Sommersemester 2017, Struktur und Simulation

# Übung: Calculating gradients and forces

aka "vectors, gradients and forces till you bleed"
**Assignment due date: 25.4.2017**

In this exercise, we will go through all the steps of deriving a expression for a force from a potential energy function. This will involve taking some derivatives. If it's been a long time since you've calculated a derivative this might look scary, but with a little practice it shouldn't be too difficult. We will first repeat some things already mentioned in the lectures and try to clear up any misunderstandings. Then we will go through some gradient calculations step by step. Finally, there is a homework assignment where you will calculate some gradients yourself.

# Notation

A vector $\vec{r}$

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

has Euclidean length $r$

$$r = \sqrt{x^2 + y^2 + z^2}$$

Mathematicians call the length a "norm" (there are other norms apart from the Euclidean norm, but that's not important here).

The unit vector $\hat{r}$ (it has a length of 1) in direction $\vec{r}$ is given by

$$\hat{r} = \frac{\vec{r}}{r}$$

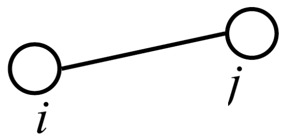
We will use the convention $\vec{r_{ij}} = \vec{r_i} - \vec{r_j}$.

# Gradient of the length of a vector

Many potential energy functions are functions of distances, which means we will need the gradient of the length of a vector later on. The gradient ($\nabla$ or $\frac{\partial}{\partial \vec{r}}$) of the Euclidean length turns out to have the simple form:

$$\nabla r = \frac{\partial}{\partial \vec{r}} r = \frac{\vec{r}}{r} = \hat{r}$$

# Calculating forces: harmonic bonds




Many potential energy functions are functions of distances. A simple example is the harmonic potential for bonds. Given a harmonic bond between particles $i$ and $j$, the potential energy is

$$U(r_{ij}) = \frac{k}{2}\left(r_{ij} - r_0\right)^2$$

where $k$ and $r_0$ are force constants.

As mentioned above $\vec{r_{ij}} = \vec{r_i} - \vec{r_j}$.

If we want to work out the gradient $\nabla_{\vec{r_i}} U(r_{ij})$, we have to use the chain rule:

$$\nabla_{\vec{r_i}} U(r_{ij}) = \left(\frac{d}{dr_{ij}} U(r_{ij})\right)\left(\nabla_{\vec{r_i}}\, r_{ij}\right) = k(r_{ij} - r_0)\frac{\vec{r_{ij}}}{r_{ij}}$$

The force $\vec{F_i}$ acting on particle $i$ is

$$\vec{F_i} = -\nabla_{\vec{r_i}} U(r_{ij}) = -k(r_{ij} - r_0)\frac{\vec{r_{ij}}}{r_{ij}}$$

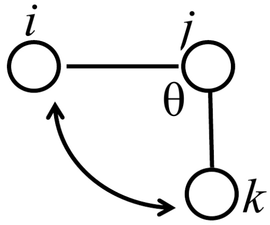
For particle $j$, we have

$$\nabla_{\vec{r_j}}\, r_{ij} = -\frac{\vec{r_{ij}}}{r_{ij}}$$

and therefore the force $\vec{F_j}$ acting on particle $j$ is

$$\vec{F_j} = -\nabla_{\vec{r_j}} U(r_{ij}) = k(r_{ij} - r_0)\frac{\vec{r_{ij}}}{r_{ij}} = -\vec{F_i}$$

# Calculating forces: angle potentials



A more complicated energy function is the potential energy for bond angles

$$U(\vec{r_i}, \vec{r_j}, \vec{r_k}) = \frac{k}{2} \left(\cos\theta_{ijk} - \cos\theta_0\right)^2$$

The cosine of the angle $\theta_{ijk}$ is

$$\cos\theta_{ijk} = \frac{\vec{r_{ij}} \cdot \vec{r_{kj}}}{r_{ij}\, r_{kj}} = \hat{r_{ij}} \cdot \hat{r_{kj}}$$

Here we have used the dot product $\vec{a} \cdot \vec{b}$ between vectors

$$\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_M \end{pmatrix}$$

$$\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix}$$

which is given by

$$\vec{a} \cdot \vec{b} = \sum_{k=1}^{M} a_k b_k$$

The dot product is also called the scalar or inner product.

To calculate the gradient of the bond angle potential, we again use the chain rule:

$$\nabla_{\vec{r_i}} U = \left( \frac{d}{d\cos\theta_{ijk}} \frac{k}{2} \left(\cos\theta_{ijk} - \cos\theta_0\right)^2 \right) \nabla_{\vec{r_i}} \cos\theta_{ijk} = k \left(\cos\theta_{ijk} - \cos\theta_0\right) \nabla_{\vec{r_i}} \cos\theta_{ijk}$$

So we are left with computing $\nabla_{\vec{r_i}} \cos\theta_{ijk}$. Using the definition for $\cos\theta_{ijk}$ from above and the quotient rule, we get

$$\nabla_{\vec{r_i}} \cos\theta_{ijk} = \nabla_{\vec{r_i}} \frac{\vec{r_{ij}} \cdot \vec{r_{kj}}}{r_{ij}\, r_{kj}} = \frac{\vec{r_{kj}}\, r_{ij}\, r_{kj} - \vec{r_{ij}} \cdot \vec{r_{kj}} r_{kj} \hat{r_{ij}}}{r_{ij}^2\, r_{kj}^2} = \frac{1}{r_{ij}} \left(\hat{r_{kj}} - \left(\hat{r_{ij}} \cdot \hat{r_{kj}}\right) \hat{r_{ij}}\right)$$

Here we have used the fact that

$$\nabla_{\vec{a}}\, \vec{a} \cdot \vec{b} = \vec{b}$$

Substituting $\cos\theta_{ijk} = \hat{r_{ij}} \cdot \hat{r_{kj}}$, we get

$$\nabla_{\vec{r_i}} \cos\theta_{ijk} = \frac{1}{r_{ij}} \left(\hat{r_{kj}} - \cos\theta_{ijk} \hat{r_{ij}}\right)$$

We can do a similar calculation to get $\vec{F_k} = -\nabla_{\vec{r_k}} U$, and then calculate $\vec{F_j}$ with the help of

$$\vec{F_i} + \vec{F_j} + \vec{F_k} = 0$$

## Checking gradients with the gradient theorem

As you have seen, there are many possibilities for subtle errors when calculating gradients. The gradient theorem can help us check a gradient numerically.

Given a function $U$ and its gradient $\nabla U$, the line integral between any two points $\vec{a}$ and $\vec{b}$ is equal to the difference of the function $U$ evaluated at the endpoints:

$$\int_{\vec{a}}^{\vec{b}} (\nabla U(\vec{r})) \cdot \vec{dr} = U(\vec{b}) - U(\vec{a})$$

As before, the $\cdot$ is a dot product.

Discretising the line integral, we get:

$$\sum_{k=1}^{N} (\nabla U(\vec{r_k})) \cdot \vec{\Delta r} \approx U(\vec{b}) - U(\vec{a})$$

with

$$\Delta r = \frac{\vec{b} - \vec{a}}{N}$$

and

$$\vec{r_k} = \vec{a} + k\,\vec{\Delta r}$$

For large $N$, the error should become very small if we implemented the gradient $\nabla U$ correctly.

# Assignment

1. For the harmonic bond potential, the force $\vec{F_i}$ was

$$\vec{F_i} = -k(r_{ij} - r_0)\frac{\vec{r_{ij}}}{r_{ij}}$$

   What is the length of $\vec{F_i}$ ?

   Hint: what is the length of $\frac{\vec{r_{ij}}}{r_{ij}}$ ?

2. Calculate the gradient $\nabla_{\vec{r_i}} U(r_{ij})$ of the Lennard-Jones potential

$$U(r_{ij}) = 4\epsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right)$$

   Hint: use the chain rule as we did for the harmonic bond potential. This means you only have to calculate $\frac{d}{dr_{ij}}U(r_{ij})$ and then combine it with the gradient $\nabla_{\vec{r_i}}\vec{r_{ij}}$ the get the final answer.

3. Show that

$$\nabla r = \frac{\vec{r}}{r}$$

where the vector $\vec{r}$ has components

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

and length

$$r = \sqrt{x^2 + y^2 + z^2}$$

# Bonus assignments (optional)

1. Implement the harmonic potential (or another one if you want) and its gradient. Show that your implementation is correct with the help of the gradient theorem, for example for a system of three particles and two bonds (a bond between particles 1 and 2 and another between particles 2 and 3). If you feel even more adventurous, choose another potential such as the angle potential or the Lennard-Jones potential.

2. Show that

$$\nabla_{\vec{a}}\, \vec{a} \cdot \vec{b} = \vec{b}$$

We used this formula when we calculated the gradient for the bond angle potential.

# Appendix A: Derivatives

We will write $f'(x)$ for $\frac{d}{dx} f(x)$ here.

The differential operator $\frac{d}{dx}$ is a linear operator. An operator in Mathematics is a function that takes a function as input and returns a function. Being a linear operator means:

$$\begin{aligned} (f(x) + g(x))' &= f'(x) + g'(x) \\ (af(x))' &= af'(x) \end{aligned}$$

With just a few additional rules, we can differentiate any function:

### Product rule

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$$

**Quotient rule**

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$$

Note: $g^2(x) = g(x)g(x)$.

**Chain rule**

$$(f(g(x)))' = f'(g(x))g'(x)$$

Some important derivatives:

**Constant functions**

$$(1)' = 0$$

**Derivatives of powers**

$$\begin{aligned}
(x^n)' &= nx^{n-1} \\
\text{Examples :} & \\
\left(x^1\right)' &= (x)' = 1 \\
\left(x^2\right)' &= 2x \\
\left(x^3\right)' &= 3x^2 \\
\left(\frac{1}{x^n}\right)' &= \left(x^{-n}\right)' = -nx^{-n-1} \\
\left(\sqrt{x}\right)' &= = \left(x^{\frac{1}{2}}\right)' = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}} \\
\left(\sqrt[3]{x}\right)' &= \left(x^{\frac{1}{3}}\right)' = \frac{1}{3}x^{-\frac{2}{3}} = \frac{1}{3\sqrt[3]{x^2}} \\
\text{etc.} &
\end{aligned}$$

**Exponential function**

$$\begin{aligned}
(e^x)' &= e^x \\
(a^x)' &= \left(\left(e^{\log a}\right)^x\right)' = \left(e^{x\log a}\right)' = (\log a)e^{x\log a} \\
(2^x)' &= \left(\left(e^{\log 2}\right)^x\right)' = \left(e^{x\log 2}\right)' = (\log 2)e^{x\log 2} \\
\text{etc.} &
\end{aligned}$$

**Trigonometric functions**

$$(\sin(x))' = \cos(x)$$
$$(\cos(x))' = -\sin(x)$$

Quick self-check:

$$\begin{aligned}
\left(4x^3 - 7x^2 + 3x - 3\right)' &= ? \\
\left(\frac{1}{x}\right)' &= ? \\
\left(\frac{4}{x^3}\right)' &= ? \\
\left(e^{3x^2-5x+3}\right)' &= ? \\
\left(\cos(4x^2+3)\right)' &= ? \\
\left(\frac{5x^2+3}{\sin(3x^2)}\right)' &= ?
\end{aligned}$$

# Appendix B: Partial derivatives

Given a function of two variables $f(x, y)$, we can take the partial derivative with respect to $x$, treating $y$ as a constant. We write the partial derivative with respect to $x$ as $\frac{\partial}{\partial x}$. Vice versa, we can take the partial derivative $\frac{\partial}{\partial y}$ and treat $y$ as the variable and $x$ as a constant.

$\frac{\partial}{\partial x}$ tells us how much $f$ changes if we change $x$, keeping all other variables constant.
Examples:

$$\begin{aligned}
\frac{\partial}{\partial x}\left(x^2 + 3xy + 3y^3\right) &= 2x + 3y \\
\frac{\partial}{\partial y}\left(x^2 + 3xy + 3y^3\right) &= 3x + 9y^2
\end{aligned}$$

All the rules of normal derivatives (linearity, product/quotient/chain rule) apply to partial derivatives as well.

# Appendix C: Gradients

A gradient is just a vector consisting of partial derivatives. Here's a small example:

Given a vector $\vec{r}$

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix}$$

and a function $f$ acting on that vector that returns a number (scalar)

$$f(\vec{r}) = f(x, y) = x^2 + 3xy + 3y^3$$

the gradient ($\nabla$ or $\frac{\partial}{\partial \vec{r}}$) is:

$$\frac{\partial}{\partial \vec{r}} f(\vec{r}) = \nabla f(\vec{r}) = \begin{pmatrix} \frac{\partial}{\partial x} f \\ \frac{\partial}{\partial y} f \end{pmatrix} = \begin{pmatrix} 2x + 3y \\ 3x + 9y^2 \end{pmatrix}$$

So when we evaluate the gradient at a point

$$\vec{r} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

the gradient is the vector

$$\begin{pmatrix} 14 \\ 48 \end{pmatrix}$$

The gradient represents (locally!) the direction in which the function $f$ increases the most, and the negative of the gradient (the opposite direction) is the direction in which it decreases the most.

Note: the $\nabla$ symbol is called “nabla”. When talking about vector calculus, you pronounce $\nabla f$ as “grad f” or “del f”.

Because they are vectors of partial derivatives, all the rules of normal derivatives (linearity, product/quotient/chain rule) apply to gradients as well.

Sometimes, we have a large vector and a function acting on it, and we only want the gradient with respect to some of the variables.

For example, let $\vec{r}$ be the vector containing all $x, y, z$ coordinates of all particles

$$\vec{r} = \begin{pmatrix} r_{1,x} \\ r_{1,y} \\ r_{1,z} \\ \vdots \\ r_{n,x} \\ r_{n,y} \\ r_{n,z} \end{pmatrix}$$

and the vector $\vec{r_i}$ contains the $x, y, z$ coordinates of particle $i$:

$$\vec{r_i} = \begin{pmatrix} r_{i,x} \\ r_{i,y} \\ r_{i,z} \end{pmatrix}$$

If we have a function $U(\vec{r})$ (which could be our potential energy), we could take the gradient of $U$ with respect to the coordinates of particle $i$:

$$\frac{\partial}{\partial \vec{r_i}} U(\vec{r}) = \nabla_{\vec{r_i}} U(\vec{r}) = \begin{pmatrix} \frac{\partial}{\partial r_{i,x}} U \\ \frac{\partial}{\partial r_{i,y}} U \\ \frac{\partial}{\partial r_{i,z}} U \end{pmatrix}$$