Local Optimization: Gradient Descent
Potential Energy Surface (PES)

Diatomic Molecule (H₂, O₂, etc.)

Energy required to break a bond

Optimal Bondlength
Potential Energy Surface (PES)

Diatomice Molecule (H\textsubscript{2}, O\textsubscript{2}, etc.)

Two Noble Atoms (Ar\textsubscript{2}, etc.)

Potential Energy, $V$

Bond length, $r$

Energy required to break a bond

Optimal Bondlength
How can we generate a PES of a chemical system on a computer?

In general there are two classes for how we can generate a PES:

<table>
<thead>
<tr>
<th>Electronic Structure Methods</th>
<th>Empirical Potentials (Force Fields)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Electronic Structure Methods take into account all electrons and nuclei in an atomic system.

**Hydrogen Molecule**

Potential Energy (V):

\[
V = k_e \frac{e_1 e_2}{r} + k_e \frac{p_1 p_2}{r} - k_e \frac{p_1 e_1}{r} - k_e \frac{p_1 e_2}{r} - k_e \frac{p_2 e_1}{r} - k_e \frac{p_2 e_2}{r}
\]
Empirical Potentials

Assumes a functional form of the PES based on the location of the nuclei and then uses free parameters to fit the potential to have desired properties known from either experiment or electronic structure calculations.

**Morse Potential**

\[ V(r) = D_e \left( 1 - e^{-a(r-r_e)} \right)^2 \]

**Lennard-Jones**

\[ V(r) = 4\varepsilon \left( \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right) \]
How can we generate a PES of a chemical system on a computer?

In general there are two classes for how we can generate a PES:

<table>
<thead>
<tr>
<th>Electronic Structure Methods</th>
<th>Empirical Potentials (Force Fields)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCURATE SLOW Use when accuracy is required (application projects)</td>
<td>NOT AS ACCURATE FAST Use when accuracy is not as important, computational methods are slow or when exploring computational methods where we need to test accuracy</td>
</tr>
</tbody>
</table>
Local Optimization

The Blue Point is:
1. the **ideal bondlength** of the molecule,
2. a **critical point**: a point on a function where the **derivative** equals zero.
3. a **local minimum**: a critical point which is also the lowest point in that region.

The derivative is the slope of the tangent line at a point on a function.
Local Optimization

The Blue Point is:
1. the **ideal bond length** of the molecule,
2. a **critical point**: a point on a function where the derivative equals zero.
3. a **local minimum**: a critical point which is also the lowest point in that region.

Local optimization or **geometry optimization** is the process of finding local minima on a PES using an algorithm or method.
Why do geometry optimization?

• Nature likes to take physical systems to low potential energy states; low energy states of a system are the most probable!

• The **Boltzmann Distribution** describes the probability, \( p \), that you will find a particular configuration of an atomic system, \( r \), in nature with the following equation:

\[
p(r) \propto e^{-\frac{V(r)}{k_bT}}
\]

• Where \( V(r) \) is the potential energy at position \( r \), \( T \) is temperature, and \( k_b \) is a constant called the Boltzmann constant.

• The Boltzmann Distribution will come up again later in the semester!
Why numerical methods?

Another question you may be asking is why not just find the local minimum like you may have done in the past in a Calculus class? i.e. Take the derivative of \(V(r)\), set it equal to zero and solve for \(r\). This is known as the analytical solution.

For realistic chemical systems, we will not have an analytical equation we can write down and solve! So, we will have to use a numerical method.
Gradient Descent

Imagine you roll a ball down a hill let gravity take its course, what direction will the ball first start moving in? What path will the ball take?

It will take the steepest path! Thus, this methods other name, *steepest descent*. 
Gradient Descent

• Why does a ball take the steepest path?
  – The *force* of gravity

• It turns out, there is an equation which relates the force to the potential energy, $V$:

$$F(r) = -\frac{d}{dr} V(r)$$

• Don’t forget *minus* sign!

• Force is a vector meaning it has both a *magnitude* and a *direction*
Algorithm for Local Optimization

The general procedure for numerical geometry optimization is as follows:

1. Calculate the force on all atoms for some configuration of an atomic system.
2. If the magnitude of the force is less than threshold, you have found a critical point! STOP.

So the “threshold” is a small number close to zero.
Algorithm for Local Optimization

The general procedure for numerical geometry optimization is as follows:

1. Calculate the force on all atoms for some configuration of an atomic system.
2. If the magnitude of the force is less than threshold, you have found a critical point! STOP.
3. If not, move the atoms such that they go towards a critical point.
4. Repeat.
Algorithm for Gradient Descent

The general procedure for numerical geometry optimization is as follows:

1. Calculate the force on all atoms for some configuration of an atomic system.
2. If the magnitude of the force is less than threshold, you have found a critical point! STOP.
3. If not, move the atoms such that they go towards a critical point

\[ r_{n+1} = r_n + \alpha F(r_n) \]

4. Repeat.
Example Application of Optimization
Analytically

$$V(r) = (r - 3)^2$$
Example Application of Optimization Analytically

\[ V(r) = (r - 3)^2 \]

To find the local minimum analytically, take the derivative of \( V \), set it equal to 0 and solve for \( r \)

\[ \frac{d}{dr} V(r) = 2(r - 3)^1 = 0 \]

\[ r = 3 \]
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

Start with an initial guess of the local minimum:
\[ r_1 = 1 \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

1. Calculate the force on all atoms for some configuration of an atomic system.

\[ \vec{F}(r) = -\frac{d}{dr} V(r) = -2(r - 3) \]

\[ \vec{F}(1) = -2(1 - 3) = 4 \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2. If the magnitude of the force is less than threshold (0.01eV/Angstrom), you have found a critical point! STOP.

\[
\vec{F}(1) = 4
\]

\[
\|\vec{F}(1)\| = \sqrt{4 \times 4} = 4
\]

\[
4 < 0.01 \implies False
\]
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

3. If not, move the atoms such that they go towards a critical points

\[ r_{n+1} = r_n + \alpha F(r_n) \]

\[ r_2 = r_1 + \alpha F(r_1) = 1 + 0.25 \times 4 = 2 \]

4. Repeat

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Example Gradient Descent**

1. Calculate the force on all atoms for some configuration of an atomic system.

\[ V(r) = (r - 3)^2 \]

\[ \vec{F}(r) = -\frac{d}{dr} V(r) = -2(r - 3) \]

\[ \vec{F}(2) = -2(2 - 3) = 2 \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>r_n</th>
<th>F(r_n)</th>
<th>r_{n+1}=r_n+\alpha F(r_n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

2. If the magnitude of the force is less than threshold (0.01eV/Angstrom), you have found a critical point! STOP.

\[ \vec{F}(2) = 2 \]

\[ \|\vec{F}(2)\| = \sqrt{2 \times 2} = 2 \]

\[ 2 < 0.01 \rightarrow False \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

3. If not, move the atoms such that they go towards a critical points

\[ r_{n+1} = r_n + \alpha F(r_n) \]
\[ r_3 = r_2 + \alpha F(r_2) = 2 + 0.25 \times 2 = 2.5 \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

3. If not, move the atoms such that they go towards a critical points

\[ r_{n+1} = r_n + \alpha F(r_n) \]

\[ r_3 = r_2 + \alpha F(r_2) = 2 + 0.25 \times 2 = 2.5 \]

4. Repeat

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

1. Calculate the magnitude of the force on all atoms for some configuration of an atomic system.

\[ \vec{F}(r) = -\frac{d}{dr}V(r) = -2(r - 3) \]

\[ \vec{F}(2.5) = -2(2.5 - 3) = 1 \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

2. If the force is less than threshold (0.01eV/Angstrom), you have found a critical point! STOP.

\[ \vec{F}(2.5) = 1 \]

\[ \| \vec{F}(2.5) \| = \sqrt{1*1} = 1 \]

\[ 1 < 0.01 \rightarrow \text{False} \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

3. If not, move the atoms such that they go towards a critical points

\[ r_{n+1} = r_n + \alpha F(r_n) \]

\[ r_4 = r_3 + \alpha F(r_3) = 2.5 + 0.25 \times 1 = 2.75 \]

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td>1</td>
<td>2.75</td>
</tr>
</tbody>
</table>
Example Gradient Descent

\[ V(r) = (r - 3)^2 \]

3. If not, move the atoms such that they go towards a critical point:

\[ r_{n+1} = r_n + \alpha F(r_n) \]

\[ r_4 = r_3 + \alpha F(r_3) = 2.5 + 0.25 \times 1 = 2.75 \]

4. Repeat until the magnitude of the force is less than 0.01

<table>
<thead>
<tr>
<th>Step #, n</th>
<th>( r_n )</th>
<th>( F(r_n) )</th>
<th>( r_{n+1} = r_n + \alpha F(r_n) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>2.5</td>
<td>1</td>
<td>2.75</td>
</tr>
</tbody>
</table>
Goals of Lab1-part2

• Explore the Lennard-Jones (L-J) potential
• Implement the gradient descent method for the L-J potential
• Explore the pros and cons of different values of alpha, $\alpha$