

The Monte Carlo Method in Astrophysics *(Part 2)*

Luis A. Aguilar

aguilar@astrosen.unam.mx

3-Dimensional case

The multi-dimensional case can be seen as a trivial extension of the methods we saw for the 2-dimensional case, and indeed, it is for separable functions. However, we will use the 3-dimensional case to illustrate a subtle point that we should be careful about.

We will use a spherically symmetric **Plummer model** as our example:

$$\rho_P(r) = \left(\frac{3M}{4\pi r_o^3} \right) \left(1 + \frac{r^2}{r_o^2} \right)^{-5/2}$$

This function was introduced by H. Plummer at the beginning of the XX century to describe the density profile of globular clusters.

M is the total mass and r_o is a characteristic length called the core radius. It is important in N -body simulations because it corresponds to a smoothed point mass potential.

In our case, since this is a spherically symmetric model, we can obtain the cumulative distribution function along the radial coordinate, find its inverse, and proceed as we did before.

3-Dimensional case

Let us work with a dimensionless version of this model:

$$\delta_p(\eta) = (1 + \eta^2)^{-5/2}, \quad \text{where } \eta \equiv r / r_o, \quad \text{and} \quad \delta_p(\eta) = \rho_p(\eta) / (3M / 4\pi r_o^3)$$

The cumulative distribution function can be obtained integrating $\delta_p(\eta)$ in spherical shells:

$$4\pi \int_0^\eta s^2 \delta_p(s) ds = \frac{4\pi}{3} \frac{\eta^3}{(1 + \eta^2)^{3/2}} \rightarrow \frac{4\pi}{3}, \quad \text{when } \eta \rightarrow \infty$$

So we define the normalized cumulative distribution function as:

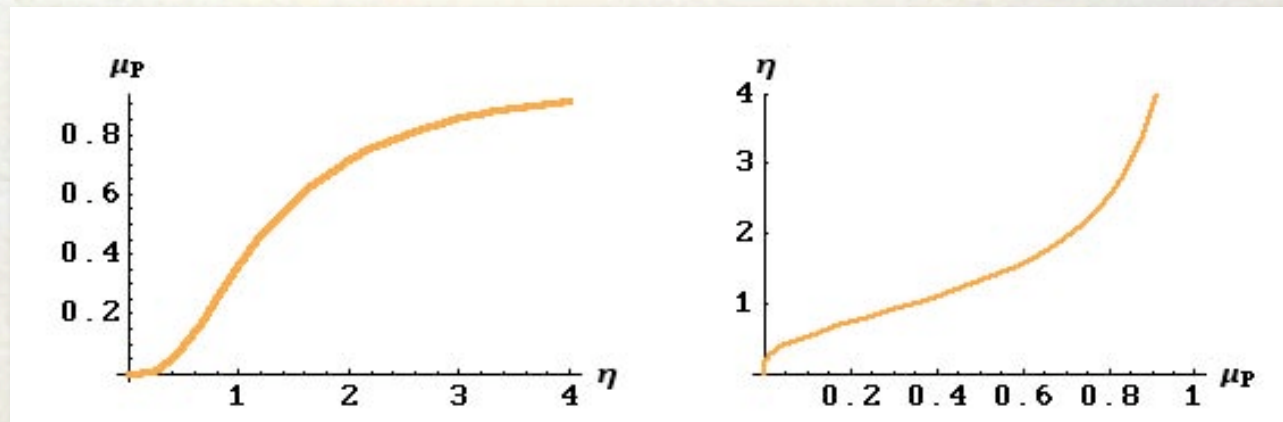
$$\mu_p(\eta) \equiv \frac{\eta^3}{(1 + \eta^2)^{3/2}}$$

The inverse function is simply,

$$\eta = \sqrt{\frac{\mu_p^{2/3}}{1 - \mu_p^{2/3}}}$$

3-Dimensional case

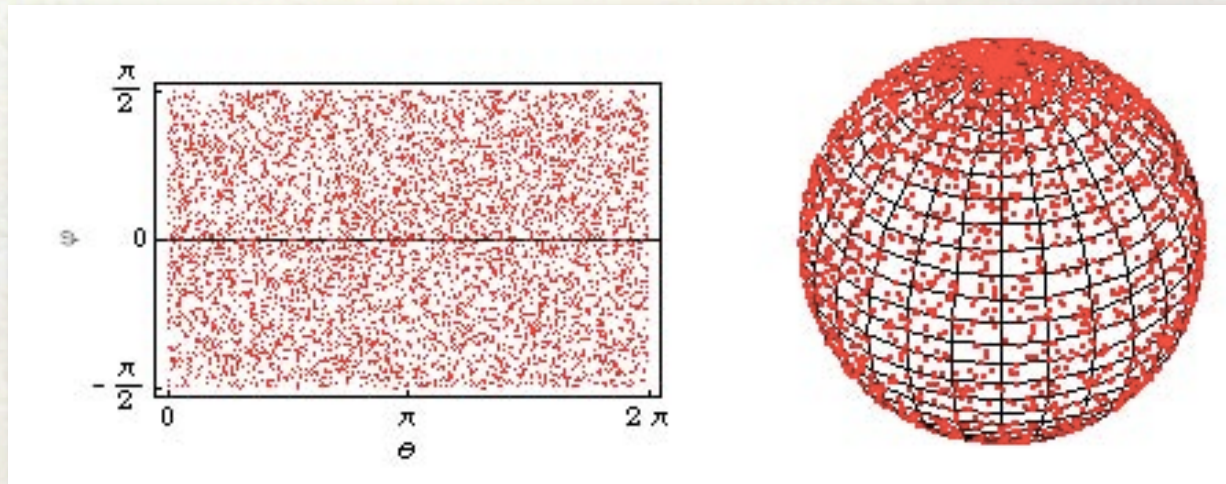
We can now draw random values for μ_p between 0 and 1 and use the inverse function to get the respective values for the normalized radial coordinate.



The next step is to draw random angles for the azimuth $0 < \theta < 2\pi$ and latitude $-\pi/2 < \phi < \pi/2$.

3-Dimensional case

It may seem that this is a straightforward draw of random values in these intervals. The figures below show what happens. The left figure is a Cartesian plot and the right figure shows the result on the unit sphere:



It is immediately apparent that something is wrong, as the density of points increases from the equator toward the poles on both hemispheres of the unit sphere.

What is going on?

3-Dimensional case

To understand this we should go back to consider a probability distribution function that depends on r only:

$$f(x, y, z) = f(\sqrt{x^2 + y^2 + z^2}) = f(r)$$

We remember that the fraction of the sample in a given region has to be the same, regardless of the coordinate system used:

$$f(x, y, z) dx dy dz = f(r) r^2 \cos(\phi) dr d\theta d\phi$$

where $0 < \theta < 2\pi$ is the azimuth, and $-\pi/2 < \phi < \pi/2$ is the latitude:

$$x = r \cos(\phi) \cos(\theta), \quad y = r \cos(\phi) \sin(\theta), \quad z = r \sin(\phi)$$

If we want to express this probability distribution function as the product of three independent 1-dimensional distribution functions, then:

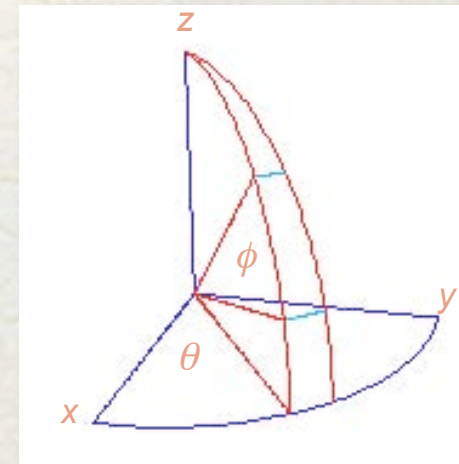
$$f(r) r^2 \cos(\phi) dr d\theta d\phi = [f_r dr] [f_\theta d\theta] [f_\phi d\phi]$$

3-Dimensional case

It is obvious that, $f_r = f(r)r^2$, $f_\theta = 1$, $f_\phi = \cos(\phi)$

The radial distribution function is the one we obtained before. The azimuthal distribution function is just the constant probability function. However, the latitudinal distribution function is not uniform! It varies with the cosine of the latitude.

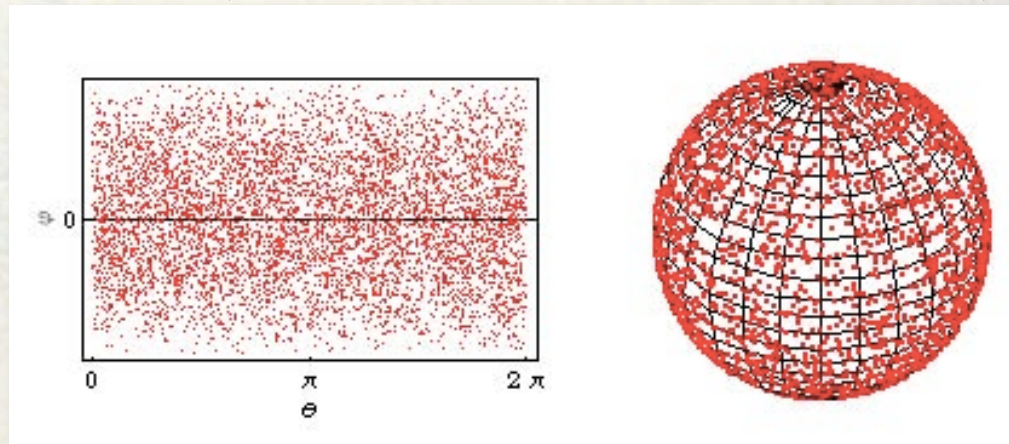
It is easy to understand why this is so, by looking at the figure: as the latitude increases from the equator to the pole, a fixed interval in azimuth corresponds to smaller arcs, and this is precisely the effect seen before. If we draw azimuth values uniformly regardless of their latitude, we will be assigning the same expected number of points to ever smaller arcs as we move toward the poles, we then generate a spurious overdensity there.



3-Dimensional case

The proper procedure is to use the cumulative function of f_ϕ to generate ϕ values. An easy way to accomplish this is to generate random uniform values for $\sin(\phi)$ in the $[-1, +1]$ interval and then just apply the arcsine function:

$$(\theta_i, \phi_i) = (2\pi \times \text{ran}(), \text{Sin}^{-1}[2 \times \text{ran}() - 1])$$



The mapping of points on the unit sphere is now uniform.

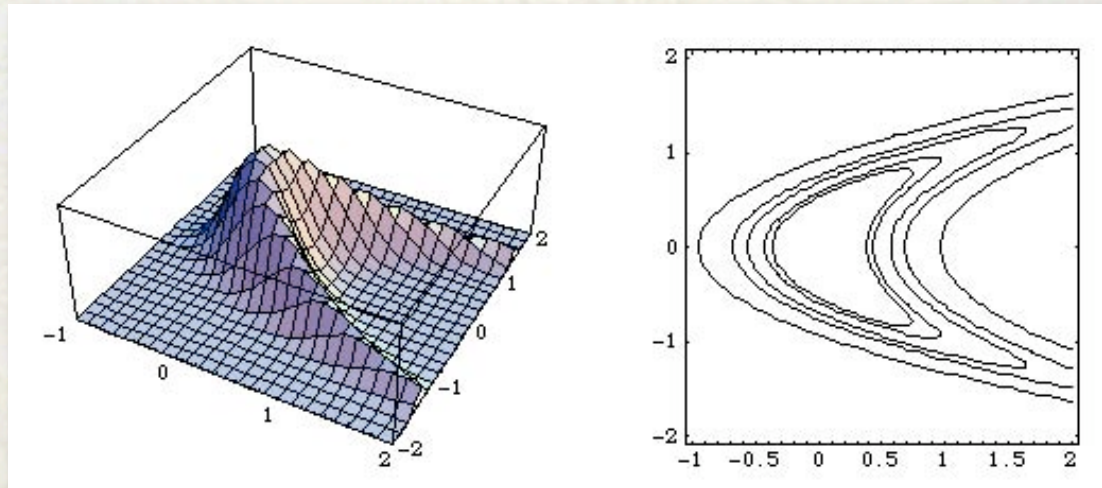
We must always remember to take into account the fact that probabilities are conserved when changing coordinate systems.

Von Neumann Rejection Technique

Up to now we have only dealt with highly symmetric, multi-dimensional distribution functions, where a simple change of variables, or change of coordinate system, allows us to express the distribution function as the product of 1-dimensional distribution functions on each variable.

We have, in fact, not really dealt with truly multidimensional distribution functions.

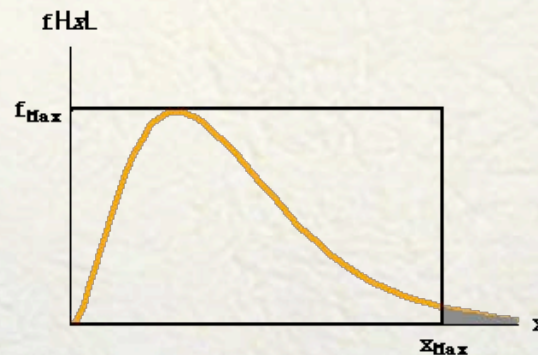
What do we do when we have a distribution function that can not be split?



John von Neumann introduced a very powerful technique that is charmingly simple: the Von Neumann Rejection Technique.

Von Neumann Rejection Technique

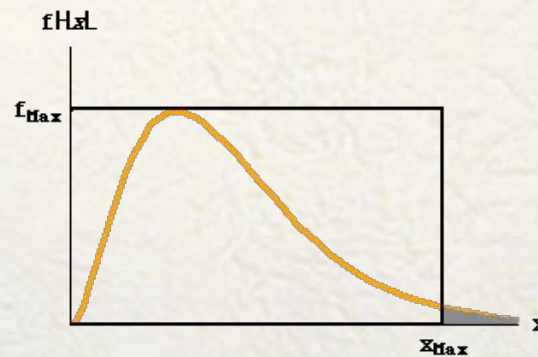
We will first illustrate its basic idea with a 1-dimensional function. Although a probability distribution function may have poles, or an infinite range in its domain, since it is normalized, we can always define a finite domain and a finite range in its value, such that a fraction, say f_0 , of the total probability, is within the bounded region.



Now, say we are building a random realization of size N . It is clear that, with such realization, a fraction smaller than $1/N$, can not be represented. If we choose a bounded region such that $f_0 = 1 - 1/N$, we won't be neglecting anything that can be represented with a sample of this size.

Von Neumann Rejection Technique

We will first illustrate its basic idea with a 1-dimensional function. Although a probability distribution function may have poles, or an infinite range in its domain, since it is normalized, we can always define a finite domain and a finite range in its value, such that a fraction, say f_o , of the total probability, is within the bounded region.

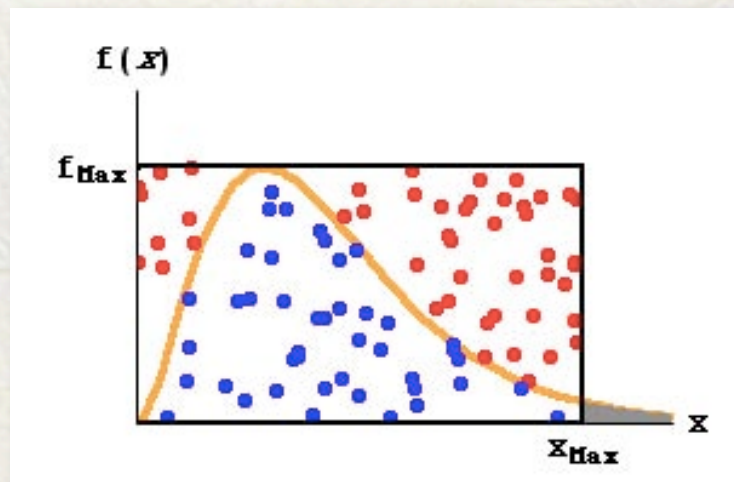


So, say we have chosen a range $0 < x < x_{max}$ and $0 < f(x) < f_{Max}$, where the upper bound in $f(x)$ is such that the distribution function is within the range for all x within the chosen interval.

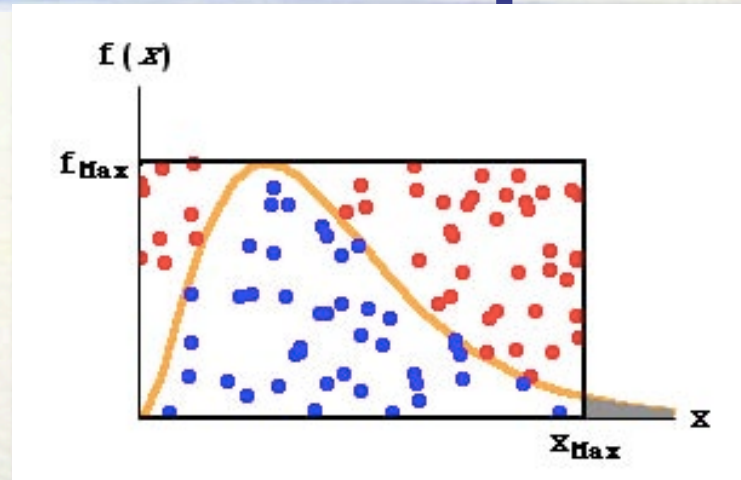
Von Neumann Rejection Technique

The procedure is now to randomly draw values in x and f within this region and check whether the points represented by these values lie below or above the distribution function. If they are below, we accept them in our sample, if they don't, we reject them. And this is it!

You can see this in the figure, where accepted points are painted blue and the rejects are painted red. The gray shaded tail represents the region we are leaving out.



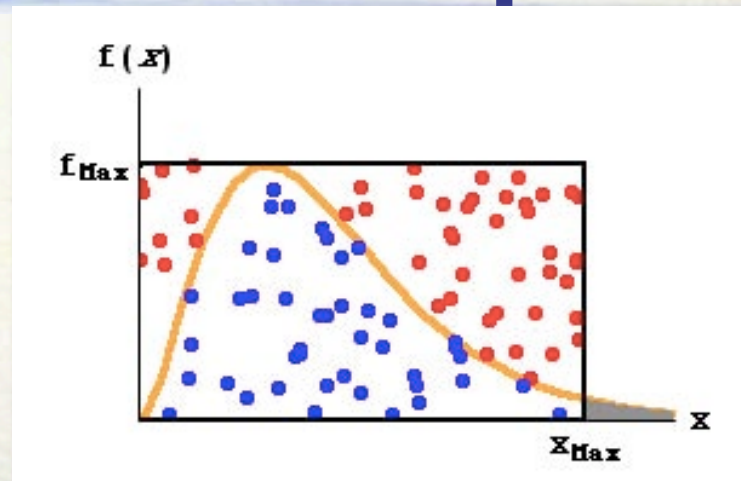
Von Neumann Rejection Technique



You may wonder how it is that this disarmingly simple procedure works. From the figure, it is clear that we are drawing points completely at random and yet, their distribution will follow the desired function.

One way of looking at it, is that once an arbitrary abscissa is chosen, the probability of this abscissa being included in the sample, goes as the ratio of the height of the function at this point, to the height of the bounded region.

Von Neumann Rejection Technique



It may seem that changing f_{Max} changes the odds of a given abscissa being accepted, and indeed, it does. However, what is changed is the absolute probability that the abscissa will be accepted within this bounded region.

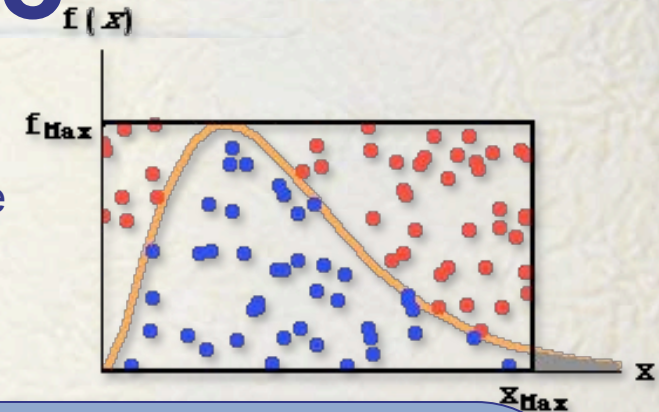
If we change the height of the box, we change this absolute probability, but not the relative acceptance probability of this particular x with respect to other values, and this is what counts:

$$\text{prob}(x_i) \propto f(x_i) / f_{Max}, \quad \text{prob}(x_j) \propto f(x_j) / f_{Max}, \quad \Rightarrow \quad \text{prob}(x_i) / \text{prob}(x_j) = f(x_i) / f(x_j)$$

Increasing the height of the box only changes the efficiency of the technique, as more points will be rejected. Obviously, the height should not be shrunk to the point where the distribution function is chopped within the considered range in x .

Von Neumann Rejection Technique

An algorithm for the Von Neumann Rejection Technique would go as follows:



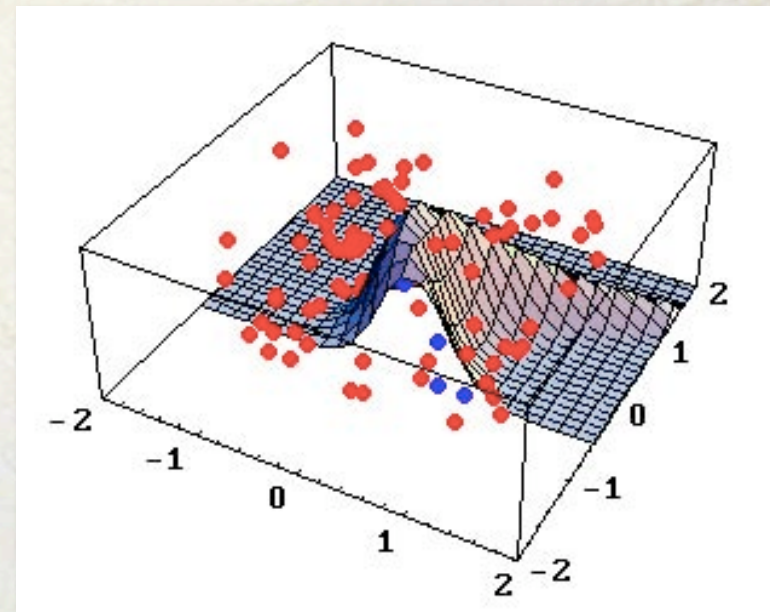
- Find cumulative distribution function: $F(x)$.
- Find value x_{Max} such that $F(x_{Max}) = 1 - (1/N)$, where N is the size of the desired random realization.
- Find f_{Max} such that $f(x) < f_{Max}$ for all $0 < x < x_{Max}$, or if it goes beyond, the neglected probability is less than $1/N$.
- Draw a random point (x_i, f_i) within $0 < x_i < x_{Max}$ and $0 < f_i < f_{Max}$.
- Evaluate distribution function at x_i : $f(x_i)$.
If $f_i < f(x_i)$, accept f_i , otherwise reject it.
- Keep iterating two previous steps until N values x_i have been found.

Von Neumann Rejection Technique

The extension of the von Neumann technique to multi-dimensional distribution functions is pretty straightforward:

- Choose a bounded region in the domain of the function and an upper bound of it within this domain.
- Randomly draw points within the restricted domain and values of the function within its bound range.
- If the randomly drawn value of the function is smaller than the value of the function at this point in the domain, we accept the point we have drawn, otherwise we reject it.

Notice that the shape of the chosen region in the domain of the function is completely arbitrary.



Von Neumann Rejection Technique

The beauty of this technique is its simplicity and its applicability. Its drawback is its inefficiency, as many points that will be rejected are generated.

If the distribution function can be split in 1-dimensional functions, we will be better off in general using the methods presented before.

However, if the function can not be split, the von Neumann technique will fit the bill.

The inefficiency of this technique quickly increases with the dimensionality of the distribution function, but when the other techniques can not be applied, the von Neumann technique will always work, if we are patient enough (and have a computer powerful enough).

Homework

2nd Assignment - I

- Write a program that generates random realization of a Plummer sphere (positions only).
- Write separately a routine that generates radial coordinates, and another that generates random points on the unit sphere. Combine the results to generate cartesian coordinates for each point.
- Make a plot of the theoretical radial density profile vs. a histogram of the radial distribution of particles, normalized to the volume of the spherical shell used for each radial bin.
- Determine the global *rms* deviations between both curves as a function of the size of the random realization.

Homework

2nd Assignment - II

- Write a program that approximates the value of the integral:

$$A = \frac{1}{\sqrt{2\pi}} \iint_{x^2 + y^2 \leq 1} \exp[-(x^2 + y^2) / 2] dx dy$$

using the von Neumann Rejection Technique.

- List the approximated values A_n obtained using n points, as well as the fractional errors: $(A_n - A_o) / A_o$; where A_o is the exact value.
- Make a plot of the fractional error as a function of n .

End of second lecture