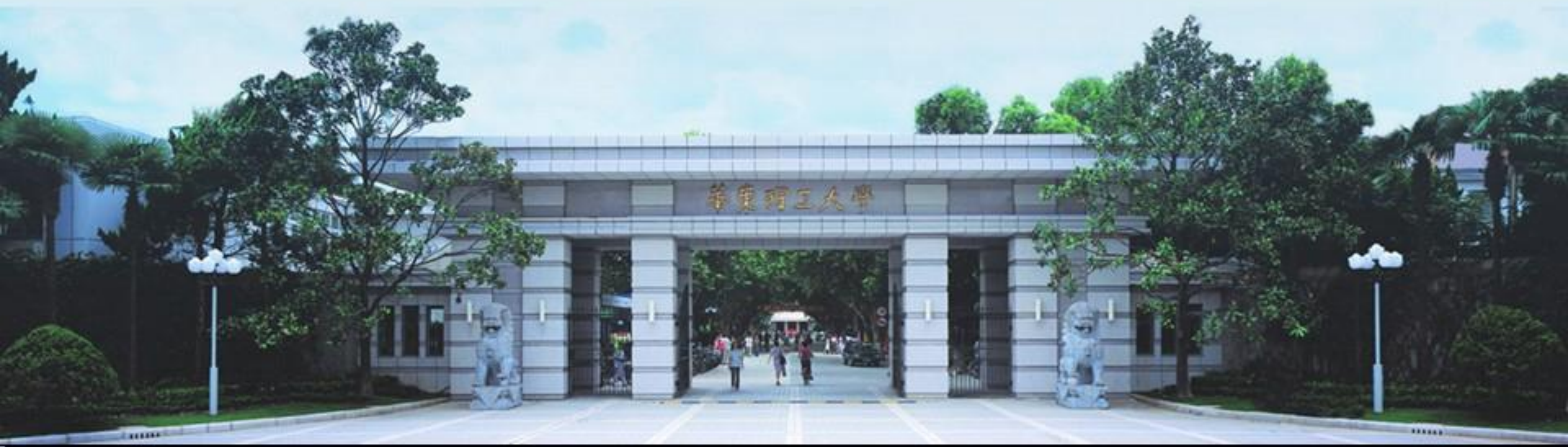
分子模拟概述

Overview of Molecular Simulation



计算机模拟的作用

按照统计力学原理，系统的宏观性质是相应微观量的系综平均值或时间平均值。计算机分子模拟的主要内容就是为模型系统提供大量的微观状态，以及对应的微观量。模拟得到的结果也称为“机器实验数据”。这类数据的突出应用价值有：

1. 检验与改进理论

应用统计力学理论解决实际问题时，为了克服数学上的困难，常常需要假设一定的物理模型。在此基础上进行数学求解，建立数学模型。在数学求解过程中，不可避免地需要引入简化，这种简化是否合理？可以通过计算机模拟结果与理论模型结果的比较进行检验。

计算机模拟的作用

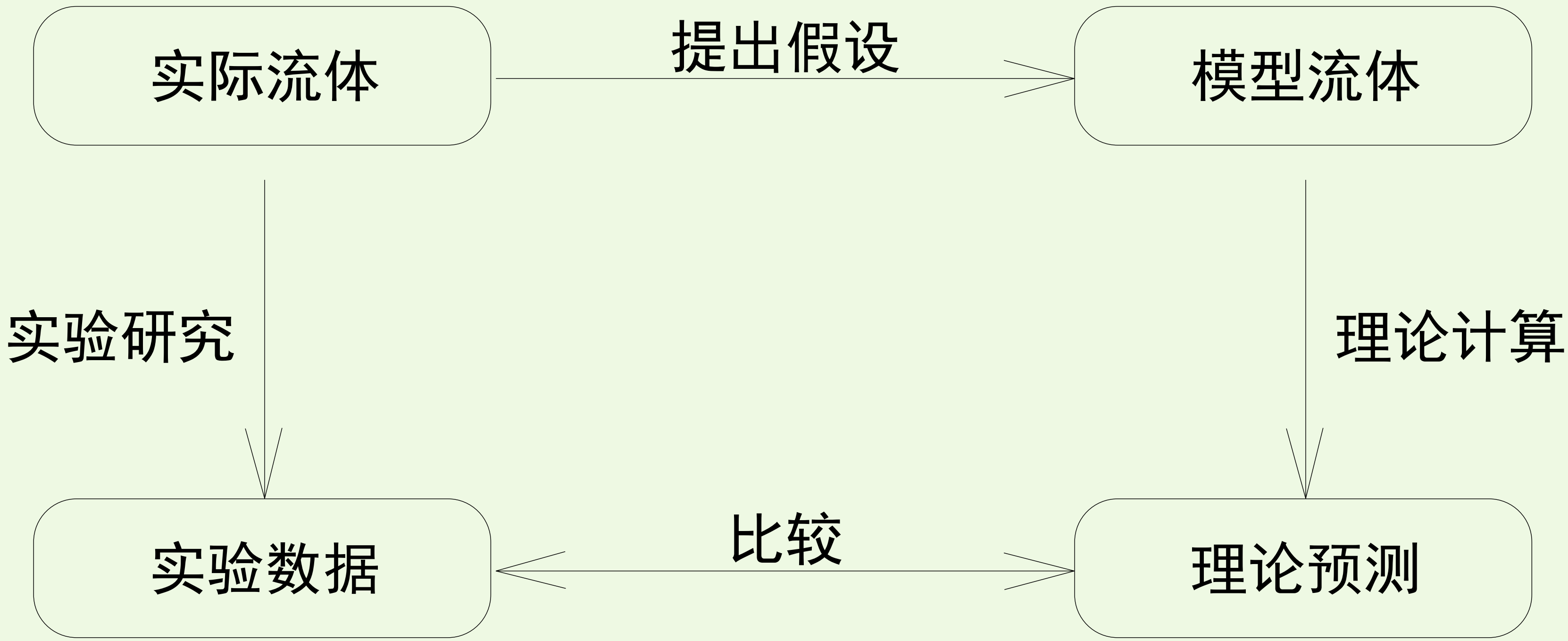
2. 检验与改进物理模型

物理模型是从实际中抽象出来的，是实际系统本质的近似反映。这种物理模型是否合理？可以通过计算机模拟结果与实验结果的比较进行检验。

3.提供极端条件下的系统性质

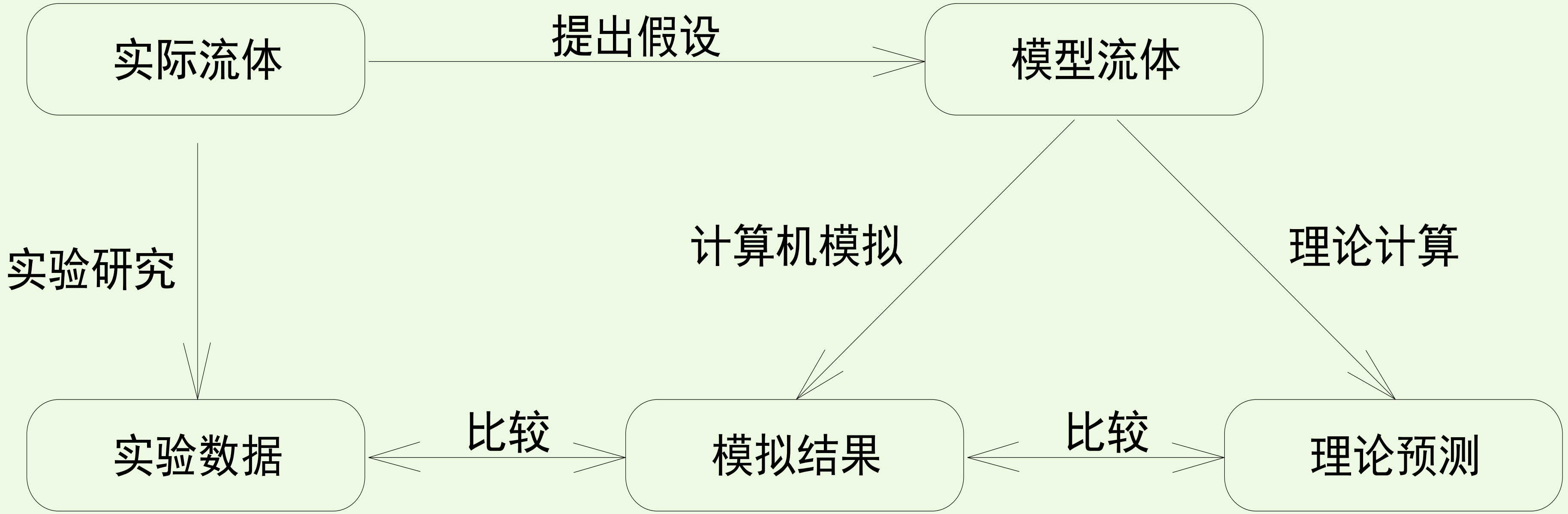
某些极端条件下系统的性质很难在实验室进行测定，可以采用计算机模拟的方法获得。如：地层中高温高压下油气的状态、温度敏感性物质的性质、剧毒物质、放射性物质等的性质。

计算机模拟的作用



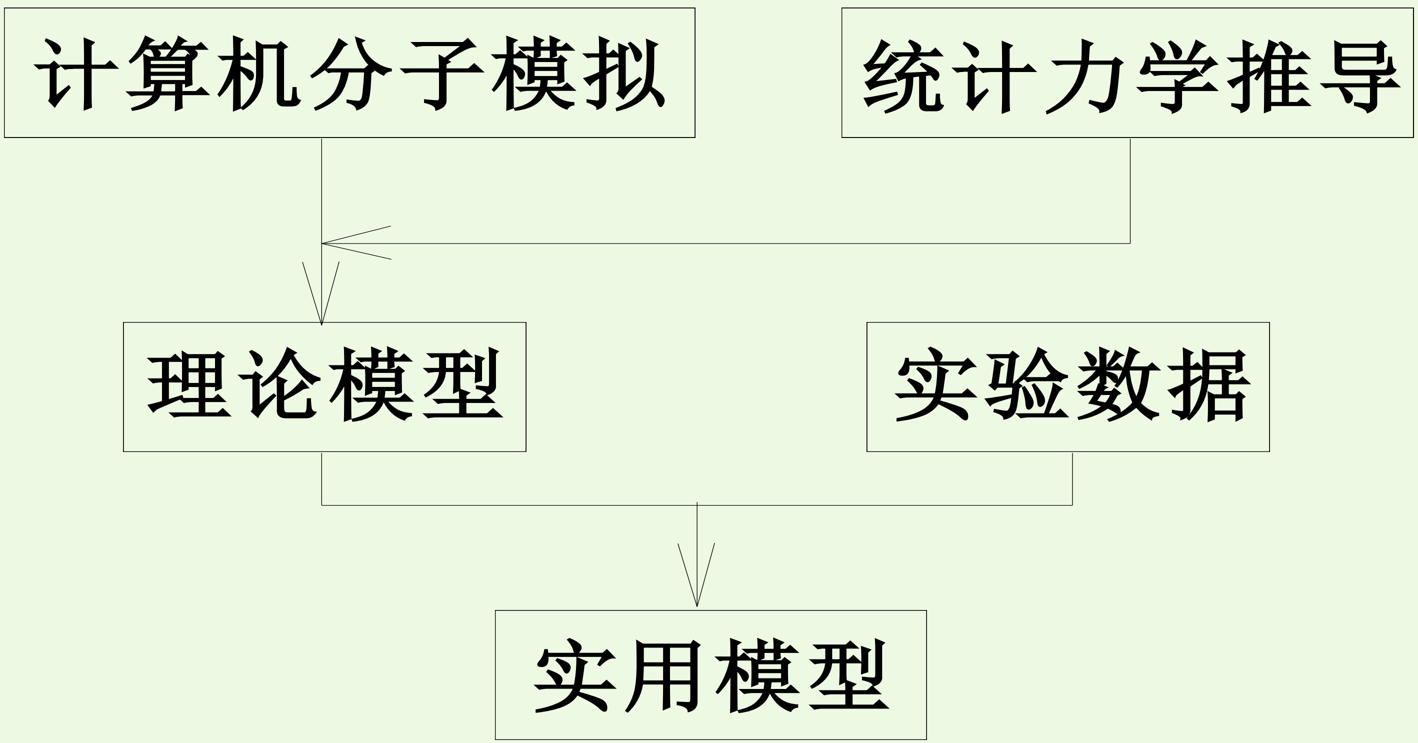
传统的分子热力学研究方法

计算机模拟的作用



用计算机模拟检验模型和检验理论

计算机模拟的作用



现代分子热力学研究方法

计算机分子模拟的分类

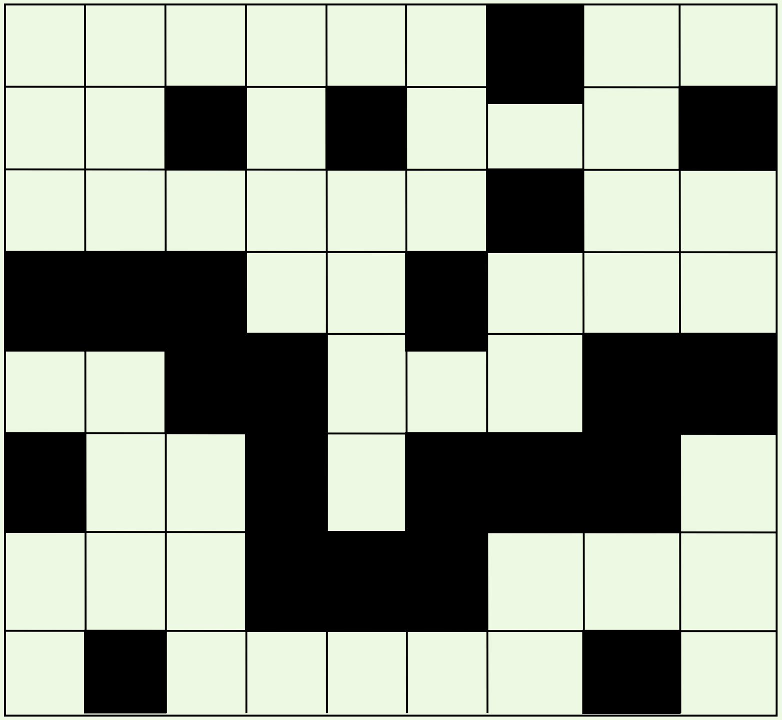
分子模拟的方法大致可分为两类：随机性模拟和确定性模拟。

（1）随机性模拟：在模拟中使用了随机数。通常只能模拟系统的平衡性质，不能模拟分子的真实运动轨迹，过程不能反演。最常用的有Monte Carlo (MC)模拟。最终可得到系统中各个粒子的平衡位置。

（2）确定性模拟：按经典运动规律来预测分子的运动轨迹，过程可反演。既可以模拟系统的平衡性质，也能够模拟系统的动力学性质。最常用的有分子动力学(Molecular Dynamics, MD)模拟。最终可得到系统中各个粒子的位置和动量（速度）。

随机性模拟的简单示例—逾渗问题

随机性模拟的一个简单例子是逾渗问题：为简单起见，考虑一个二维方格，每个格点可以被占也可以为空，格点被占的概率为$p \in [0,1]$，空的概率为1-p。如果这个二维方格为无穷大，则当p等于什么数值的时候，可以从盒子的一边出发，通过被占格点到达对边，即形成逾渗通道？



随机性模拟的简单示例—逾渗问题

为了求解上述逾渗问题，我们可以这样进行：先产生一个L×L的有限方格，然后以p的概率随机填充这些格点，最后观察是否形成逾渗通道。

从算法的角度看，它分两个步骤：**第一步**是产生一个被随机占住的二维方格；**第二步**是检验逾渗通道的存在与否。

随机性模拟的简单示例—逾渗问题

第一步可以这样实现：

1、建立一个L×L的二维数组，数组的元素全部置0。

2、然后访问全部格点。对于每个访问到的格点，**产生一个均匀分布的随机数** $R \in [0,1]$，如果R小于p，则认为该格点被占，其对应的元素置为1。这样产生的数组称为一个实现或位形。如果用FORTRAN语言，则程序为：

随机性模拟的简单示例—逾渗问题

```
      subroutine percolation(lattice, L, P)
      real P
      integer L, lattice(1:L, 1:L)
      do 20 I=1, L
         do 10 J=1, L
            lattice(I,J)=0
10    continue
20    continue
      do 40 I=1, L
         do 30 J=1,L
            R=uniform()
            If (R.LT.P) lattice(I,J)=1
30     continue
40     continue
       return
       end
```

产生一个二维数组，全部元素置0

访问全部格点，以*p*的概率将对应的元素置1

注：uniform()为均匀分布的随机数发生器

伪随机数

在随机性模拟中，需要大量的随机数，通常是在[0,1]范围均匀分布的随机数，其它分布的随机数往往可以从均匀随机数导出。

产生随机数的方法可以是物理的，也可以是数学的。物理方法主要是计算机的固有噪声，所产生的随机数质量很高，但无法再现，因此无法对计算结果进行验证，而且需要配备特殊的设备。数学方法主要是利用某些合适的递推公式，当初值给定后，产生的随机数是确定的，并非完全独立；另外由于受计算机字长的限制，递推公式产生的随机数迟早都会出现周期性的循环现象。由于这种随机数并非真正的随机数，所以称为伪随机数。实际使用中需要注意所使用的随机数不能超过伪随机数的循环周期，否则需要进行特殊处理。一般高级语言都有随机数发生器(函数语句)，在程序中直接调用即可。下面是一个伪随机数发生器的例子，其循环周期大于2×10^{18}。

附：伪随机数发生器程序RAN2(idum)

```
      FUNCTION ran2(idum)
      INTEGER idum, IM1, IM2, IMM1, IA1, IA2, IQ1, IQ2, IR1, IR2, NTAB, NDIV
      REAL ran2, AM, EPS, RNMX
      PARAMETER (IM1=2147483563, IM2=2147483399, AM=1./IM1, IMM1=IM1-1,
     *              IA1=40014, IA2=40692, IQ1=53668, IQ2=52774, IR1=12211,
     *              IR2=3791, NTAB=32, NDIV=1+IMM1/NTAB, EPS=1.2e-7, RNMX=1.-EPS)
       INTEGER idum2, j, k, iv(NTAB), iy
      SAVE iv, iy, idum2
      DATA idum2/123456789/, iv/NTAB*0/, iy/0/
      if (idum.le.0) then
         idum=max(-idum,1)
         idum2=idum
         do 11 j=NTAB+8,1,-1
            k=idum/IQ1
            idum=IA1*(idum-k*IQ1)-k*IR1
            if (idum.lt.0) idum=idum+IM1
            if (j.le.NTAB) iv(j)=idum
          enddo 11
          iy=iv(1)
     endif
  k=idum/IQ1
  idum=IA1*(idum-k*IQ1)-k*IR1
      if (idum.lt.0 ) idum=idum+IM1
      k=idum2/IQ2
      idum2=IA2*(idum2-k*IQ2)-k*IR2
      if (idum2.lt.0) idum2=idum2+IM2
      j=1+iy/NDIV
      iy=iv(j)-idum2
      iv(j)=idum
      if (iy.lt.1) iy=iy+IMM1
      ran2=min(AM*iy, RNMX)
      return
      END
```

随机数发生器的均匀性检验

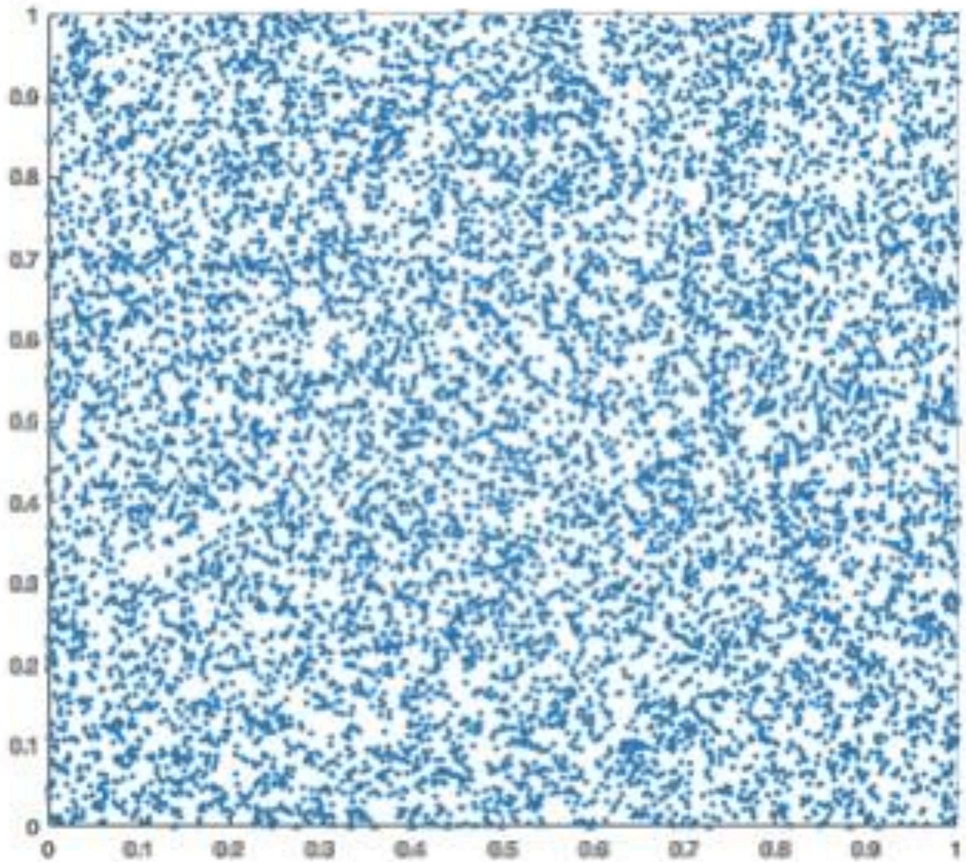
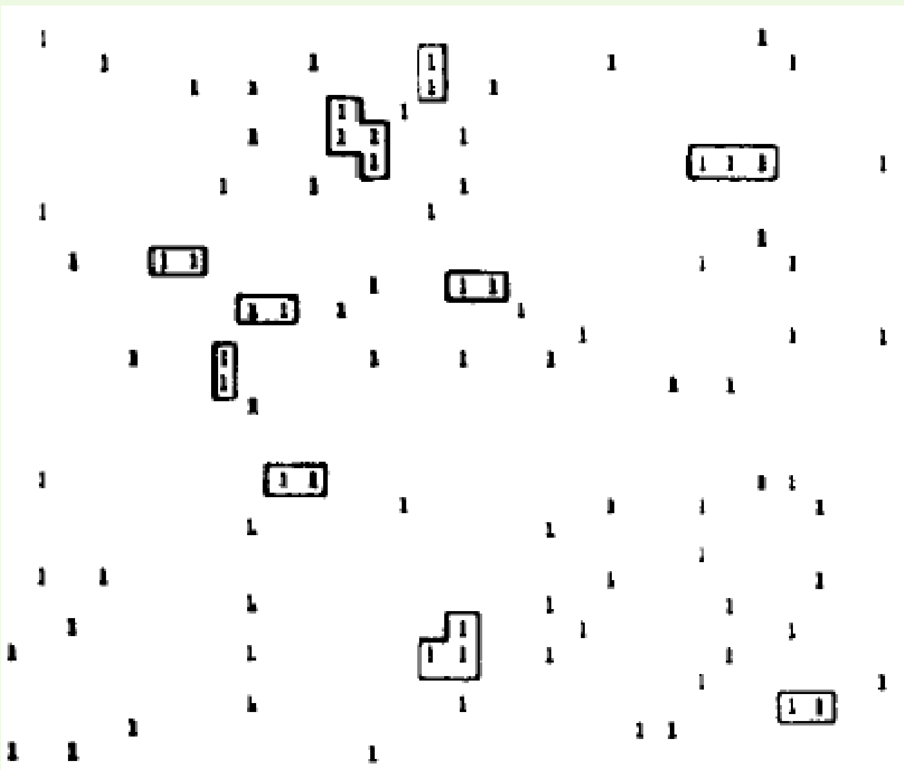


图 1: x 和 y 在 (0,1) 区间内任意取值作为坐标，描点绘图。

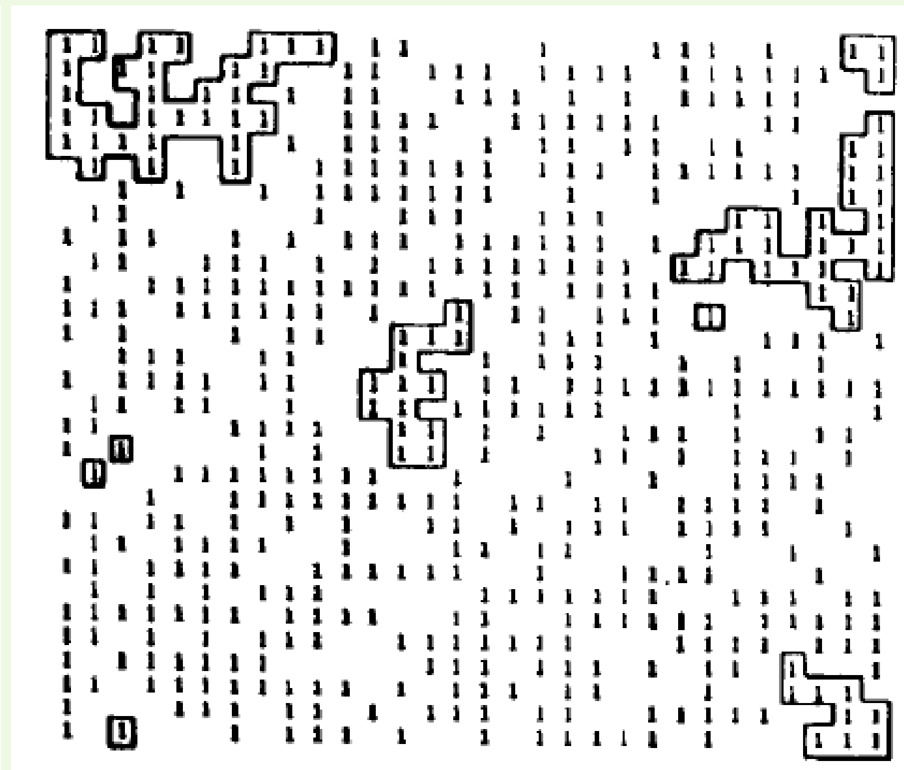
随机数均匀性检验效果图

随机性模拟的简单示例—逾渗问题

第二步的实现比较困难。对于二维的小系统，我们可以将第一步产生的位形打印出来，然后观察是否有逾渗通道。



p=0.1



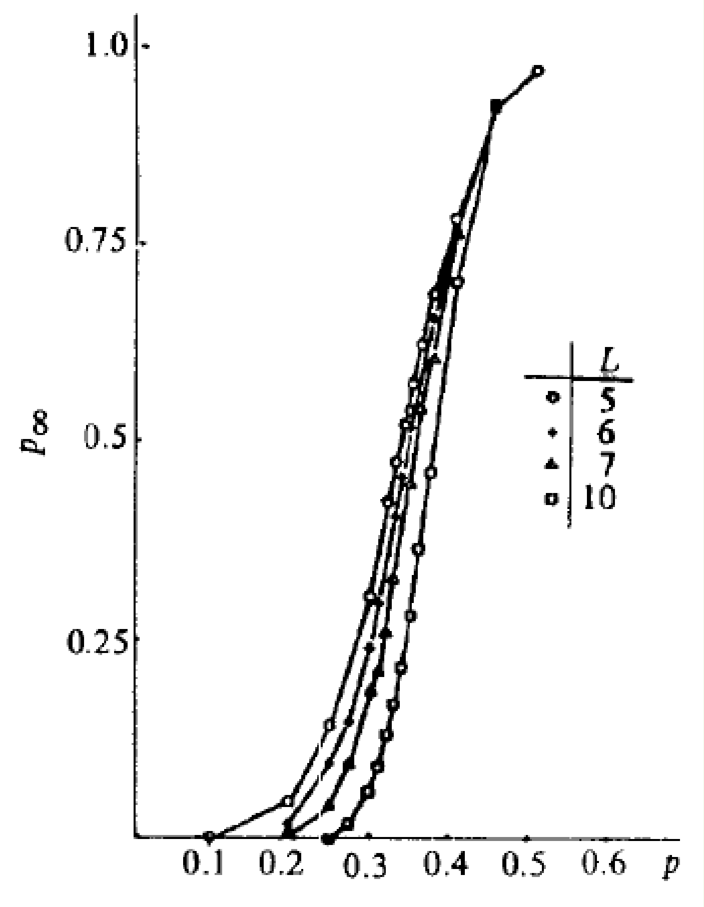
p=0.6

随机性模拟的简单示例—逾渗问题

由于某一个格点是否被占是随机的，在相同的p值条件下产生的不同位形中，有的存在逾渗通道，有的不存在逾渗通道。为了定量表达逾渗通道存在的可能性，我们可以在一定的p值条件下，重复产生N个位形，并进行观察，设其中有N^*次存在逾渗通道，

则逾渗概率为： $p^{\infty} = N^* / N$

随机性模拟的简单示例—逾渗问题



三维逾渗问题的模拟结果

将p^∞ 对p作图，可以得到逾渗概率曲线。可以发现，对于有限的系统，曲线是逐渐变化的。如果对不同大小的系统进行模拟，我们可以发现，逾渗概率曲线的斜率是随系统大小而变化的，系统愈大，曲线愈陡峭。

随机性模拟的简单示例—逾渗问题

逾渗问题的模拟存在三个困难：

1、逾渗概率的模拟结果与系统大小有关。

2、逾渗概率的确定与样本数目有关。

3、随机数是否存在偏差？

上述三个问题在所有随机模拟方法中都存在。

随机性模拟的简单示例—数值积分

逾渗问题本身是一个随机性问题，所以可以用随机模拟方法加以解决。还有一些问题，它本身是一个确定性问题，但我们可以将它转化为随机性问题来予以解决，如定积分问题。

考虑一维定积分：

$$\theta = \int_0^1 f(x)\mathrm{d}x = \int_0^1 \frac{e^x - 1}{e - 1}\mathrm{d}x$$

由概率论可知， θ可以看作是矩形分布(一定范围内的均匀分布)上函数$f(x)$的数学期望（平均值）。根据这个模型，抽样在[0,1]均匀分布上进行，然后建立估计量：

随机性模拟的简单示例—数值积分

$$\bar{\theta} = \frac{1}{n}\sum_{i=1}^{n} f_i = \frac{1}{n}\sum_{i=1}^{n} f(\xi_i)$$

可以证明，$\bar{\theta}$ 是θ的无偏估计。其程序可编写如下：

```
      subroutine integration(N, Fx, Cita)
      real Cita
      integer N
      external function Fx
     Cita=0.0
      do 10 I=1,N
          R=uniform()
          Fi=Fx(R)
          Cita=Cita+Fi
10     Continue
       Cita=Cita/N
       return
      end
```

```
function Fx(x)
real x
Fx=(exp(x)-1.0)/(exp(1.0)-1.0)
return
end
```

这个问题的真值为θ=0.418。我们可以发现模拟结果与真值的偏差与样本数N有关。

随机性模拟的简单示例——数值积分

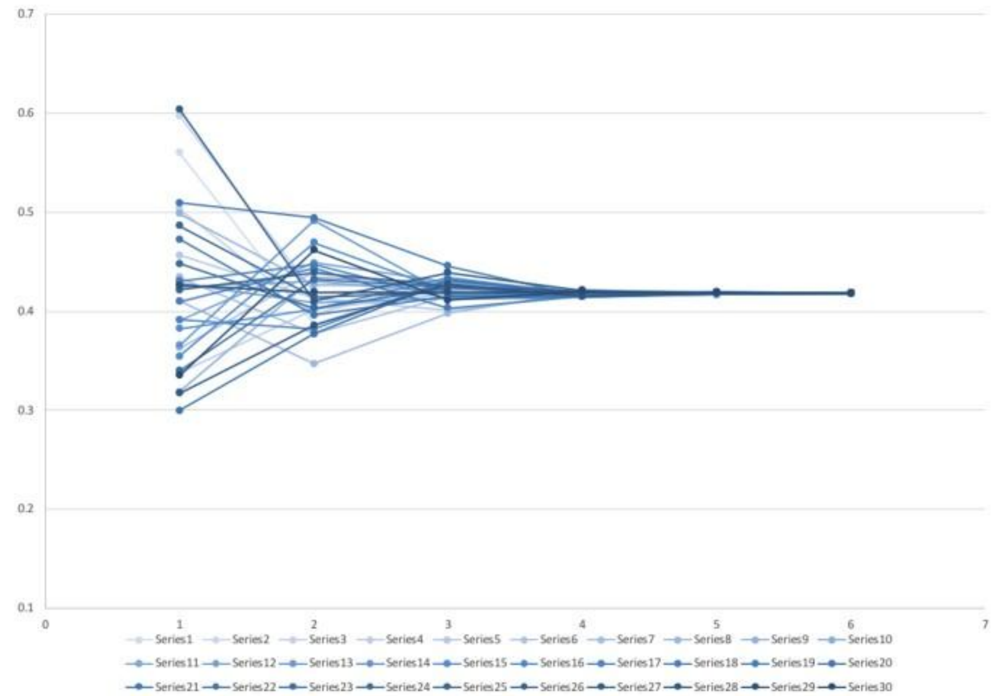


图 2-1：利用随机数求定积分 30 次，绘制得到近似估计值与绘散点趋势图。

注：横轴为随机取值次数的对数，纵轴为计算定积分结果，同一颜色代表一次运行结果，真值=0.418023。运行过程截图详见：附录-2。

模拟结果与真值的偏差与样本数*N*有关！

随机性模拟的简单示例—数值积分

思考题：

请采用随机性模拟方法求解下列定积分问题。

$$\theta=\int_0^1\int_0^1\left[\frac{e^{(x^2+y^2+xy)}+xy}{xy+1}\right]\mathrm{d}x\mathrm{d}y$$

随机性模拟的简单示例—数值积分

要求： 利用随机数计算二元定积分的确定性问题。

积分函数：

$$S = \int_0^1 \int_0^1 \left[\frac{e^{(x^2+y^2+xy)} + xy}{xy+1} \right] dxdy$$

结果： 计算其积分结果，$\bar{S} \approx 2.47$ 。

数学处理：

$$\bar{S} = \int_0^1 \int_0^1 f(x,y)dxdy = \frac{1}{n}\sum_{i=1}^{n} f(\xi_{x_i}, \xi_{y_i})$$

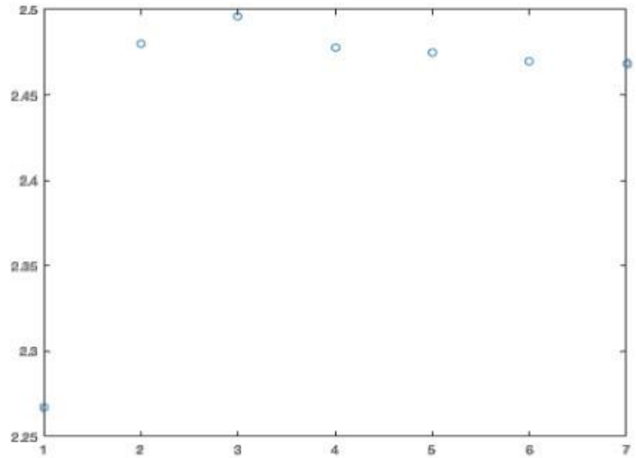
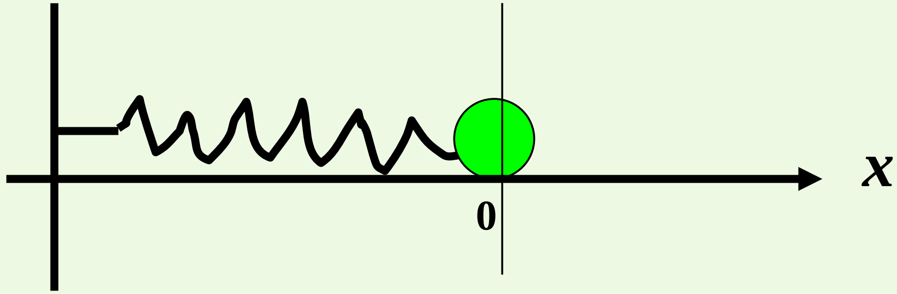


图 3：利用随机数求定积分，绘制得到近似估计值。

确定性模拟的简单示例——一维谐振子



问题：计算在弹性力位势中粒子的运动轨迹

运动轨迹包含两部分内涵：

（1）粒子的位置　　　　（2）粒子的动量（速度）

描写粒子运动的哈密顿量(动能与势能之和)为：

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2$$

确定性模拟的简单示例——一维谐振子

粒子的初始位置和速度必须给定： $(x(0), p(0))$

$$H = \frac{p^2}{2m} + \frac{1}{2}kx^2$$

我们取运动方程的哈密顿形式，即：

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H}{\partial p} = \frac{p}{m} \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial x} = -kx$$

我们的目的是要通过求解运动方程，获得粒子的运动轨迹$(x(\mathrm{t}), p(\mathrm{t}))$ 。我们用数值积分的方法求解这个问题。首先我们用差分代替微分，差商代替导数，即：

确定性模拟的简单示例—一维谐振子

$$\frac{\mathrm{d}f}{\mathrm{d}t}=\frac{f(t+\Delta t)-f(t)}{\Delta t}=\frac{f(t+h)-f(t)}{h}$$

h*=△*t **为时间步长。将上述方法引入运动方程，得：**

$$\frac{\mathrm{d}x}{\mathrm{d}t}\approx\frac{1}{h}[x(t+h)-x(t)]=\frac{p(t)}{m} \qquad \frac{\mathrm{d}p}{\mathrm{d}t}\approx\frac{1}{h}[p(t+h)-p(t)]=-kx(t)$$

$$\Downarrow$$

$$x(t+h)=x(t)+\frac{hp(t)}{m} \qquad p(t+h)=p(t)-hkx(t)$$

确定性模拟的简单示例——一维谐振子

$$x(t+h)=x(t)+\frac{hp(t)}{m} \qquad p(t+h)=p(t)-hkx(t)$$

给出与给定能量值不矛盾的初始位置$x(0)$和初始动量$p(0)$后，从时刻0出发，由上式可以算出时刻h的位置和动量，然后再依次计算t=2h, 3h, …时刻的位置和动量，最后得到粒子运动的整个轨迹。该问题的FORTRAN程序可编写为：

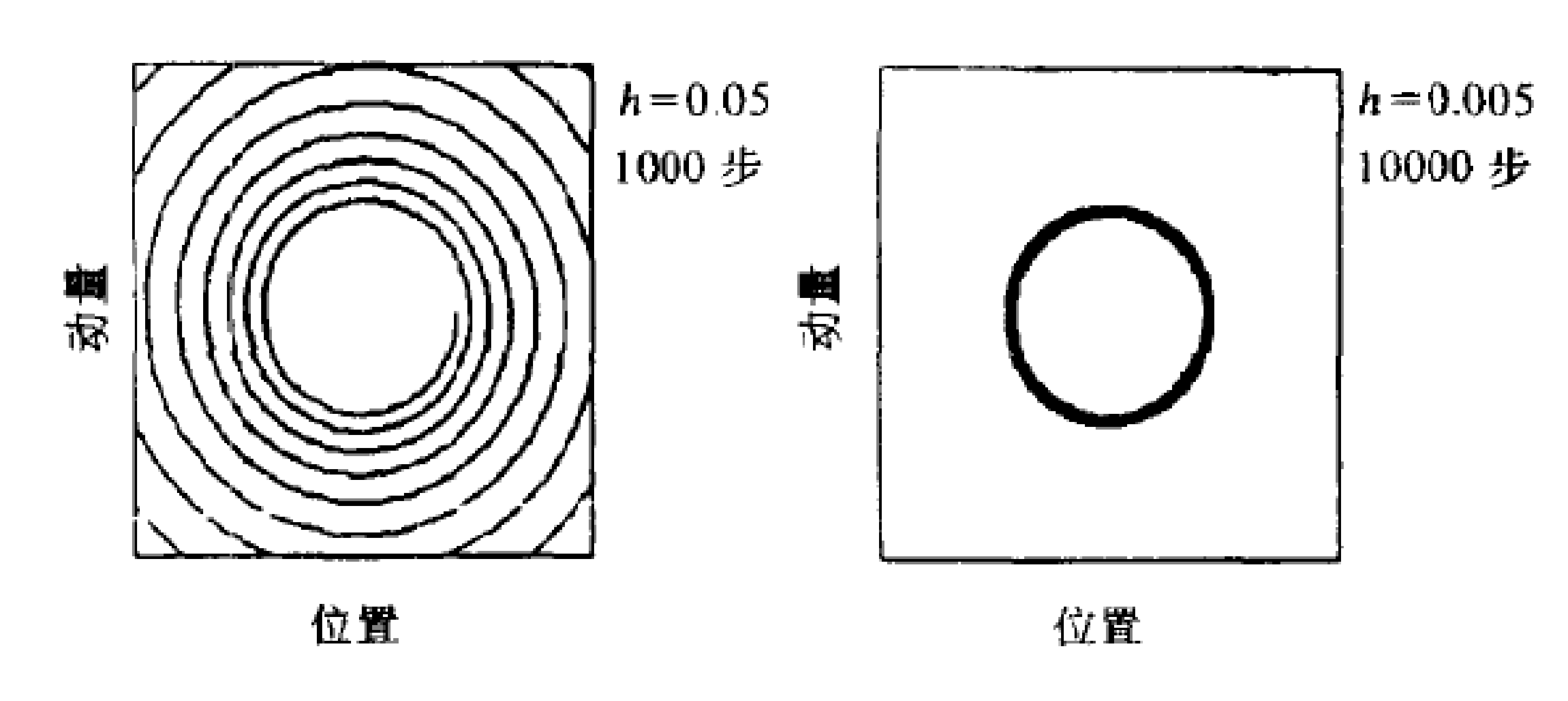
```
      subroutine particletrack(h, N, x,
p,m,k)
      integer N
      real h, m, x(0:N), p(0:N)
      do 10 I=1,N
         x(I)=x(I-1)+h*p(I-1)/m
         p(I)=p(I-1)-h*k*x(I-1)
10    continue
      return
      end
```

确定性模拟的简单示例—一维谐振子

确定性模拟的困难：由于采用差商代替导数，而时间步长不可能无限小，所以每步计算都存在误差。这些误差会随着计算步数的增多而逐渐累加。如何控制计算误差非常重要！

上例中，有限大小的时间步长，算出的轨道将偏离真正的轨道，步长愈大，偏离愈远。另一方面，很小的时间步长将导致计算量的大幅增加。即**精度和计算量是矛盾的**。

确定性模拟的简单示例——一维谐振子



步长太大，能量不守恒

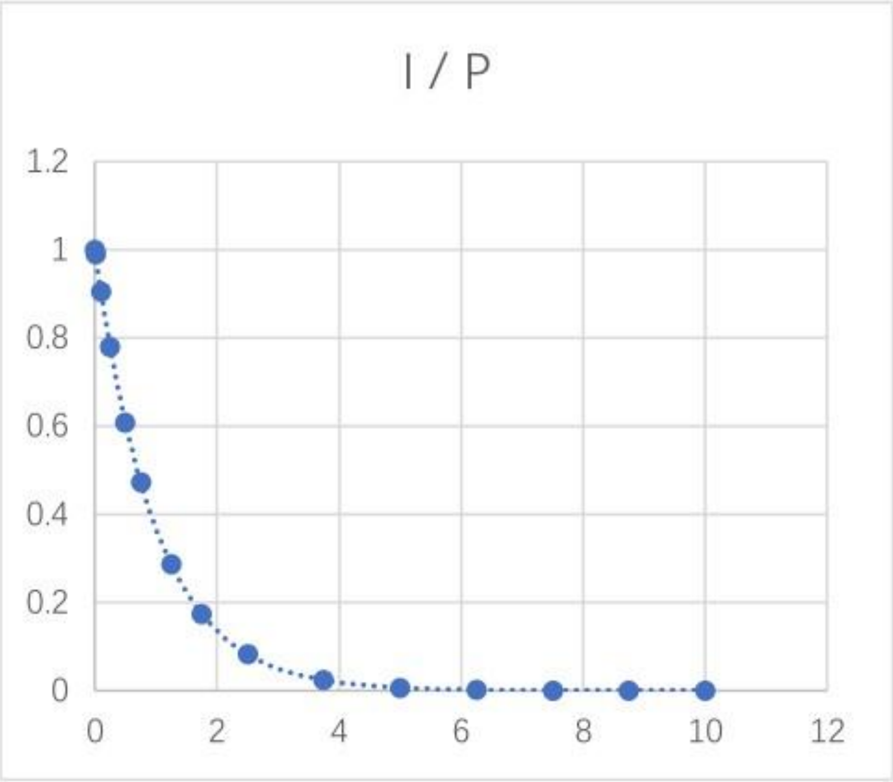
步长合适，能量近于守恒

上机习题

1、构作一个L^3的三维方格，初始时对所有格点赋值0，然后产生三个伪随机数，转换成$[1,L]$之间的整数后代表选中的格点，如果该格点为0，则赋值1。经过多次挑选和填充后，赋值为0的格点愈来愈少。当次数达到$I*L^3$次时，空格子数的比例应按指数规律$\exp(-I)$下降。选择一个随机数发生器，并检验之。

上机习题解答

结果： 对出现空格的概率取对数拟合，$y = 0.9997x + 0.0007$、$R^2 = 1$。



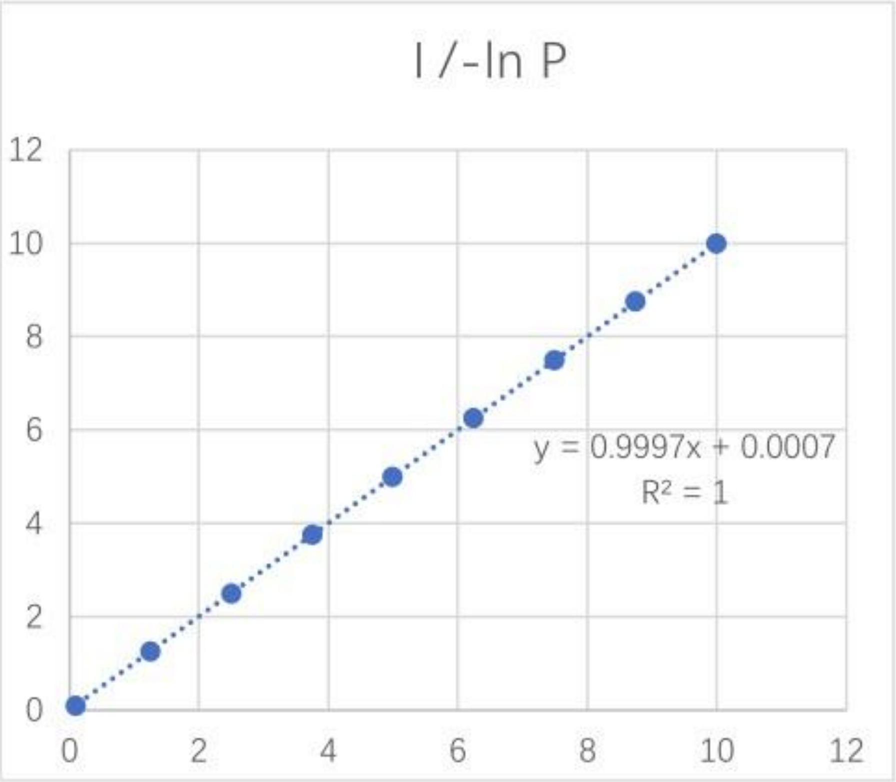


图4：利用随机数计算L^3的空间方格中的空格比例，以填入次数的指数下降。

注： 计算次数 $N = I * L^3$，横轴为次数变量 I，纵轴为空格占总空间的比例为 P。运行过程截图详见：附录-4。

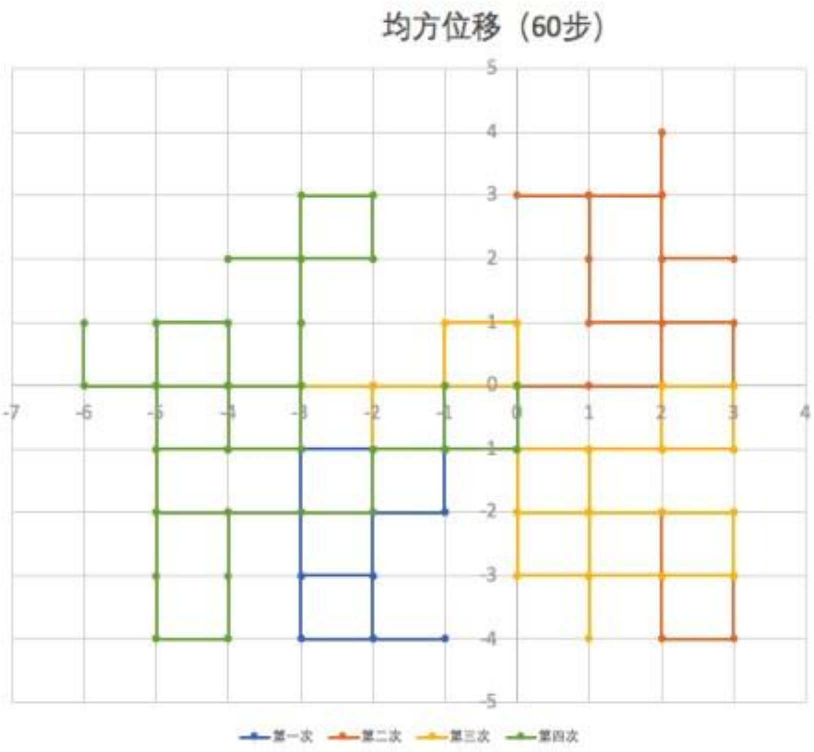
思考题

2、为二维逾渗问题写一个计算机程序，判断所产生的一个位形是否存在逾渗通道。

3、考虑一个二维方格，一个粒子可以在其上运动，只允许这个粒子从一个格点运动到最近邻的格点。从原点出发，选一个随机数，这个随机数决定粒子运动到$q(=4)$个最近邻格点中的哪一个格点，继续这样走下去，共走N步。计算每次行走的均方位移，并对结果进行平均，画出均方位移与N的函数关系图。

思考题 3 解答

结果： 可见随着模拟计算次数的增加，均方位移增大，离散程度变大。



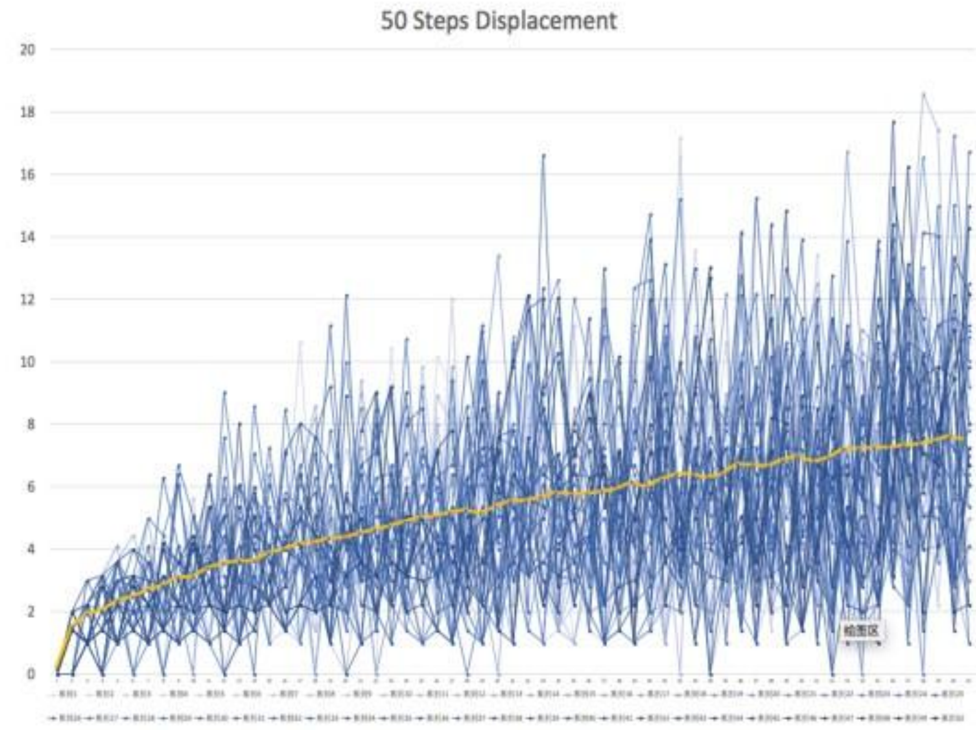


图6：多次模拟计算均方位移的模拟结果如左图，统计终点位移与模拟步数的关系如右图。