……

# Random Number Generation

## Uniform distribution

$$P(a \leq x \leq b) = \int_a^b 1 dx = b - a \qquad 0 \leq a \leq b \leq 1$$

MATLAB function:

rand : gives a single number in (0,1), distributed uniformly in [0,1]

rand(n,1) : give a vector of n numbers, …

rand(n) : gives a n by n matrix ….

Let $X_k, \quad k = 1, 2, \ldots$ be random numbers distributed uniformly in $[0, 1]$. What can we say about

$$S_n = \frac{1}{n} \sum_{j=1}^{n} X_k$$

# Random Variables Attaining a few values

Let a random variable attain two values,

$$\begin{aligned} P(x=1) &= \lambda \\ P(x=0) &= 1-\lambda \end{aligned}$$

To generate such a random variable:

$$\text{Pick} \quad 0 \le r \le 1, \quad \text{according to U[0,1].}$$

$$\begin{array}{lll} if & r \le \lambda & \text{Set X=1} \\ else & & \text{Set X=0} \end{array}$$

Later we will regard the event X=1 as a jump.

# Basic MATLAB (use MATLAB help!!)

Vectors:

```
v = ones(n,1);          %generate a row vector of ones
a = v’;                 % a is a column vector of ones
b = zeros(n,1);         % a row vector of zeroes
v = [1:10];             % a vector [1 2 3 .. 10]
a = [1 2 3; 4 5 6];     % a 2 by 3 matrix
```

Matrix Operations:

```
A = B + C;   A = B – C;
A*v;                    % v must be a column vector
```

Loops:

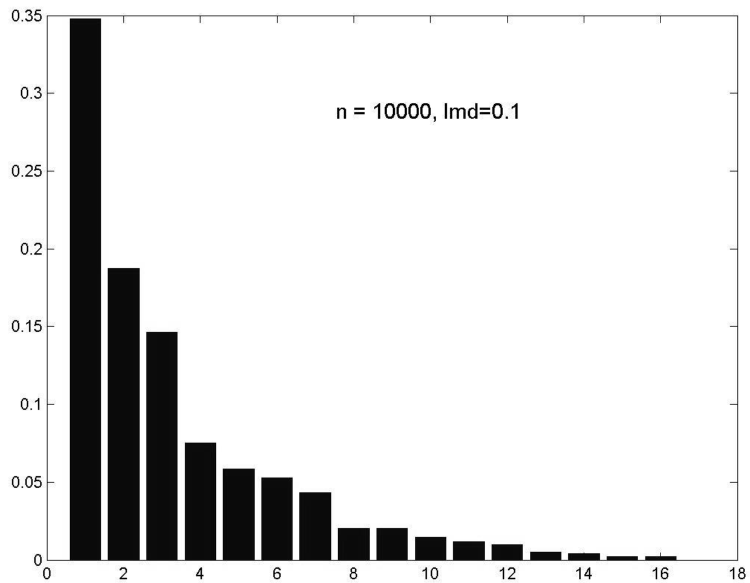
```
for i=1:10
        ….
end
```

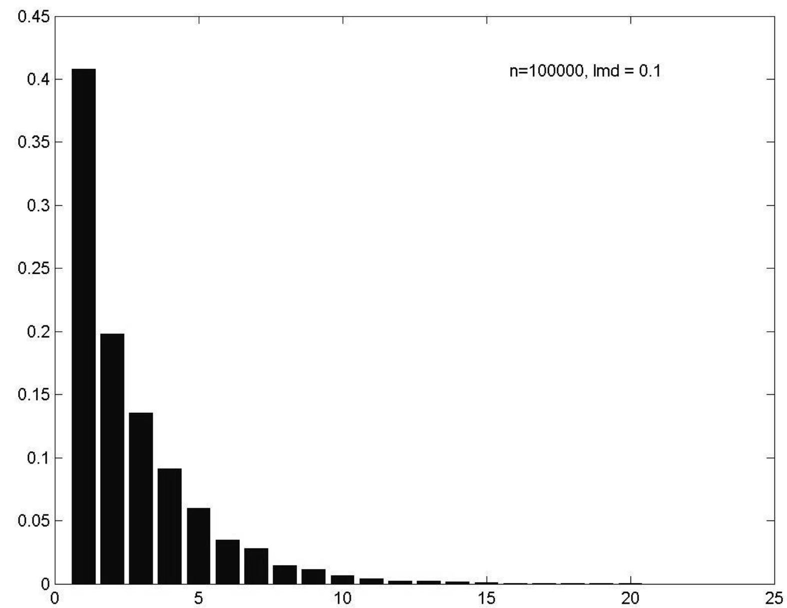
# more MATLAB commands

```
C = fix(10*rand(3,2));     %create a 3 by 2 matrix, of integers 0->9
fix:                       %round toward zero
plot(x);                   % plot the values in a vector x
bar(x);                    % create a bar graph form a vector
sum(x);                    %
hist(x);                   % create a histogram from the values of x
hist(x,n);                 % histogram with n bins
x = rand(n,1);             % a vector of n random numbers U[0,1]
Ind = find(x>=0.5);        % ind is an array of indices satisfying ….
clear all;                 % clear memory
```

# Studying Jump Time Intervals

```
%jumpTime.m
% find the distribution of jump time intervals
n=10000;
Lamd = 0.1;
r = rand(n,1);
Ind = find(r<=Lamd);        %Ind is an array of indices at which jumps will
                            %occur
for i=1:length(Ind)-1
    j(i) = Ind(i+1)-Ind(i);
end
v = hist(j,20);             %construct histogram of the j values
w = v/sum(v);               %normalize
bar(w);
```


n = 10000, lmd=0.1
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 2 4 6 8 10 12 14 16 18


n=100000, lmd = 0.1
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0 5 10 15 20 25

**Normal distribution:** $\mathcal{N}(0,1)$

$$P(a \leq x \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

MATLAB function

randn : a single random number

randn(n,1) : n vector of n random numbers

randn(n) : an n by n matrix of random numbers

# Some Limit Theorems

**Law of Large Numbers:** Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. random variables with finite mean $\mu$ and non-zero finite variance $\sigma^2$, then $S_n = \sum_{j=1}^{n} X_k$ satisfies

$$\frac{1}{n} S_n \to \mu$$

**Central Limit Theorem:** Let $X_k, S_n$ be as above. Then

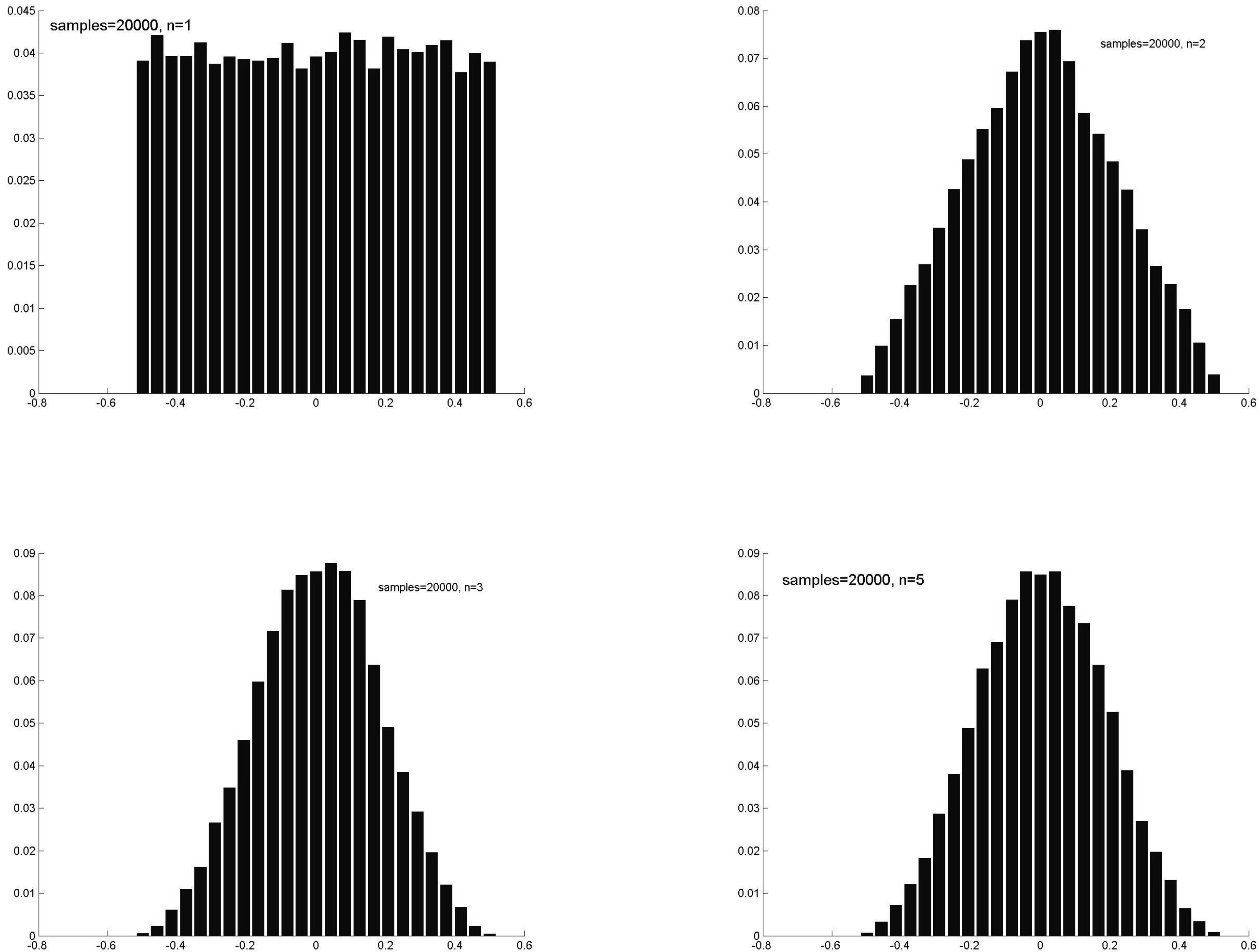
$$\frac{1}{\sqrt{n\sigma^2}}(S_n - n\mu) \to \mathcal{N}(0,1)$$

# Converges of Averages (LLN)

```
% average.m
n=1000;
y = rand(n,1);
s(1)=y(1);
for i=2:n
    s(i) = s(i-1)+y(i);
end
for i=2:n
    s(i) = s(i)/i;
end
plot(s);
```

# Behavior of  Averages (CLT)

```
% CLT.m
n=1;
s = zeros(n,1);
c = ['r','g','b','c','m','y','k'];
for l=1:20000
    y = rand(n,1);
    s(1)=y(1);
    for i=2:n
        s(i) = s(i-1)+y(i);
    end
    z(l) = (s(n)-0.5*n)/sqrt(n);
end
m=25;
w=hist(z, m);
v=linspace(-0.5,0.5,m);
bar(v,w/m);
```


samples=20000, n=1
samples=20000, n=2
samples=20000, n=3
samples=20000, n=5

# Other Distribution Functions

Constructing random numbers according to a specified distribution $F(z)$

Let $Y_k$ be distributed uniformly in [0,1]. Let $f(x)$ be a desired density function. Define

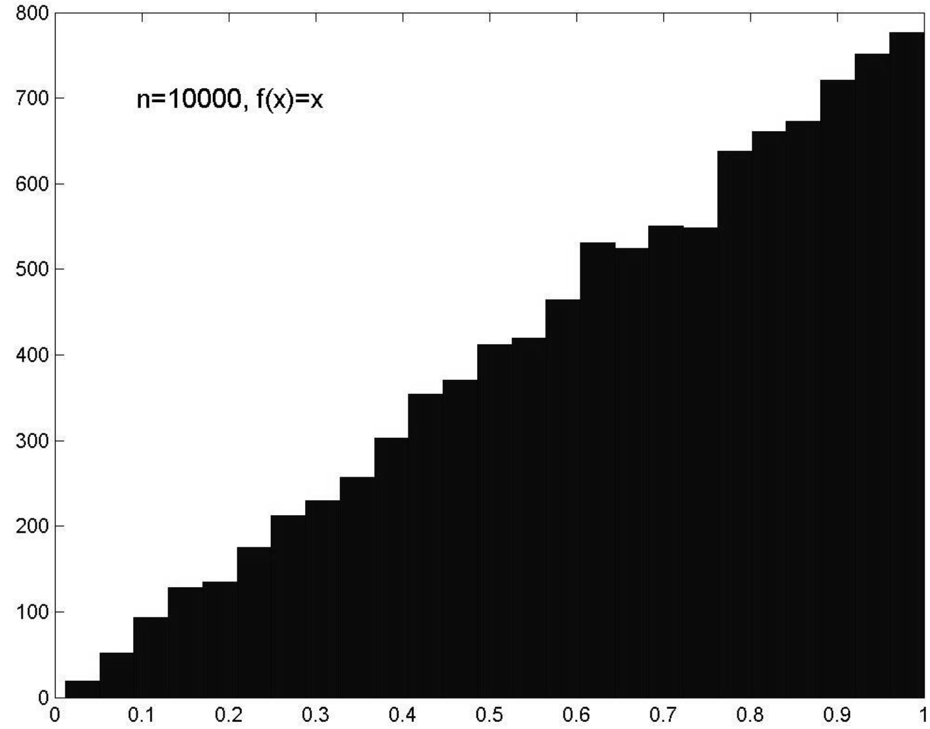
$$X_k = F^{-1}(Y_k) \qquad\qquad F(y) = \int^{y} f(s)ds$$

Then

$$P(a \leq X_k \leq b) = \int_a^b f(z)dz$$

# Non-uniform Distribution

```
%nonuniformDist.m
% density f(x) = 2x, x in [0,1]
% F(z) = int_0^z f(x)dx = x^2
n=10000;
y = rand(n,1);
x = sqrt(y);
m=25;
hist(x,m);
```

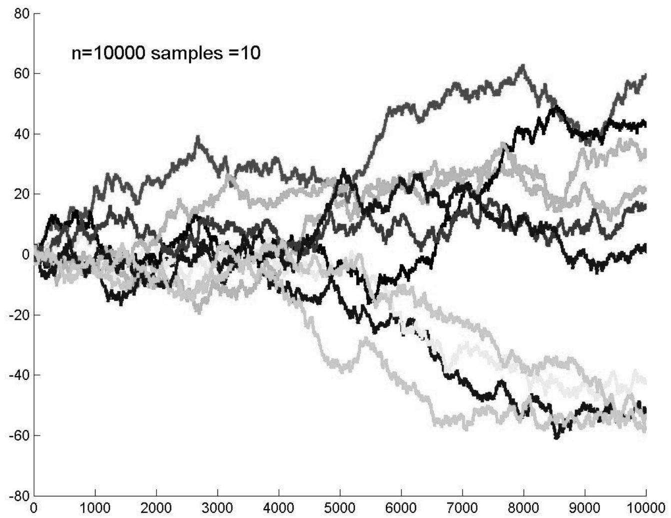


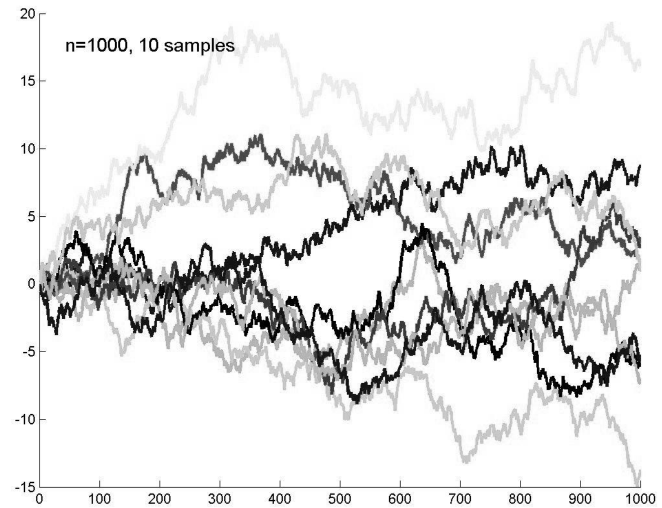
# A Random Walk: 1D

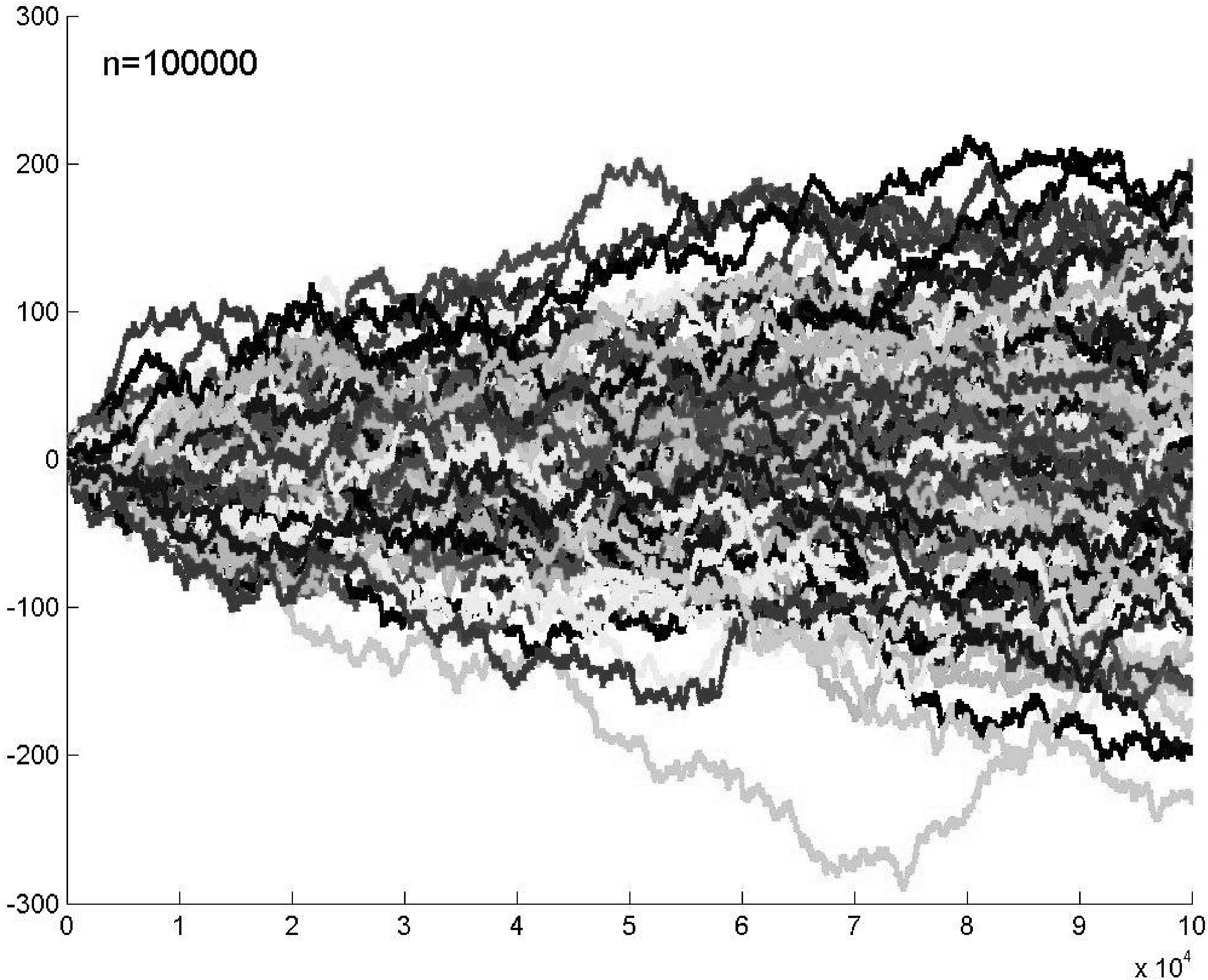
Let the position of a particle at time $k = 1, 2, \ldots$ be denoted by $X_k$. We assume that the particle changes its position according to

$$P(X_{k+1} = X_k \pm h) = \frac{1}{2}$$

Sample paths




n=100000
300
200
100
0
-100
-200
-300
0
1
2
3
4
5
6
7
8
9
10
x 10^4

# MATLAB code

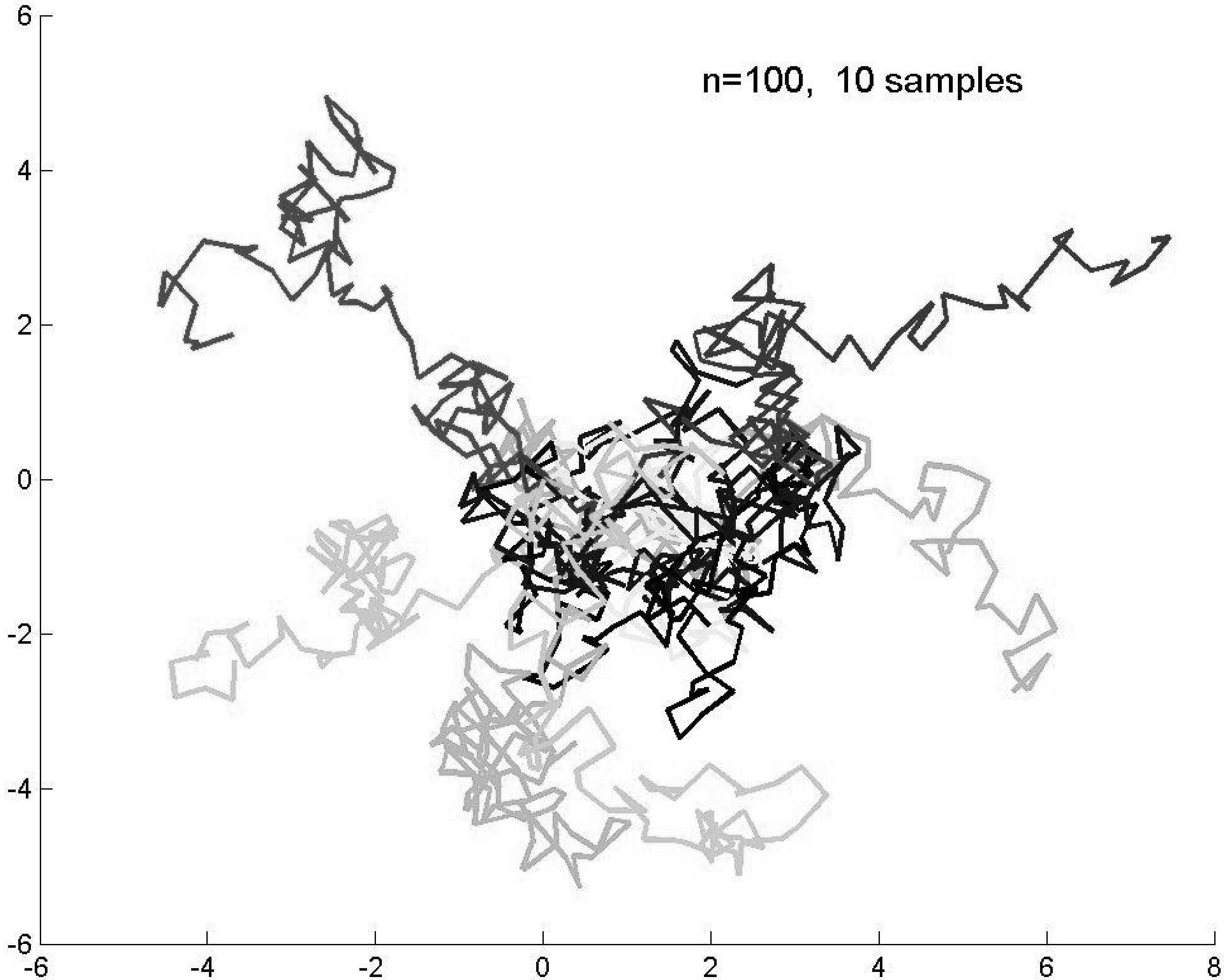
```
% randWalk1D.m
% (symmetric)  random walk in 1D
n=100000;                    %number of steps
figure;
hold on;
c = ['r','g','b','c','m','y','k'];
for l=1:100
      %x = (rand(n,1) - 0.5*ones(n,1));  %uniformly distributed jumps
      %x = randn(n,1);                   % normal
                                         %jumps are +-1
      x = (rand(n,1) - 0.5*ones(n,1));
      indP = find(x>0);
      indM = find(x<=0);
      x(indP) = 1;
      x(indM) = -1;
      z = zeros(n,1);
      for i=2:n
                z(i) = z(i-1)+ x(i);
      end
      plot(z,'Color', c(mod(l,7)+1),'LineWidth',2);
end
```

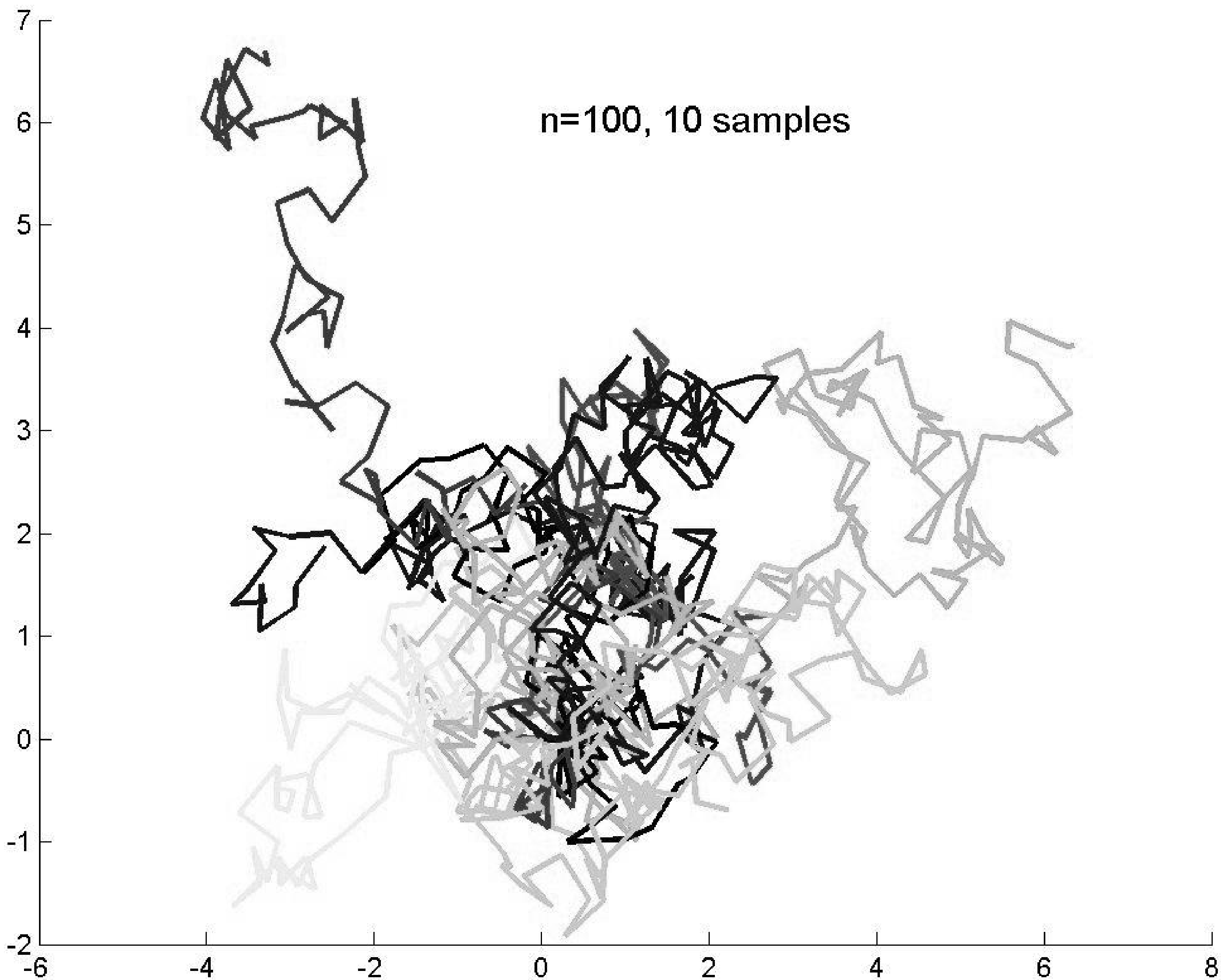
# Random Walk: 2D

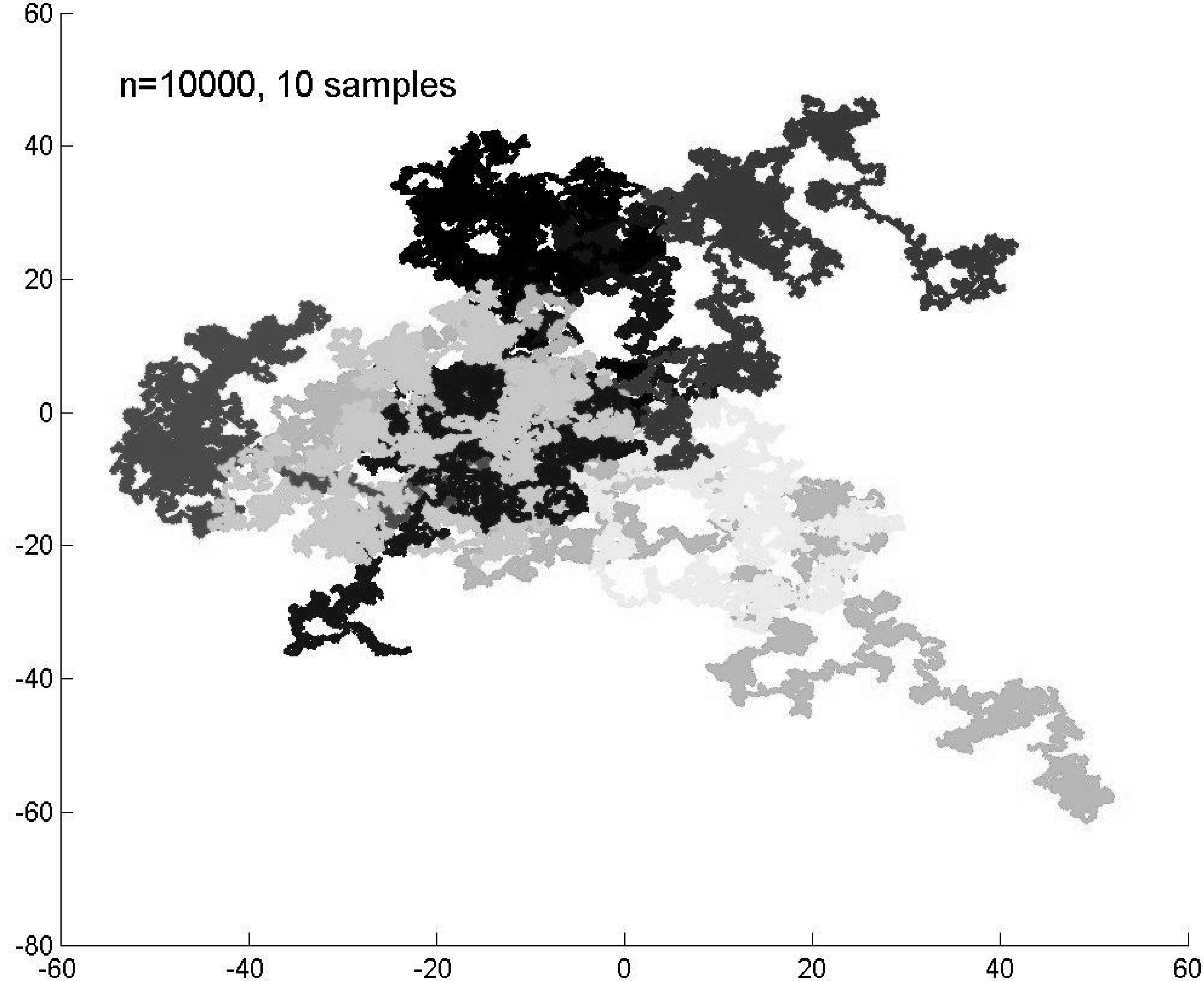
Let the position of a particle at time $k = 1, 2, \ldots$ be denoted by $\mathbf{X_k} \in R^2$
We assume that the particle changes its position according to

$$P(\mathbf{X_{k+1}} = \mathbf{X_k} + \alpha h) = \frac{1}{4} \qquad \alpha = \{(0,1),(0,-1),(1,0),(-1,0)\}$$

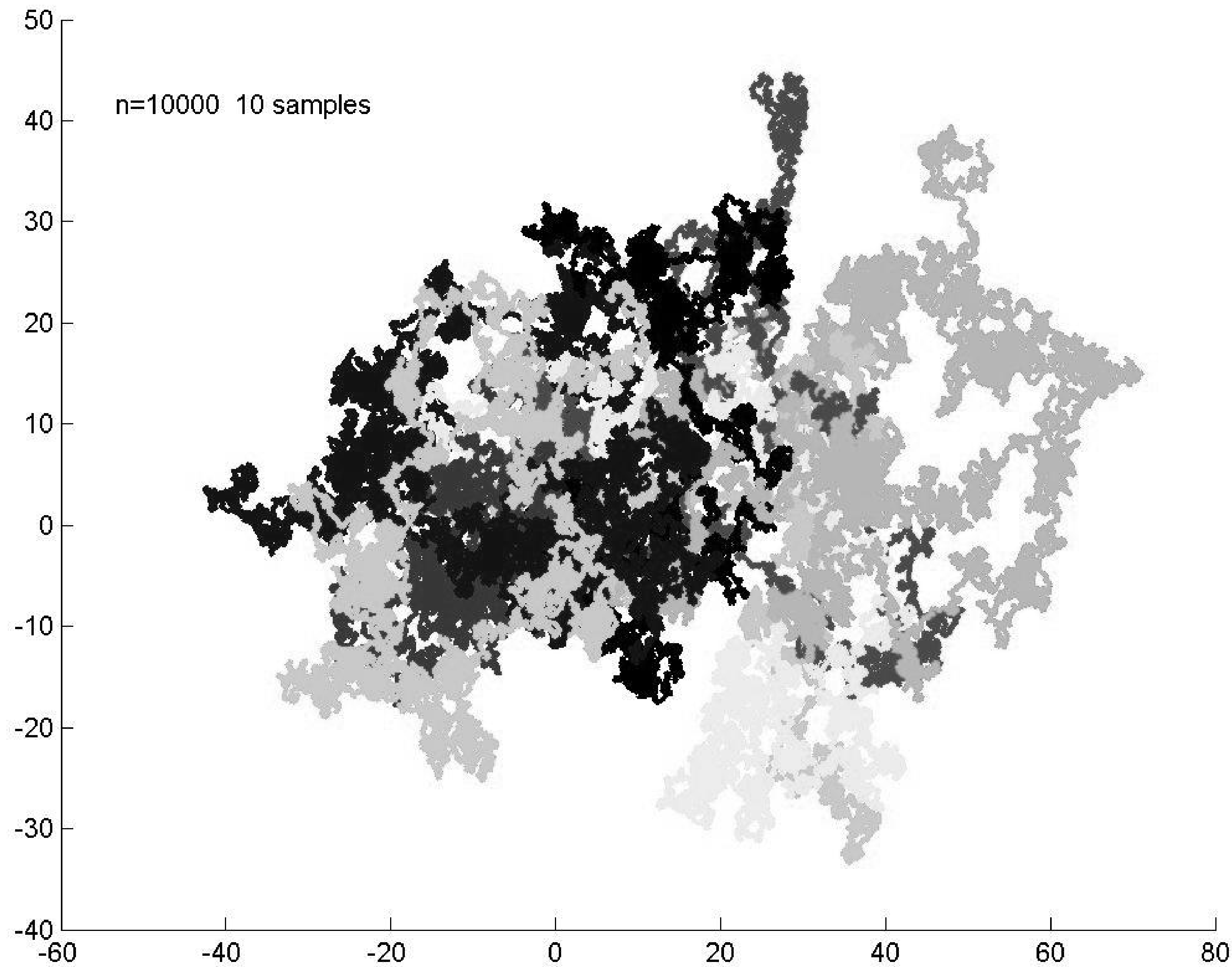
Sample paths:


n=100, 10 samples
6
4
2
0
-2
-4
-6
-6
-4
-2
0
2
4
6
8


n=100, 10 samples
7
6
5
4
3
2
1
0
-1
-2
-6
-4
-2
0
2
4
6
8



n=10000, 10 samples
60
40
20
0
-20
-40
-60
-80
-60
-40
-20
0
20
40
60


n=10000  10 samples
50
40
30
20
10
0
-10
-20
-30
-40
-60
-40
-20
0
20
40
60
80

# randWalk2D.m

```
%random walk in 2D, uniformly dist jumps
n=100000;
figure;
hold on;
c = ['r','g','b','c','m','y','k'];
for l=1:10
    x = (rand(n,1) - 0.5*ones(n,1));
    y = (rand(n,1) - 0.5*ones(n,1));
   z1 = zeros(n,1);
   z2 = zeros(n,1);
   for i=2:n
      z1(i) = z1(i-1) + x(i);
      z2(i) = z2(i-1) + y(i);
   end
plot(z1,z2,'Color',
c(mod(l,7)+1),'LineWidth',2);
end
```

Can change the jump distribution to be normal (randn),

Or just +=1 jumps

Code for this is on the web page!

# Some Questions

Starting from the origin, what is the probability that a particle will reach a specified location (area), during a given time interval?

1. What is the probability of a molecule to reach the nucleus using random walk, starting from the membrane?

2. Can a macrophage find a bacterium, using just a random walk?

Starting from the origin, what is the average time that it takes a particle to reach a specific location (area)?

1. How long on the average it takes a signaling protein to reach the nucleus, starting from the membrane, if only random walk is used?

# Questions (cont.)

How do populations of particles behave?

How do particles (viruses, bacteria) spread (based on random walk only)? Can this explain actual observations?

Given observations (real data), can we approximate the process by some random walk?

Are there more random walks that are interesting to study? Yes!!!

# Master Equation

**Back to 1D random walk.** (more general than before). Let $\alpha \leq 1/2$ and assume the following transition probabilities

$$P(X_{k+1} = X_k \pm h) = \alpha \qquad P(X_{k+1} = X_k) = 1 - 2\alpha$$

Position j correspond to a distance jh from the origin.

A particle at position j and time n, was at position j-1, j+1,j at time n-1. Let $P_j^n$ be the probability of finding a particle at position j and time n . The following equation is satisfied,

$$P_j^n = (1 - 2\alpha)P_j^{n-1} + \alpha(P_{j-1}^{n-1} + P_{j+1}^{n-1})$$

# Master Equation cont.

If the domain is bounded we need to supply also equations at the boundaries.

Particles may have different behavior

absorption by boundary

reflection from boundary

periodic domain (a particle that leaves form $j=N$, comes back at $j=1$)

# Master Equation: matrix form

Let $P^n = [P_1^n, P_2^n, \dots]^T$. The evolution of $P^n$ can be written as

$$P^n = AP^{n-1}$$

where

$$A = \begin{bmatrix} & \dots & & & & \\ & 1-2\alpha & \alpha & 0 & \dots & \\ \dots & \alpha & 1-2\alpha & \alpha & \dots & \dots \\ & 0 & \alpha & 1-2\alpha & \alpha & \\ & \dots & & & & \end{bmatrix}$$

# Boundary conditions

Need to specify equations for $P_1^n, P_N^n$

Based on the behavior of particles at the end of the domain.

(i) Infinite domain: no need to introduce B.C. (matrix is infinite)
(ii) Finite domain

(1) absorbing boundary (particle disappear)

$$P_1^n = (1-2\alpha)P_1^{n-1} + \alpha P_2^{n-1}$$
$$P_N^n = (1-2\alpha)P_N^{n-1} + \alpha P_{N-1}^{n-1}$$

# cont..

Matrix A:

$$A = \begin{bmatrix} 1-2\alpha & \alpha & 0 & \dots & 0 \\ \alpha & 1-2\alpha & \alpha & 0 & \dots \\ \dots & & & & \\ \vdots & & & & \\ & & & & \end{bmatrix}$$

# B.C. (cont.)

(2) reflecting boundary

$$P_1^n = (1-\alpha)P_1^{n-1} + \alpha P_2^{n-1}$$
$$P_N^n = (1-\alpha)P_N^{n-1} + \alpha P_{N-1}^{n-1}$$

$$A = \begin{bmatrix} 1-\alpha & \alpha & 0 & \dots & 0 \\ \alpha & 1-2\alpha & \alpha & 0 & \dots \\ \dots & & & & \\ \vdots & & & & \end{bmatrix}$$

# B.C. cont.

(3) periodic B.C.

$$P_1^n = (1-2\alpha)P_1^{n-1} + \alpha P_2^{n-1} + \alpha P_N^{n-1}$$

$$P_N^n = (1-2\alpha)P_N^{n-1} + \alpha P_{N-1}^{n-1} + \alpha P_1^{n-1}$$

First and last rows in A change

$$A = \begin{bmatrix} 1-2\alpha & \alpha & 0 & \dots & \alpha \\ \alpha & 1-2\alpha & \alpha & 0 & \dots \\ \dots & & & & \\ \vdots & & & & \\ & & & & \end{bmatrix}$$

# Evolution

Where will a particle be at time n, if at time n=0 it was at position j=0?

We will answer this it in two ways:

simulation using particles

here we have to modify our MATLAB code randWalk1D.m to treat the different boundaries.

using the master equation

here we will use sparse matrices in MATLAB for efficient calculation. Essentially we need to apply A to the initial data n times.

# Some Matrix Algebra

We want to understand the evolution given by

$$P^{n+1} = AP^n$$

Assume that A has a complete set of eigenvectors,

$$Au_j = \lambda_j u_j \quad j = 1, \ldots, N$$

Decomposing

$$P^0 = \sum_{j=1}^{N} a_j u_j$$

And using

$$A^k u_j = \lambda_j^k u_j$$

Gives

$$P^n = \sum_{j=1}^{n} a_j \lambda_j^n u_j$$

Asymptotic behavior of our probability function depends on the largest eigenvalues, and eigenvectors of A!

Boundary conditions play important role here (they affect the eigenvalues).

# Back to Simulation

Place many particles at j=0, and monitor their evolution. Define

$$m_j^k = \sum_l \delta_{j,X_l^k} \qquad \delta_{i,j} = \begin{cases} i = j & 1 \\ i \neq j & 0 \end{cases}$$

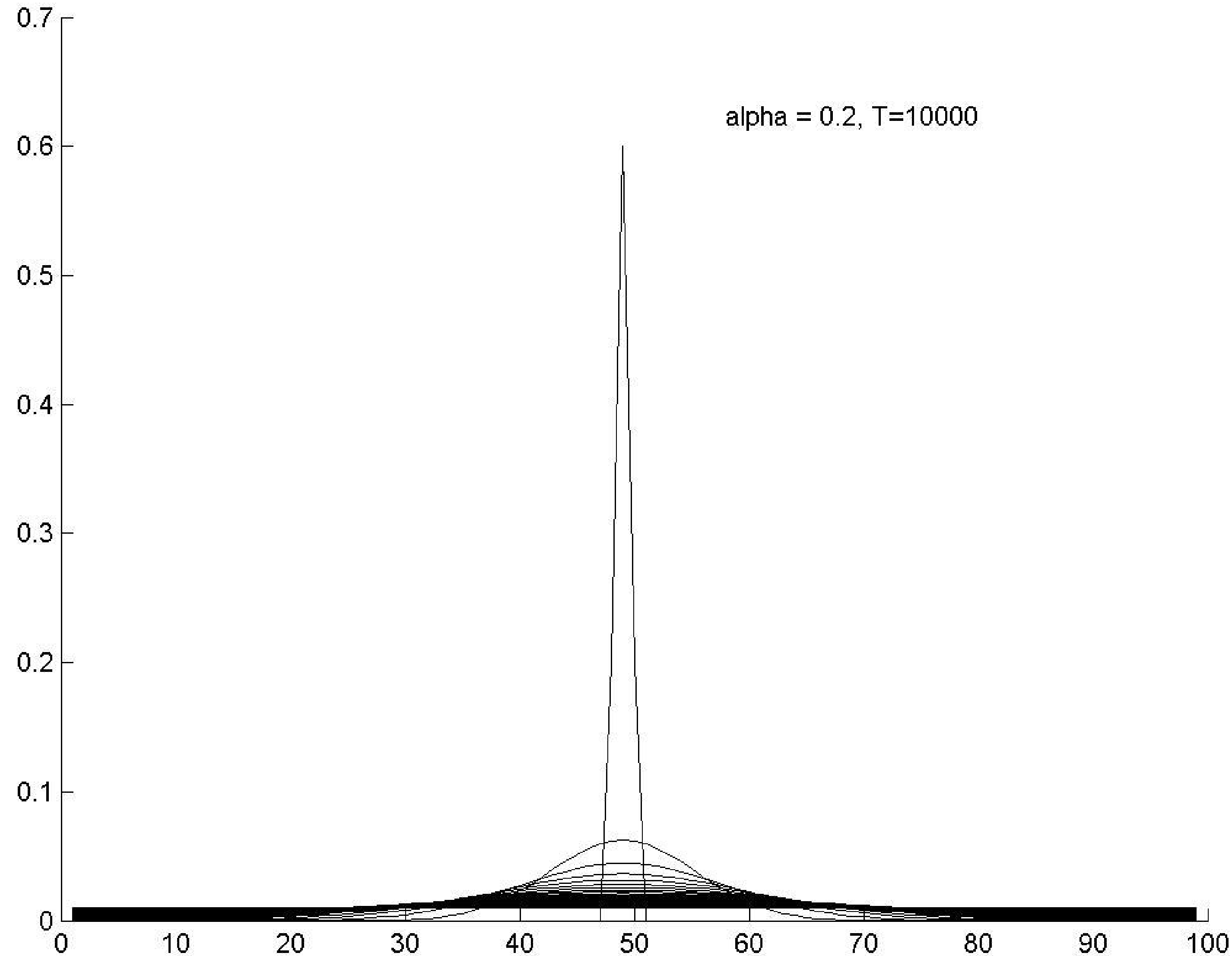
$m_j^k$ Total number of particles at position j, time k

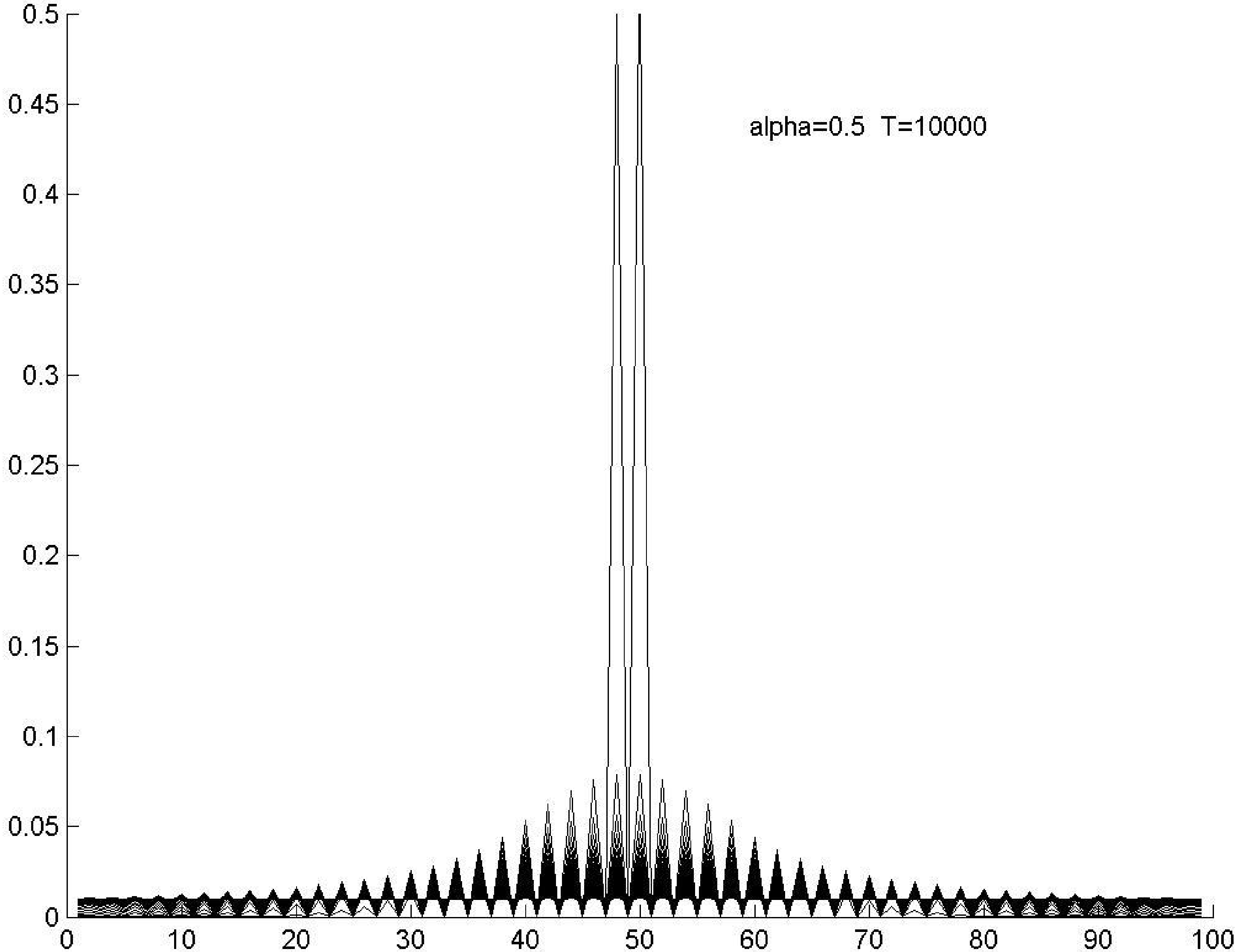
$X_j^k$ Position of particle l, at time k

We are interested in the evolution of $m_j^k$

# Evolution using Master Equation

```
% evolution.m. : evolution of probability density func.
nMid = 49;    n=2*nMid + 1;
alpha = 0.4;   figure;      hold on;
c = ['r','g','b','c','m','y','k'];
vs = (1.0-2.0*alpha)*ones(1,n);
D = sparse(1:n,1:n,vs,n,n);
E = alpha*sparse(2:n,1:n-1,ones(1,n-1),n,n);
S = E+D+E';
v = zeros(n,1);    v(nMid) = 1.0;
for i=1:10000
    v = S*v;
    if mod(i,100) == 1
            plot(v, 'Color', 'k');
    end
end
```


alpha = 0.2, T=10000
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
80
90
100


alpha=0.5  T=10000
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
10
20
30
40
50
60
70
80
90
100

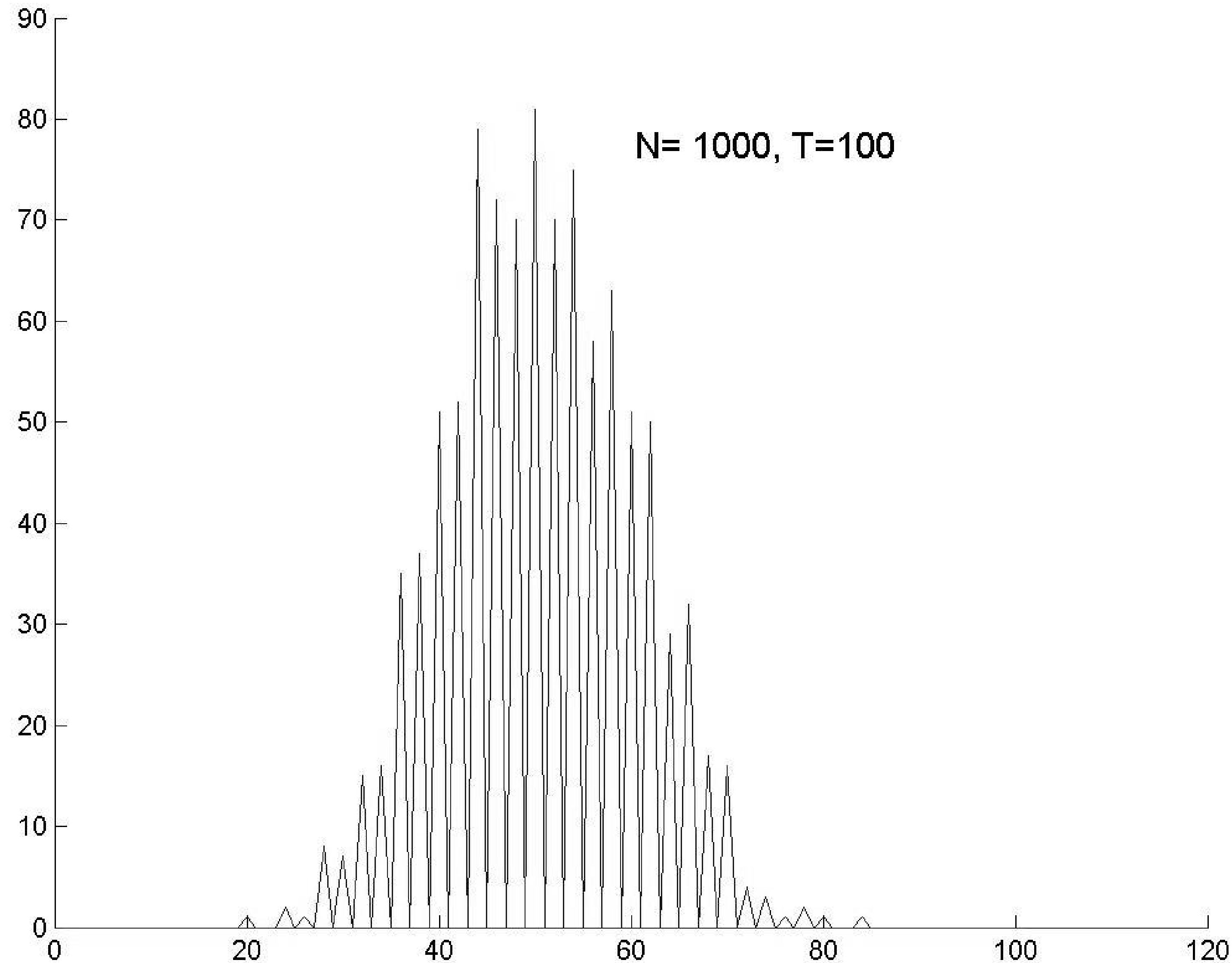
# Questions

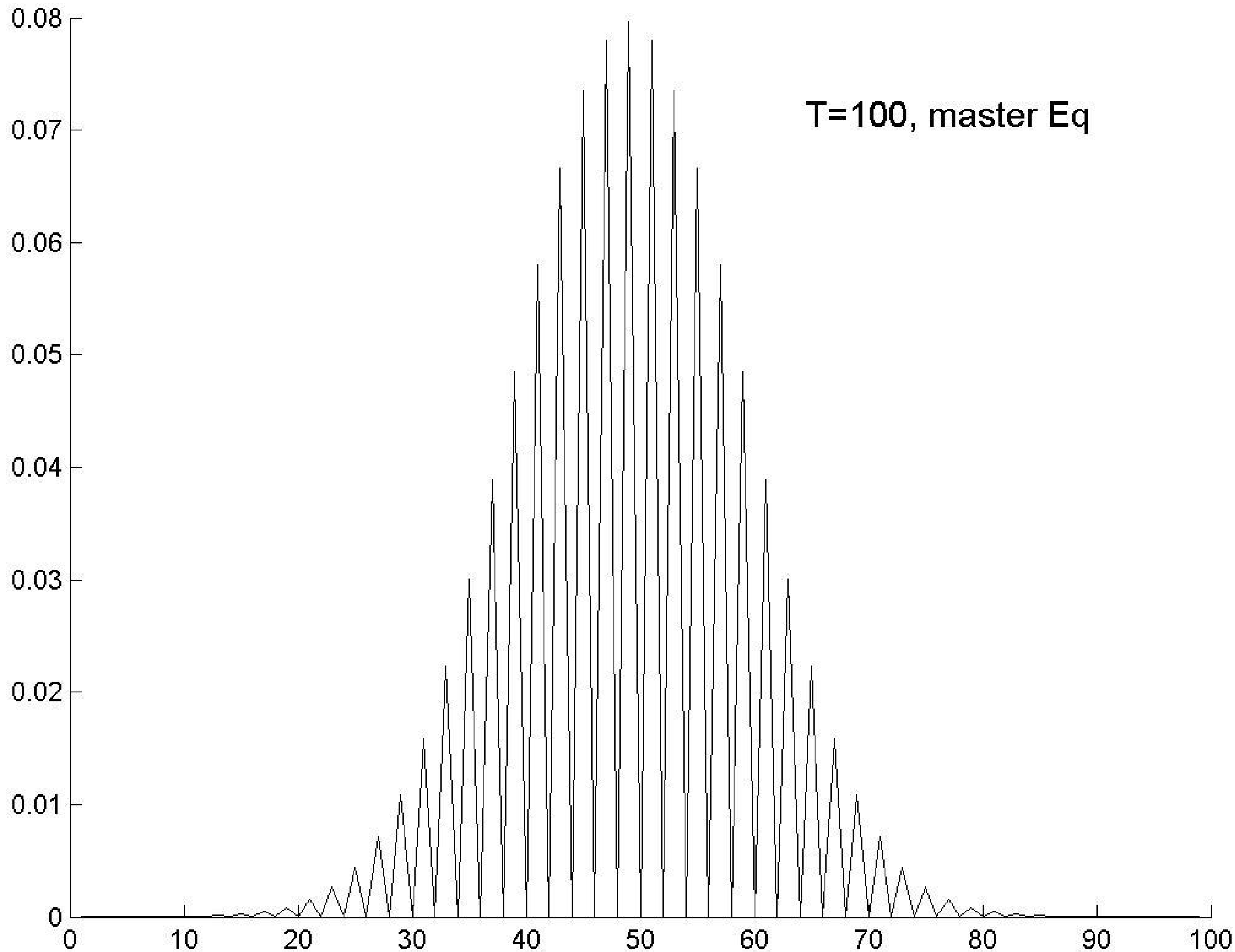
The computation shows the evolution of the probability density function, starting from a configuration where all particles are at the origin.

1. Where do the oscillation come from?
2. Can we predict the appearance of such oscillations?
3. What is the effect of different BC?
   - absorbing BC
   - reflecting BC

# Particle Simulation

```
% randWalk1D2.m
T=150;      %number of time steps
nMid=50;
n=2*nMid+1;
figure;
hold on;
h = zeros(n,1);
c = ['r','g','b','c','m','y','k'];
for l=1:2000                        %number of particles
    x = (rand(T,1) - 0.5*ones(T,1));
    indP = find(x>0);
    indM = find(x<=0);
    x(indP) = 1;
    x(indM) = -1;
    z = nMid;
    for i=1:T
     z = z + x(i);
    end
    h(z) = h(z) + 1;
end
plot(h)
```


N= 1000, T=100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100
120


T=100, master Eq
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
10
20
30
40
50
60
70
80
90
100

# Probability to Capture

Suppose that we start a random walk at the origin. What is the probability of reaching a target at x=b, by a prescribed time T?

Assume also that a particle that reach x=b does not move anymore.

Using a particle simulation: start with many particle at the origin and simulate until time T. Check the fraction of particles that reached b.

Sounds easy

Can we do it using the master Equation? We need to express the fact that a particle that reaches the target disappears.

# Time to Capture

Suppose that we start a random walk at the origin. What is the mean time to reach a target at $x=b$?