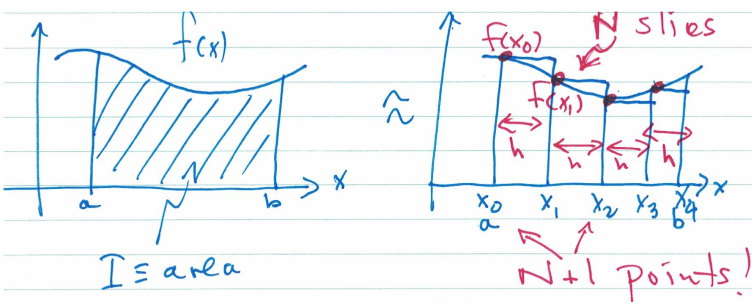
Physics 395: Guide 2 Notes

Numerical Integration

Many problems in physics require us to calculate definite integrals,

$$I = \int_a^b f(x)\mathrm{d}x \ . \tag{1}$$

Doing an integral amounts to calculating an area under a curve.



The simple Riemann sum is just to add up the rectangles:

$$\int_a^b f(x)\mathrm{d}x \approx \sum_{i=0}^{N-1} f(x_i)h \ . \tag{2}$$

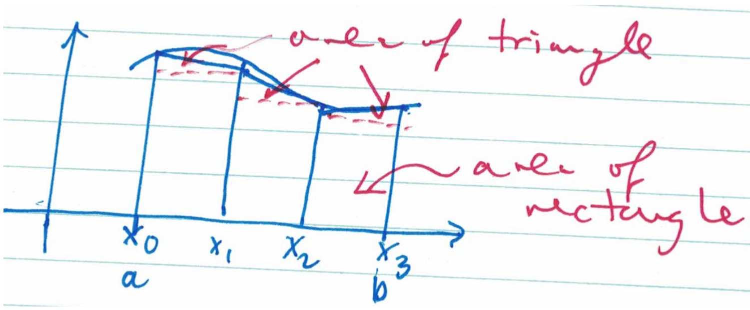
This method often needs large N to get low error.

Q: How can we do better? Let's discuss better ways to approximate the area.

Trapezoid Rule

This uses a linear interpolation to better represent the area.

So we approximate the integral by the areas of the trapezoids, i.e., the sum of the areas of the rectangles plus the areas of the triangles (see the equation in the Activity Guide).

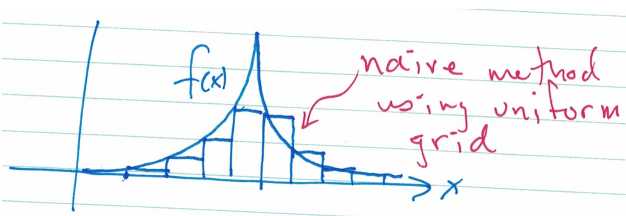


Simpson's Rule

Instead of using a line, this method improves the area estimates by using a quadratic approximation. The derivation is a bit involved.

Pathological functions

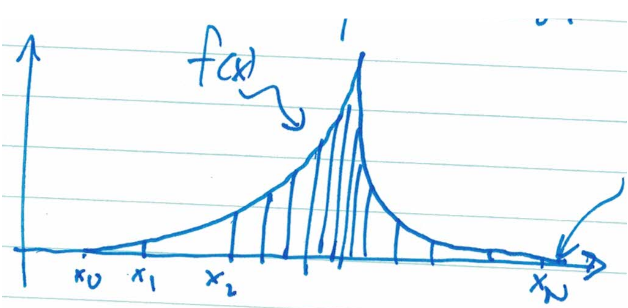
Try to integrate:



If we were to use a uniform grid, this would give a horrible estimate. We would need to use large N to get a low error.

Quadrature

This method uses a non-uniformly spaced set of points to estimate the integral.



The x_i are determined using a polynomial representation of the function $f(x)$.

Advantage: can use much smaller N compared to a grid-based approach.

Disadvantage: For this method you need a closed-form expression of the function you are trying to integrate (i.e., this will not work for integrating data; instead use trapezoid or Simpson's method).

Monte Carlo methods

In statistical physics, we are often faced with high-dimensional integrals,

$$I = \int \cdots \int f(x_1, ..., x_{n+m}) \mathrm{d}x_1 \cdots \mathrm{d}x_n \ . \tag{3}$$

Using a grid-based approach would be impractical as the number of points blows up with increasing dimension.

Solution: sample the function using a set of randomly chosen points. The average of a function is related to the integral by

$$\int_a^b f(x) \mathrm{d}x = (b-a)\langle f \rangle \ . \tag{4}$$

And you estimate the average by

$$\langle f \rangle \approx \frac{1}{N} \sum_{i=1}^{N} f(x_i) \ . \tag{5}$$

So pick a random set of $\{x_i\}$ and compute the average of $\{f(x_i)\}$.

Advantage: can get good estimates of high-dimensional integrals.

Disadvantage: can waste time sampling from unimportant regions.