

Today's lecture: what are general features of the potential energy surface and how can we locate and characterize minima on it

Derivatives of the potential energy function

Forces

For any atomic configuration, we can compute the force acting on each atom from the negative energy derivative with respect to the particle coordinates:

$$f_{x,i} = -\frac{\partial U(\mathbf{r}^N)}{\partial x_i} \quad f_{y,i} = -\frac{\partial U(\mathbf{r}^N)}{\partial y_i} \quad f_{z,i} = -\frac{\partial U(\mathbf{r}^N)}{\partial z_i}$$

In shorthand notation,

$$\mathbf{f}^N = -\frac{\partial U(\mathbf{r}^N)}{d\mathbf{r}^N}$$

Alternatively,

$$\mathbf{f}^N = -\nabla U(\mathbf{r}^N)$$

That is, the force stems from the gradient of the potential energy.

We can simplify this expression if our potential energy function is built from pairwise interactions,

$$U(\mathbf{r}^N) = \sum_{i<j} u(r_{ij})$$

We have that:

$$r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2$$

Consider computing the force on atom i . There are $N - 1$ terms in the pairwise summation that include the coordinates of atom i . Thus the energy derivative with respect to x_i is:

$$\begin{aligned} f_{x,i} &= -\frac{\partial U}{\partial x_i} \\ &= -\frac{\partial}{\partial x_i} \sum_{j \neq i} u(r_{ij}) \end{aligned}$$

$$= - \sum_{j \neq i} \frac{\partial r_{ij}}{\partial x_i} \frac{\partial}{\partial r_{ij}} u(r_{ij})$$

Using the definition of r_{ij} above and taking the derivative,

$$2r_{ij} \frac{\partial r_{ij}}{\partial x_i} = 2(x_i - x_j)$$

$$\frac{\partial r_{ij}}{\partial x_i} = \frac{(x_i - x_j)}{r_{ij}}$$

Hence,

$$f_{x,i} = - \sum_{j \neq i} \frac{(x_i - x_j)}{r_{ij}} \frac{\partial}{\partial r_{ij}} u(r_{ij})$$

Doing the same for the y and z coordinates,

$$\mathbf{f}_i = \sum_{j \neq i} \frac{\mathbf{r}_{ij}}{r_{ij}} \frac{\partial}{\partial r_{ij}} u(r_{ij})$$

where

$$\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$$

For example, if our pair potential is the so-called **soft-sphere** interaction,

$$u(r) = \epsilon \left(\frac{r}{\sigma} \right)^{-n}$$

where n is a positive integer, then

$$\mathbf{f}_i = \sum_{j \neq i} \frac{\mathbf{r}_{ij}}{r_{ij}} [-n\epsilon\sigma^n r_{ij}^{-n-1}]$$

Simplifying,

$$\mathbf{f}_i = -n \sum_{j \neq i} \frac{\mathbf{r}_{ij}}{r_{ij}^2} \left[\epsilon \left(\frac{r_{ij}}{\sigma} \right)^{-n} \right]$$

Notice that the force acting on atom i due to atom j is just the negative of the force acting on atom j due to atom i , since

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji}$$

In simulation, we typically maintain an (N,3) array of forces. A typical way to implement pairwise force and potential energy calculation in computer code is the following:

```

set the potential energy variable and elements of the force array equal to zero
loop over atom i = 1 to N - 1
  loop over atom j = i+1 to N
    calculate the pair energy of atoms i and j
    add the energy to the potential energy variable
    calculate the force on atom i due to j
    add the force to the array elements for atom i
    subtract the force from the array elements for atom j
  
```

Second derivatives and the Hessian

The curvature of the potential energy function surrounding a given atomic configuration \mathbf{r}^N is given by second derivatives. There are $3N \times 3N$ different second derivatives, attained by cross-permuting different particle coordinates. We typically write the second derivatives as a matrix called the **Hessian**:

$$\mathbf{H}(\mathbf{r}^N) = \begin{bmatrix} \frac{d^2U(\mathbf{r}^N)}{dx_1^2} & \frac{d^2U(\mathbf{r}^N)}{dx_1dy_1} & \cdots & \frac{d^2U(\mathbf{r}^N)}{dx_1dz_N} \\ \frac{d^2U(\mathbf{r}^N)}{dy_1dx_1} & \frac{d^2U(\mathbf{r}^N)}{dy_1dy_1} & \cdots & \frac{d^2U(\mathbf{r}^N)}{dy_1dz_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^2U(\mathbf{r}^N)}{dz_Ndx_1} & \frac{d^2U(\mathbf{r}^N)}{dz_Ndy_1} & \cdots & \frac{d^2U(\mathbf{r}^N)}{dz_Ndz_N} \end{bmatrix}$$

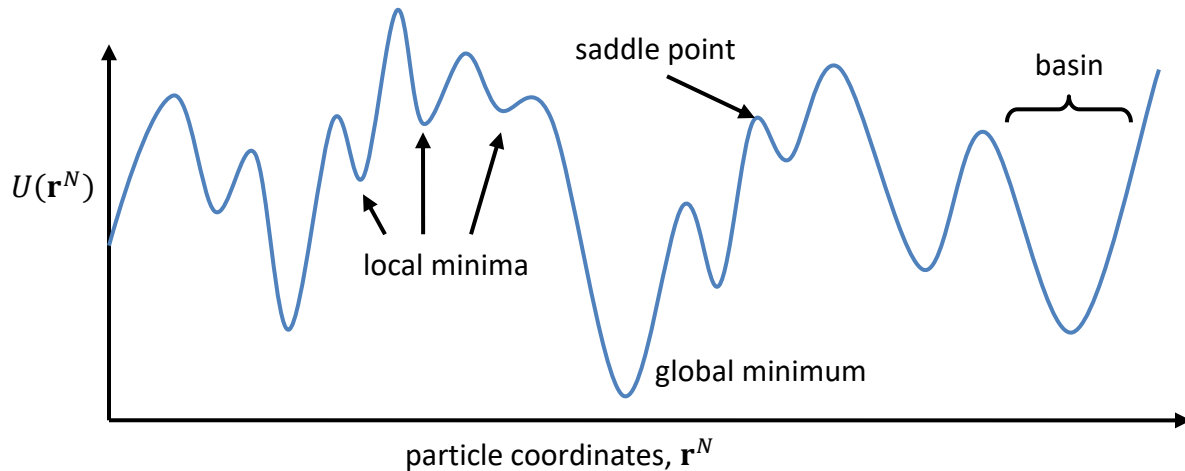
The Hessian is important in distinguishing potential energy minima from saddle points, as discussed in greater detail below. It can be a memory-intensive matrix to compute in molecular simulation, as the number of elements is $9N^2$.

The energy landscape

Basic definition and properties

The potential energy function $U(\mathbf{r}^N)$ defines all of the thermodynamic and kinetic properties of a classical atomic system. We can think of this function as being plotted in a high-dimensional space where there are $3N$ axes, one for each coordinate of every atom. An additional axis measures the potential energy, and we call the projection of the potential energy function in $3N + 1$ dimensional space as the **potential energy landscape** (PEL).

Though we cannot visualize the PEL directly, the following schematic shows some features of it:



Stationary points correspond to the condition where the slope of the PEL is zero:

$$\nabla U(\mathbf{r}^N) = \mathbf{0}$$

This is shorthand for the following simultaneous conditions:

$$\frac{\partial U(\mathbf{r}^N)}{\partial x_i} = 0 \quad \frac{\partial U(\mathbf{r}^N)}{\partial y_i} = 0 \quad \frac{\partial U(\mathbf{r}^N)}{\partial z_i} = 0 \quad i = 1, 2, \dots, N$$

Notice that these equations imply that the net force acting on each particle is zero:

$$\mathbf{f}_i = 0$$

Therefore, the stationary points correspond to **mechanically stable** configurations of the atoms in the system.

There are two basic kinds of stationary points:

- For **minima**, the curvature surrounding the stationary point is positive. That is, any movement away from the point increases the potential energy.
- For **saddles**, one can move in one or more directions around a stationary point and experience a decrease in potential energy.

Scaling of the number of minima and saddle points

PELs contain numerous stationary points. It can be shown [Stillinger, Phys Rev E (1999); Shell et al, Phys Rev Lett (2004)] that both the number of minima and saddle points grow exponentially with the number of atoms:

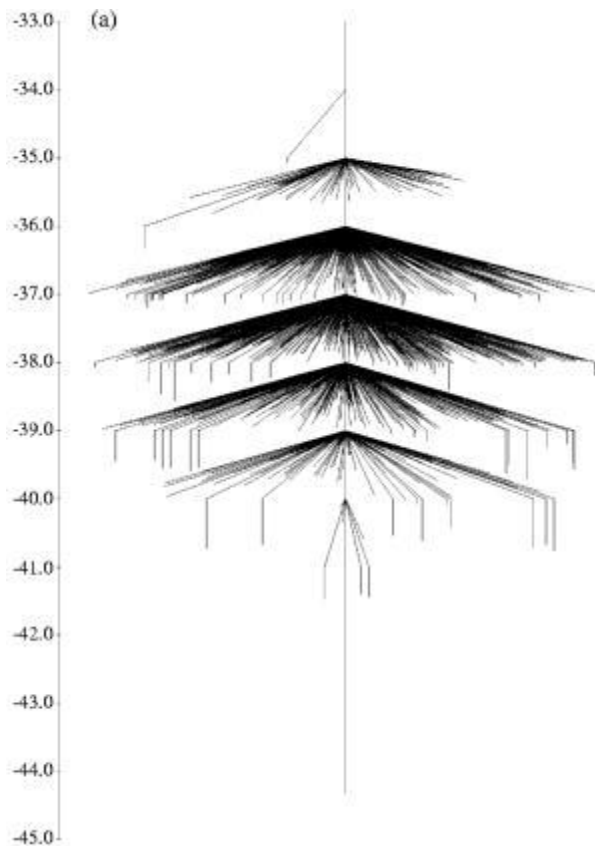
$$\text{number of minima} \sim \exp(\sigma N)$$

$$\text{number of saddle points} \sim \exp(\theta N)$$

where σ and θ are system-specific constants but independent of the number of atoms.

The implication of these scaling laws is that it is extraordinarily difficult to search exhaustively for all of the minima for even very small systems (~ 100 atoms), since their number is so great. By extension, it is extremely challenging to locate the global potential energy minimum, that is, the set of particle coordinates that gives the lowest potential energy achievable.

The following figure shows a **disconnectivity graph** for a system of just 13 Lennard-Jones atoms in infinite volume, taken from [Doye, Miller, and Wales, J. Chem. Phys. 111, 8417 (1999)].



The horizontal axis marks the potential energy and the lines show the connectivity of different minima. To move from one minimum to another, one must travel the path formed when the lines extending from two minima connect to a common energetic point. This small system has 1467 distinct potential energy minima. These results were generated using an advanced and computationally-intensive minima-finding algorithm.

Relevance of finding minima

Why should we care to find PEL minima?

- **Mechanically stable molecular structures** can be identified. When examining conformational changes within a molecule, minima give the distinct conformational states possible. For changes among many molecules, the minima provide the basic structures associated with physical events in the system (e.g., binding). Both can often be compared with experiment.
- Differences in the energies of different minimized structures provide a first-order perspective on the **relative populations** of these (neglecting entropies). Again, this can be compared to experiment and often is an important force field diagnostic.
- **Initial structures** in molecular simulations often need to be “relaxed” away from high-energy states. Minimization can take an initial structure and remove atomic overlaps or distorted bond and angle degrees of freedom.

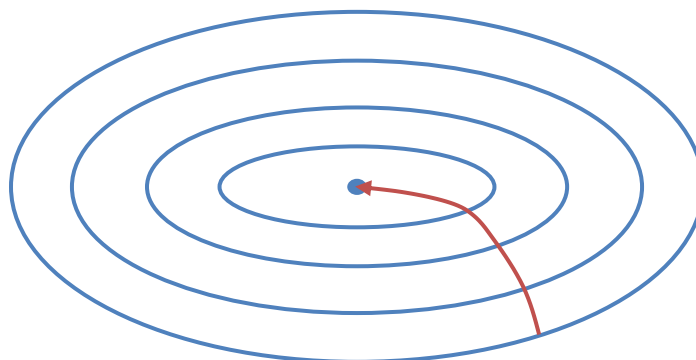
Locating local energy minima

Problem formulation

If we are anywhere within the basin of a minimum, we can define a **steepest descent** protocol that takes us to the minimum:

$$\begin{aligned} \frac{d\mathbf{r}^N}{ds} &= -\nabla U(\mathbf{r}^N) \quad (\lim s \rightarrow \infty) \\ &= \mathbf{f}^N(\mathbf{r}^N) \end{aligned}$$

Here, s is a fictitious time-like variable. The solution to this first order differential equation in the limit that $s \rightarrow \infty$ is the set of coordinates at a minimum \mathbf{r}_{\min}^N . Here, we essentially follow the negative gradient of the energy (i.e., the forces) down the potential energy surface until we reach its bottom—the velocities vectors of each atom follow the negative forces. In a two-dimensional contour plot of energies (as a function of coordinates), this would look like:



The steepest descent protocol is equivalent to an instantaneous quench of a system to absolute zero: we continuously remove energy from the system until we cannot remove any more.

In practice, typically we would use a finite difference approach to implement a steepest descent protocol:

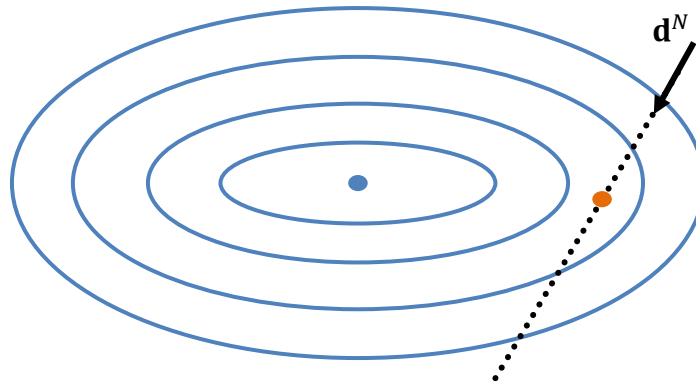
$$\frac{\Delta \mathbf{r}}{\Delta s} = \mathbf{f} \quad \rightarrow \quad \mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} + \Delta s \cdot \mathbf{f}^{(k)}$$

where Δs is a small interval that serves like a time step for propagating positions from one time k to the next, $k + 1$.

Line searches

It is very inefficient to implement the above equation because one must take s to infinity, and because the closer we get to the minimum, the slower we are to converge due to the fact that the forces grow increasingly small. In the above finite difference equation, note that for fixed time step Δs the update to positions decreases in magnitude near the minimum where \mathbf{f} approaches zero.

A number of much more efficient minimization methods are available. These typically use a so-called **line search** as a part of the overall approach. A line search simply says to find a minimum along one particular direction in coordinate space, \mathbf{d}^N . In two dimensions,



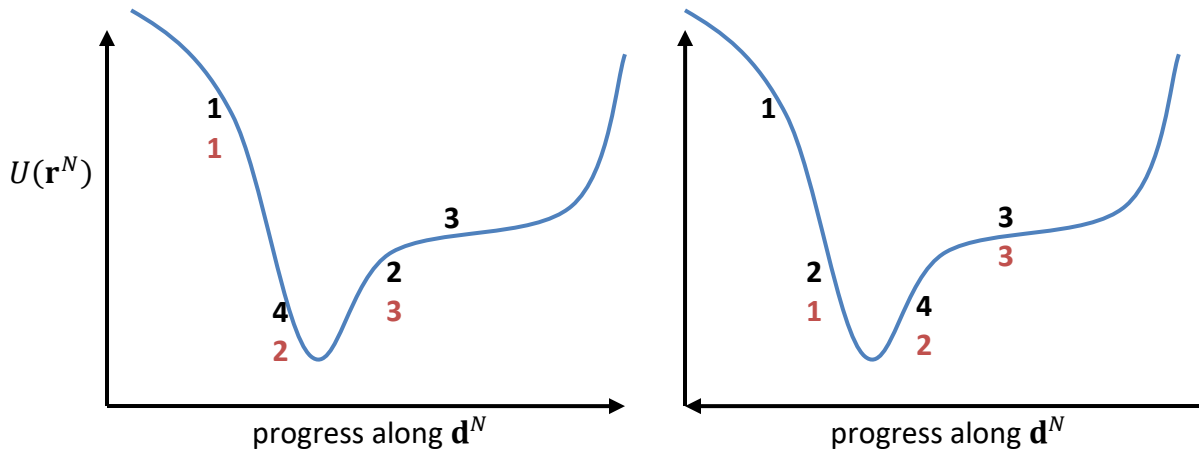
The procedure for a line search is broken into two parts.

1. Start with an initial set of coordinates \mathbf{r}_0^N and a search direction \mathbf{d}^N , chosen to be in the downhill direction of the energy surface.
2. Let $\mathbf{r}_1^N = \mathbf{r}_0^N$. Then, find two more coordinates $\mathbf{r}_2^N = \mathbf{r}_0^N + \delta \mathbf{d}^N$ and $\mathbf{r}_3^N = \mathbf{r}_0^N + 2\delta \mathbf{d}^N$ by taking small steps δ along the search direction. δ should be chosen to be small relative to the scales of energy changes with the particle coordinates.

3. Is $U(\mathbf{r}_3^N) > U(\mathbf{r}_2^N)$? If so, the three coordinates bracket a minimum and we should move to the next part of the algorithm.
4. If not, we need to keep searching along this direction. Let $\mathbf{r}_1^N \leftarrow \mathbf{r}_2^N$, $\mathbf{r}_2^N \leftarrow \mathbf{r}_3^N$, and $\mathbf{r}_3^N \leftarrow \mathbf{r}_3^N + \delta \mathbf{d}^N$. Go back to step 3.

The second part of the algorithm consists of finding the minimum once we've bracketed it. One possibility is to bisect pairs of coordinates:

1. Find a new coordinate \mathbf{r}_4^N . If $|\mathbf{r}_1^N - \mathbf{r}_2^N| > |\mathbf{r}_2^N - \mathbf{r}_3^N|$, pick $\mathbf{r}_4^N = \frac{1}{2}(\mathbf{r}_1^N + \mathbf{r}_2^N)$. Otherwise, pick $\mathbf{r}_4^N = \frac{1}{2}(\mathbf{r}_2^N + \mathbf{r}_3^N)$.
2. Of the four coordinates $(\mathbf{r}_1^N, \mathbf{r}_2^N, \mathbf{r}_3^N, \mathbf{r}_4^N)$ pick new values for $(\mathbf{r}_1^N, \mathbf{r}_2^N, \mathbf{r}_3^N)$ from these that most closely bracket the minimum. Whichever distance produces the lowest energy should be the new middle coordinate ("2").
3. Go back to step 1 until the successive values of the energies found at each new \mathbf{r}_4^N changes by less than some fractional tolerance, typically 1 in $10^6 - 10^8$.



An alternative to step 1 here, which converges faster, is to use the three points $(\mathbf{r}_1^N, \mathbf{r}_2^N, \mathbf{r}_3^N)$ to fit a parabola and find the coordinates at the minimum from this approximation:

$$d_{12} = |\mathbf{r}_2^N - \mathbf{r}_1^N| \quad d_{13} = |\mathbf{r}_3^N - \mathbf{r}_1^N|$$

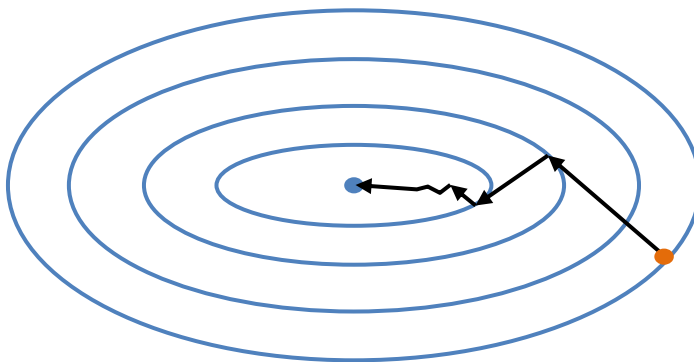
$$\Delta U_{12} = U(\mathbf{r}_2^N) - U(\mathbf{r}_1^N) \quad \Delta U_{13} = U(\mathbf{r}_3^N) - U(\mathbf{r}_1^N)$$

$$\mathbf{r}_4^N = \mathbf{r}_1^N + \frac{1}{2} \left(\frac{\mathbf{r}_2^N - \mathbf{r}_1^N}{|\mathbf{r}_2^N - \mathbf{r}_1^N|} \right) \left(\frac{d_{13}^2 \Delta U_{12} - d_{12}^2 \Delta U_{13}}{d_{13} \Delta U_{12} - d_{12} \Delta U_{13}} \right)$$

Steepest descent with line search

The steepest descent approach can be made more efficient with a line search:

1. From the current set of coordinates, compute the forces \mathbf{f}^N . Choose the search direction to be equal to $\mathbf{d}^N = -\mathbf{f}^N$, the negative gradient of the potential energy surface.
2. Perform a line search to find the minimum along \mathbf{d}^N .
3. Compute the new set of forces at the configuration found in point 2 and a new search direction from them.
4. Repeat step 2 until the minimum energy found with each iteration no longer changes by some fractional tolerance, typically 1 in 10^8 .



Conjugate-gradient method

For long, narrow valleys in the PEL, the steepest descent + line search method results in a large number of zigzag-type motions as one approaches the minimum. This is because the approach continuously overcorrects itself when computing the new search direction with each iteration (new line search). Instead, a more efficient approach is to use the **conjugate gradient (CG) method**.

The CG approach differs only in the choice of a new search direction with each iteration in step 3. Like the steepest descent + line search approach, we use the gradient of the energy surface. In addition, however, we use the previous search direction as well. The CG method chooses search directions that are conjugate (perpendicular) to previous gradients found. A new search direction at the i th iteration is found using

$$\mathbf{d}_i^N = \mathbf{f}_i^N + \gamma_i \mathbf{d}_{i-1}^N$$

with

$$\gamma_i = \frac{\mathbf{f}_i^N \cdot \mathbf{f}_i^N}{\mathbf{f}_{i-1}^N \cdot \mathbf{f}_{i-1}^N}$$

Here, the dot product signifies the 3N-dimensional vector multiplication:

$$\mathbf{f}^N \cdot \mathbf{f}^N = f_{x,1}^2 + f_{y,1}^2 + f_{z,1}^2 + f_{x,2}^2 + \dots + f_{z,N}^2$$

For n -dimensional quadratic functions, the CG approach converges exactly to the minimum in n iterations (e.g., n line search operations).

An alternative equation for γ is typically used in molecular modeling, as it has better convergence properties:

$$\gamma_i = \frac{(\mathbf{f}_i^N - \mathbf{f}_{i-1}^N) \cdot \mathbf{f}_i^N}{\mathbf{f}_{i-1}^N \cdot \mathbf{f}_{i-1}^N}$$

Newton and Quasi-Newton methods

Location of minima can be enhanced by using second derivative information. Consider the PEL in the vicinity of a minimum. For the sake of simplicity, we will consider a one-dimensional function $U(x)$. We can Taylor expand $U(x)$ and its derivative to second order:

$$U(x) = U(x_0) + (x - x_0) \frac{dU(x_0)}{dx} + \frac{(x - x_0)^2}{2} \frac{d^2U(x_0)}{dx^2} \dots$$

$$\frac{dU(x)}{dx} = \frac{dU(x_0)}{dx} + (x - x_0) \frac{d^2U(x_0)}{dx^2} \dots$$

Neglecting higher terms, we can solve for the point at which the derivative is zero:

$$x = x_0 - \left(\frac{dU(x_0)}{dx} \right) \left(\frac{d^2U(x_0)}{dx^2} \right)^{-1}$$

For quadratic functions, this relation will find the minimum after only a single application. For other functions, this equation defines an iterative procedure whereby the minimum can be located through successive applications of it (each time the function being approximated as quadratic):

$$x_{i+1} = x_i - \left(\frac{dU(x_i)}{dx} \right) \left(\frac{d^2U(x_i)}{dx^2} \right)^{-1}$$

This is the **Newton-Raphson iteration** scheme for finding stationary points. The above procedure will converge to both minima and saddle points. To converge to minima, one must first move to a point on the PEL where the energy landscape has positive curvature in every direction.

For multidimensional functions $U(\mathbf{r}^N)$ the update equation above involves an inverse of the Hessian matrix of second derivatives:

$$\begin{aligned}\mathbf{r}_{i+1}^N &= \mathbf{r}_i^N - \nabla U \cdot \mathbf{H}_i^{-1} \\ &= \mathbf{r}_i^N + \mathbf{f}_i^N \cdot \mathbf{H}_i^{-1}\end{aligned}$$

To implement this method in simulation requires computation of the $3N \times 3N$ Hessian matrix and its inverse with each iteration step. For even small systems, this task can become compute- and memory-intensive.

To remedy these issues, several **quasi-Newton methods** are available. The distinctive feature of these approaches is that they generate approximations to the inverse Hessian; upon each iteration, these approximations grow increasingly accurate. These approximations are generated from the force array. The most popular of these in use with molecular modeling methods is the **Limited Broyden-Fletcher-Goldfarb-Shanno (LBFGS)** approach. Implementations of this method can be found at online simulation code databases.

In general, Newton and quasi-Newton methods are more accurate and efficient than the conjugate gradient approach; however, both are widely used in the literature.

Locating global energy minima

Methods to find global potential energy minima are actively researched and have important applications in crystal and macromolecular structure prediction. The global minimum problem is much more challenging as it requires a detailed search of the entire PEL and it is difficult if not impossible to know *a priori* when the global minimum has been found (unless the entire surface has been explored exhaustively).

Typically global-minimizers combine local minimization with various other methods to selectively explore low-energy regions of the PEL and with bookkeeping routines to remember which parts have been searched and what low-energy configurations have been found. This search is usually partially **stochastic** in nature, often leveraging the molecular dynamics and Monte Carlo techniques we will discuss in later lectures.

Normal mode analysis

For minima

Consider the PEL in the vicinity of a minimum. Let's write the coordinates of all the atoms about this minimum be

$$\mathbf{r}_0^N$$

If we consider very small perturbations about these coordinates,

$$\mathbf{r}^N = \mathbf{r}_0^N + \Delta\mathbf{r}^N$$

we can Taylor expand the potential energy to second order. That is, we can perform a **harmonic approximation** to the PEL about a minimum:

$$U(\mathbf{r}^N) = U(\mathbf{r}_0^N) + \frac{\Delta x_1^2}{2} \frac{d^2 U(\mathbf{r}_0^N)}{dx_1^2} + \frac{\Delta x_1 \Delta y_1}{2} \frac{d^2 U(\mathbf{r}_0^N)}{dx_1 dy_1} + \dots + \frac{\Delta z_N^2}{2} \frac{d^2 U(\mathbf{r}_0^N)}{dz_N^2}$$

Notice that all of the first-order terms in the Taylor expansion are zero by the fact that $U(\mathbf{r}_0^N)$ is a minimum. A shorter way of writing this is in matrix form using the Hessian:

$$\Delta U = \frac{1}{2} \begin{bmatrix} \Delta x_1 \\ \Delta y_1 \\ \Delta z_1 \\ \Delta x_2 \\ \vdots \\ \Delta z_N \end{bmatrix}^T \begin{bmatrix} \frac{d^2 U}{dx_1^2} & \frac{d^2 U}{dx_1 dy_1} & \dots & \frac{d^2 U}{dx_1 dz_N} \\ \frac{d^2 U}{dy_1 dx_1} & \frac{d^2 U}{dy_1 dy_1} & \dots & \frac{d^2 U}{dy_1 dz_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^2 U}{dz_N dx_1} & \frac{d^2 U}{dz_N dy_1} & \dots & \frac{d^2 U}{dz_N dz_N} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta y_1 \\ \Delta z_1 \\ \Delta x_2 \\ \vdots \\ \Delta z_N \end{bmatrix}$$

which can be written in more compact form:

$$U = U_0 + \frac{1}{2} \mathbf{q}^T \mathbf{H} \mathbf{q}$$

where $U_0 = U(\mathbf{r}_0^N)$ and $\mathbf{q} = \Delta\mathbf{r}^N$. Keep in mind that the Hessian is evaluated at \mathbf{r}_0^N .

We can use a trick of mathematics to simplify the analysis, by diagonalizing the Hessian matrix \mathbf{H} . Notice that \mathbf{H} is a symmetric matrix, since the order in which the second derivatives are taken doesn't matter. That property means that we can write

$$\mathbf{H} = \mathbf{P}^T \mathbf{D} \mathbf{P}$$

where \mathbf{D} is a diagonal matrix of eigenvalues:

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}$$

and \mathbf{P} is a matrix whose rows are eigenvectors of \mathbf{H} .

Then,

$$\begin{aligned}U &= U_0 + \frac{1}{2} \mathbf{q}^T \mathbf{P}^T \mathbf{D} \mathbf{P} \mathbf{q} \\&= U_0 + \frac{1}{2} (\mathbf{P} \mathbf{q})^T \mathbf{D} (\mathbf{P} \mathbf{q}) \\&= U_0 + \frac{1}{2} \mathbf{s}^T \mathbf{D} \mathbf{s}\end{aligned}$$

In the last line, we defined a new set of $3N$ coordinates, instead of $\mathbf{r}(\mathbf{q})$, that consists of linear combinations of the atomic positions:

$$\begin{aligned}\mathbf{s} &= \mathbf{P} \mathbf{q} \\&= \mathbf{P} \mathbf{r}^N - \mathbf{P} \mathbf{r}_0^N\end{aligned}$$

Physically, the new coordinates \mathbf{s} correspond to projecting the atomic positions on a new set of axes given by the eigenvectors of the Hessian.

Using the above Taylor expansion, the fact that \mathbf{D} is a diagonal matrix means that the energy is a simple a sum of $3N$ independent harmonic oscillators along the new directions \mathbf{s} :

$$U = U_0 + \frac{1}{2} \sum_{i=1}^{3N} \lambda_i s_i^2$$

The physical interpretation of this result is the following. Starting from the minimum and at low temperatures, the system oscillates about \mathbf{r}_0^N . The eigenvectors of \mathbf{H} give the **normal modes** along which the system fluctuates, that is, the directions and collective motions of all of the atoms that move in harmonic fashion for each oscillator. There are $3N$ of these modes corresponding to the $3N$ eigenvectors.

The eigenvalues give the corresponding force constants and can be related to the frequencies of fluctuations. If all atomic masses are the same, m , then

$$\omega_i = \sqrt{\lambda_i/m}$$

In the case that the atomic masses differ, a slightly more complicated equation is used [Case, Curr. Op. Struct. Biol. 4, 285 (1994)].

Regarding the eigenvalues,

- All will be positive or zero, since the energy must increase as one moves away from the minimum. Thus, the frequencies are all real.

- If the system has translational symmetry (i.e., atoms can collectively be moved in the x, y, and z directions without a change in energy), then three of the eigenvalues will be zero.
- If the system also has rotational symmetry (i.e., atoms can be collectively rotated about the x, y, and z axes), then an additional three eigenvalues will be zero.

The spectrum of different **normal mode frequencies** informs one of motions on different time-scales:

- High-frequency modes signify very fast degrees of freedoms. The corresponding eigenmodes typically correspond to bond stretching motions, if present. High-frequency modes tend to be **localized**, i.e., involving only a few atoms.
- Low-frequency modes indicate slow degrees of freedoms. These modes tend to be **collective**, involving the cooperative movement of many atoms.

For higher-order stationary points

Minima in the PEL are not the only points that satisfy the relation:

$$\nabla U(\mathbf{r}^N) = 0$$

Saddle points are also defined by this condition. At a saddle point, the slope of the PEL is zero but there exists at least one direction that will correspond to a *decrease* in the energy of a system.

Recall that the PEL is a hypersurface in $3N + 1$ dimensional space. This means that, at any point on the PEL, there are $3N$ orthogonal directions that we can take. As a result, saddles can be categorized by the number of directions that lead to a decrease in potential energy, called the **saddle order**. First order saddles have one direction of negative curvature, second have two, and so on and so-forth. A zeroth order saddle is a minimum and a $3N$ th order saddle is a maximum.

An identical Taylor-expansion and normal mode analysis can be performed at a saddle point in the potential energy surface:

- Unlike minima, saddles will have one or more eigenvalues of the Hessian that are negative. These correspond to a decrease in the potential energy as one moves away from the saddle point.
- The corresponding eigenvectors give the saddle directions that result in this energy decrease.
- The number of negative eigenvalues gives the order of a saddle.

Such considerations can be useful when examining activated processes and barrier-crossing events in molecular systems.