Minimal Energy of 100 Point Charges on the Unit Sphere

In Problem 04-003, Folkmar Bornemann, Dirk Laurie, Stan Wagon, and Jörg Waldvogel proposed a set of five numerical problems taken from their book *The SIAM 100-Digit Challenge: A Study in High-Accuracy Numerical Computing*. These problems included the following one, contributed by Lloyd N. Trefethen of Oxford University. (See problem 4, p. 282.)

(a) If N point charges are distributed on the unit sphere, the potential energy is

$$E = \sum_{j=1}^{N-1} \sum_{k=j+1}^{N} |x_j - x_k|^{-1},$$

where $|x_j - x_k|$ is the Euclidean distance between x_j and x_k. Let E_N denote the minimal value of E over all possible configurations of N charges. What is E_{100}?

Solution by Wouter Tierens[1] (Engineering Student, University of Ghent, Belgium).

1 History

Before 1909, atoms were thought to consist of a diffuse positive charge surrounded by electrons (the "plum pudding" model). In order to predict the properties of the elements on the periodic table with this model, the distribution of electrons on a sphere had to be studied. This problem was named after Thomson, the originator of the plum pudding model and the first to study it.

In 1909, Rutherford showed the existence of nuclei with his famous gold foil experiment. His results falsified the plum pudding model, yet the problem turned out to be important for many other fields, from biology to telecommunications.

[1]Email: Wouter.Tierens@gmail.com

2 Theoretical Considerations

2.1 An upper bound

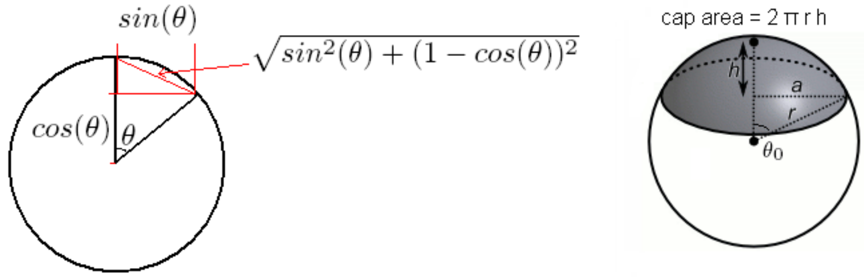

Figure 1: Left: Integrating over the unit sphere. Right : a spherical cap.

For large N, a continuous approximation can be used:

(1) $$E \approx \frac{1}{2} 4\,\pi \int_0^{2\pi} \int_0^{\pi} \frac{\rho^2 \sin(\theta)}{\sqrt{\sin^2(\theta) + (1 - \cos(\theta))^2}} d\theta d\phi = \frac{N^2}{2},$$

with $\rho = \frac{N}{4\pi}$.

Replacing a point charge by a charge density in some finite area increases the potential energy, because this area also has some potential energy due to itself. Each point becomes distributed over an area of $\frac{4\pi}{N}$. Approximating this area as a spherical cap (see Figure 1),

(2) $$\frac{4\pi}{N} = 2\pi h \implies h = \frac{2}{N},$$

(3) $$a = \sqrt{1 - (1 - h)^2},$$

(4) $$\theta_0 = \arctan\left(\frac{2\sqrt{N - 1}}{N - 2}\right).$$

The excess energy is

(5) $$E_{\text{excess}} = \frac{1}{2} \int_{cap} \int_{cap} \frac{\rho^2}{d(\vec{p}, \vec{p'})} dS\, dS' \approx \frac{1}{2} \frac{4\pi}{N} \int_0^{2\pi} \int_0^{\theta_0} \frac{\rho^2 \sin(\theta)}{\sqrt{\sin^2(\theta) + (1 - \cos(\theta))^2}} d\theta\, d\phi.$$

Here the second integral over the spherical cap was approximated as the area of the cap times the excess energy of a point at the center of the cap.

The energy becomes

(6) $$E \approx \frac{N^2}{2} - N E_{\text{excess}} = \frac{N^2}{2} \left(1 - \sin\left(\frac{1}{2} \arctan\left(\frac{2\sqrt{N - 1}}{N - 2}\right)\right)\right).$$

For the reasons mentioned above, $\frac{N^2}{2}$ is an upper bound for E_N. Formula (6) will be a reasonably good approximation.

2.2 A lower bound

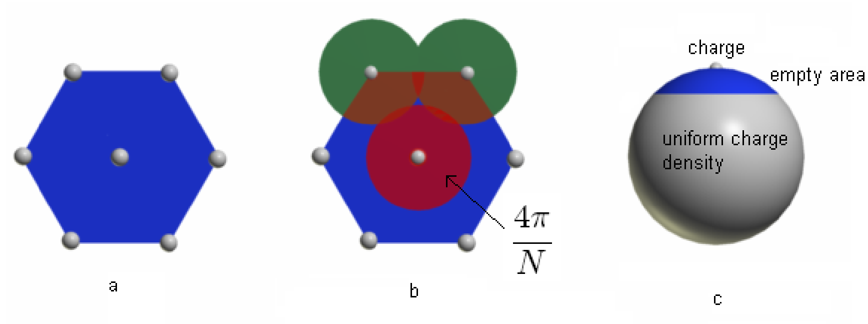

Figure 2: Hexagonal pattern, lower bound model.

Due to the repulsive nature of the interactions, it is reasonable to expect empty areas of at least $\frac{4\pi}{N}$ (actually larger) around every charge. These empty areas are actually more-or-less regular polygons (usually hexagons for large N). They can be approximated as being spherical caps.

Assuming that the charges are arranged in a hexagonal pattern, the actual empty area is given by $\frac{4\pi}{N} + 6(\frac{4\pi}{3N}) = \frac{4\pi}{N/3}$. This follows from Figure 2 (a and b) : the circles are the $\frac{4\pi}{N}$-areas, and the red parts of these circles are the parts of the $\frac{4\pi}{N}$-areas that contribute to the empty area. In other words, formulae (2) through (4) can be used for the complete empty area by simply substituting $N \to \frac{N}{3}$.

Now a lower bound for E_N can be found using the following model:

- There are N charges, all of which are equivalent. Thus $E_N = E_c N$, with E_c the energy of one charge.
- Each charge is surrounded by an empty spherical cap with area $\frac{4\pi}{N/3} = \frac{12\pi}{N}$.

In order to calculate E_c, an additional approximation is necessary: to find the effect of the $N-1$ charges on one charge, the $N-1$ charges are replaced by a uniform charge density that covers the whole unit sphere, except for the empty area around the charge under consideration (see Figure 2 c). This charge density is $\rho = (N-1)/(4\pi - \frac{12\pi}{N})$.

We can now calculate E_N:

$$E_N \approx \frac{N\rho}{2} \int_0^{2\pi} \int_{\theta_0}^{\pi} \frac{\sin(\theta)}{\sqrt{\sin^2(\theta) + (1-\cos(\theta))^2}}\, d\theta\, d\phi, \tag{7}$$

$$\theta_0 = \arctan\left(\frac{2\sqrt{(N/3)-1}}{(N/3)-2}\right). \tag{8}$$

The result is (for $N > 6$)

$$E_{N,\,\text{lower}} = \frac{N^2(N-1)}{2(N-3)} \left(1 - \sin\left(\frac{1}{2} \arctan\left(\frac{2\sqrt{3N-9}}{N-6}\right)\right)\right). \tag{9}$$

For large N, this formula has to reduce to $\frac{N^2}{2}$, and indeed it does.

Formula (9) is an underestimation because the charge density on the remaining part of the unit sphere is not really uniform : the nearest neighbor charges are just outside the empty area, causing a higher than average local charge density near the empty area.

3 Computing E_{100}

3.1 Description of the method

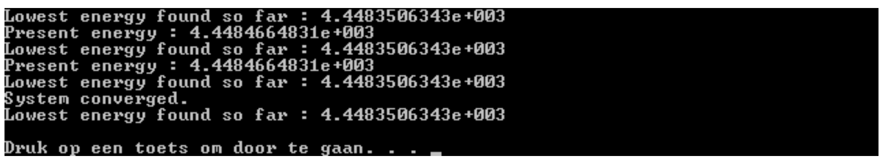

Figure 3: The program for $N = 100$.

Most approaches to finding minima of functions are more-or-less sophisticated variations on “walk downhill.” In this particular problem, “downhill” is equivalent to moving each point charge in the direction of the Coulomb force.

Given the physical nature of this problem, I opted for a physical approach. The charges are given a unit mass and a velocity. The velocity is constantly updated using Newton’s laws. In order to ensure convergence, a friction force $\vec{F}_{friction} = -\gamma\vec{v}$ is used. The system is considered to be converged when the kinetic is smaller than 10^{-8} and the changes in the potential energy are smaller than 10^{-7}, hopefully ensuring 10-digit accuracy.

In order to avoid getting stuck in a local minimum, the system is “annealed” once it has converged: the charges are given new velocities selected from a Maxwell–Boltzmann distribution. This process is repeated a few times in order to search for nearby, more optimal solutions.

3.2 Results

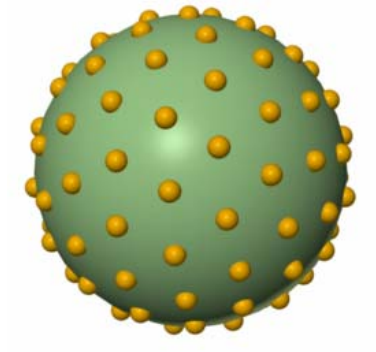

Figure 4: The minimum-energy configuration for 100 point charges on the unit sphere. Notice the hexagonal pattern, with a few pentagons.

The result for E_{100} is

$$E_{100} = 4448.350634. \tag{10}$$

3.3 Remarks

The method used here is by no means the fastest or most efficient. There are plenty of reasonably good heuristics to find a more-or-less uniform distribution of points on a sphere, and there are lots of quick algorithms to find a minimum without simulating the physics of the system. The advantage of this method is that it performs a thorough search for better minima, hopefully finding the global minimum.

4 Theory vs. Computation

The obtained result is smaller than the upper bound $\frac{N^2}{2} = 5000$ and greater than the lower bound $E_{100,lower} = 4219.211186$.

As expected, formula (6) gives a reasonable approximation for E : $E_{100,approximated} = 4500$.

Recall that, in section 2.2, the empty area was expected to be $\frac{12\pi}{N}$. This was derived under the assumption that the configuration consists of a nearly hexagonal pattern. To check the accuracy of this prediction, the average empty area is calculated for $N = 100$. In order to do this, an extremely elegant result from spherical trigonometry is used:

$$A = \alpha + \beta + \gamma - \pi. \tag{11}$$

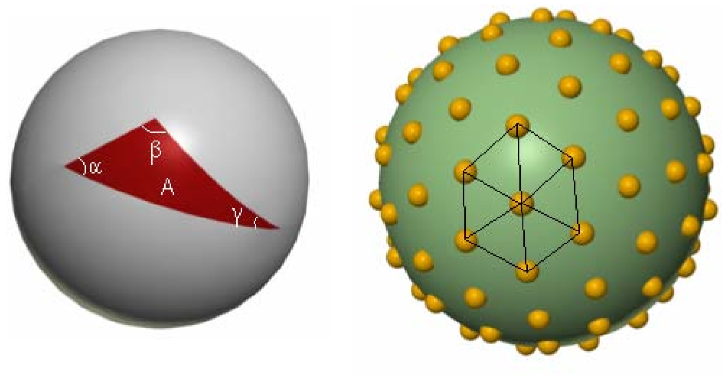

Figure 5: Left: a spherical triangle. Right: the empty area as a sum of areas of spherical triangles.

Here A is the surface of a spherical triangle, and α, β, and γ are its angles (in radians). The average empty area turns out to be 0.378166807. The predicted value was $\frac{12\pi}{100} = 0.376991116$, which is close to the computed value.

In Figure 6, E_N is shown for values of N between 10 and 100, together with the upper bound, the lower bound, and the approximation (formula (6)). The results are in line with the expectations.

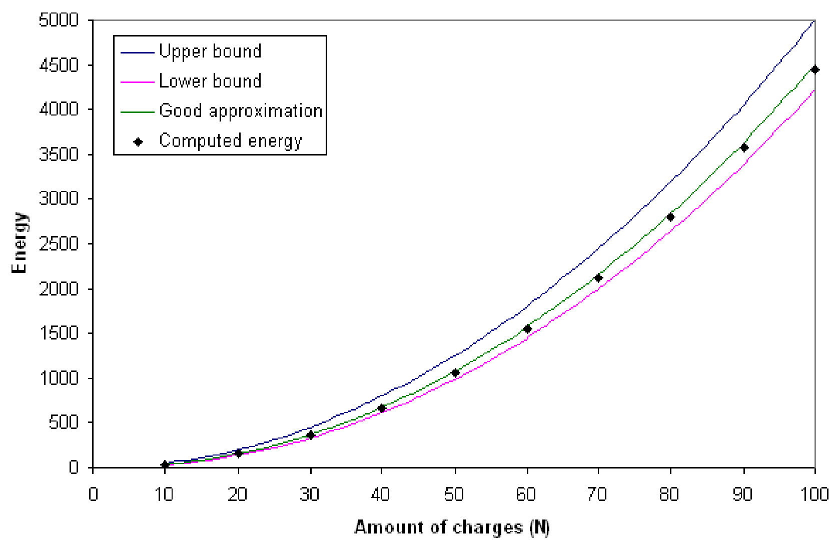

Figure 6: Results for $N = 10, 20, \ldots, 100$.