AIMD分子动力学与机器学习初探

主讲人：许楠 博士

2021年12月18-19日

1. https://www.vasp.at
2. https://github.com/deepmodeling/

第三节：AIMD计算结果的后处理

提纲：（14:00-16:30）

1. 使用VMD分析
   - 可视化轨迹
   - 计算正交体系的RDF
2. 基于VASPKIT进行后处理分析
   - 径向分布函数
   - 速度自相关函数
   - 红外光谱等
   - 均方根位移（MSD）、扩散系数
3. 基于Python进行后处理分析
   - 键长键角跟踪
   - 氢键分析

VMD软件的安装

VMD: Visual Molecular Dynamics

- 分子结构、轨迹可视化、渲染[1]
- 简单后处理功能[2]
- 简单结构建模功能[3]

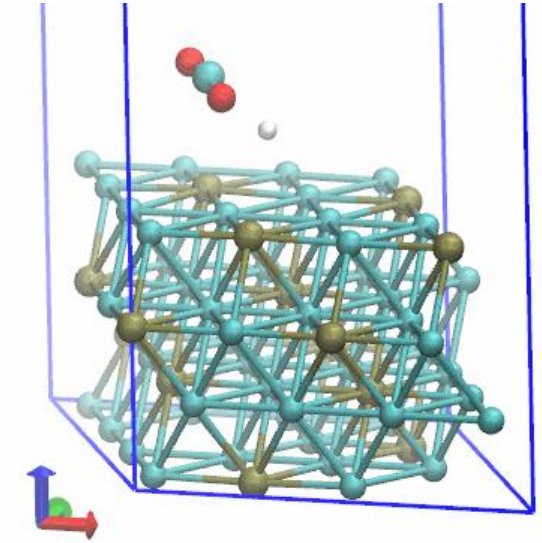
➢ VMD下载网站(需要注册，网速较慢):
https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD

Version 1.9.4 LATEST ALPHA (2021-10-19) Platforms:

Latest pre-release ALPHA test version

- **Source Code**
- **LINUX_64 (RHEL 7+) OpenGL, CUDA, OptiX RTX, OSPRay** (Linux (RHEL 7+)
- **LINUX_64 (RHEL 7+) OpenGL, CUDA, OptiX RTX, OSPRay, RTX RTRT** (Linux RTRT)
- **MacOS 11.x, ARM64 (64-bit "M1" Macs)** (Apple MacOS-X 11 or later)
- **MacOS 10.15, x86_64 (64-bit Intel x86_64)** (Apple MacOS-X 10.15 or later)
- **MacOS X OpenGL (32-bit Intel x86)** (Apple MacOS-X (10.10.x to 10.13.x) with
- **Windows 32-bit (32-bit Intel x86)** (Windows 10)
- **Windows 64-bit, CUDA, OptiX, OSPRay (64-bit Intel x86_64)** (Windows 10)

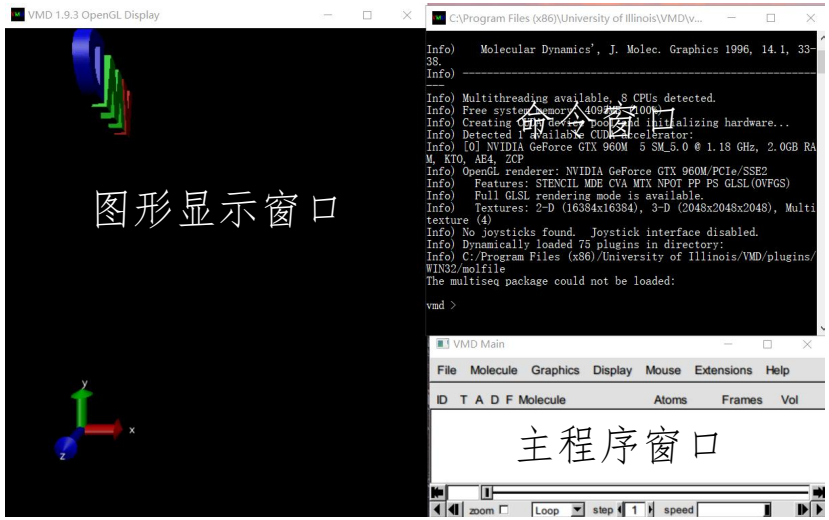
全平台支持，选择对应平台下载

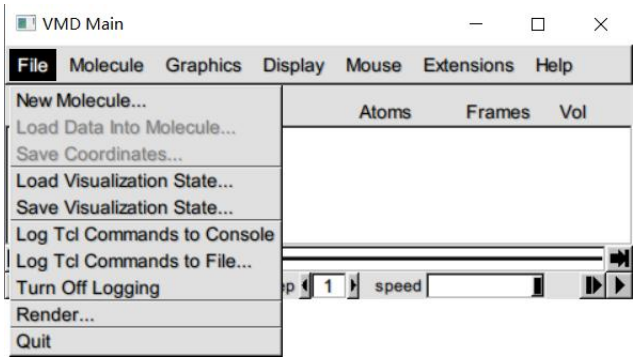


使用VMD渲染的动画

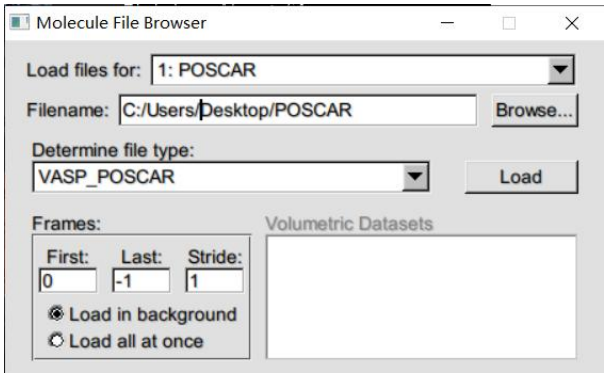
1. https://zhuanlan.zhihu.com/p/398515892
2. https://mp.weixin.qq.com/s/99OwKlbngoUIZmoZWbGFog
3. https://jerkwin.github.io/2020/03/25/VMD%E5%BB%BA%E6%A8%A1%E7%A4%BA%E4%BE%8B/

VMD软件可视化结构





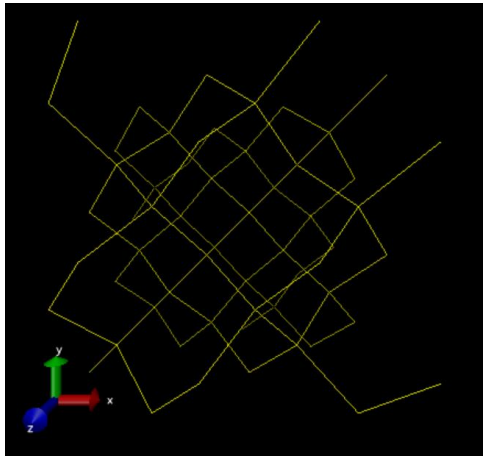
①选择载入新结构



②打开Si的POSCAR文件

VASP_CHGCAR
VASP_OUTCAR
VASP_PARCHG
VASP_POSCAR
VASP_XDATCAR
VASP_XDATCAR5
VASP_xml

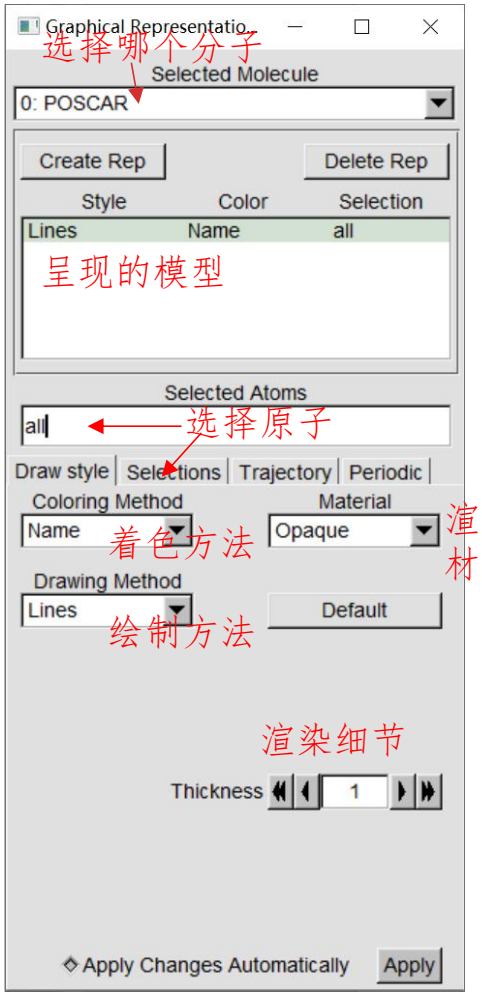
VMD支持的VASP的格式



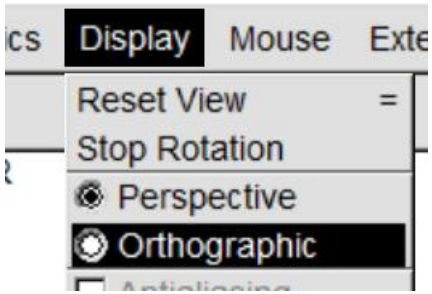
如何美化？

1. http://sobereva.com/545

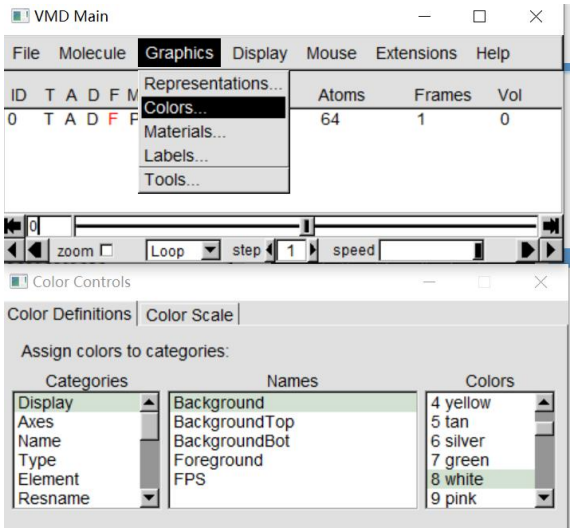
VMD软件操作设置



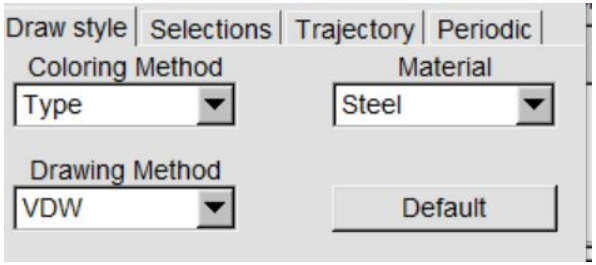
图形显示调整



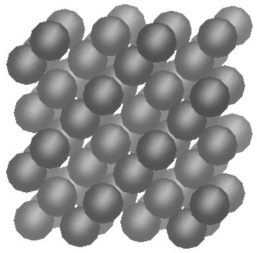
改为正交显示



背景改为白色



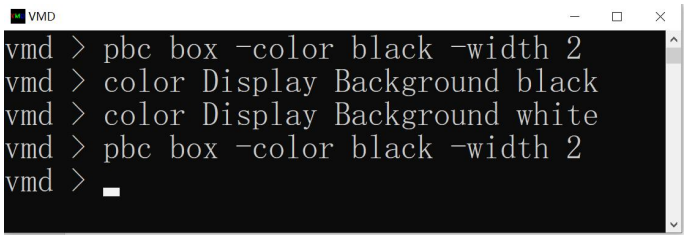
按照原子类型着色，
材质选为金属，绘制方法选择VDW，即VDW球



可按照sob的推荐[1]设置VMD默认的显示模式

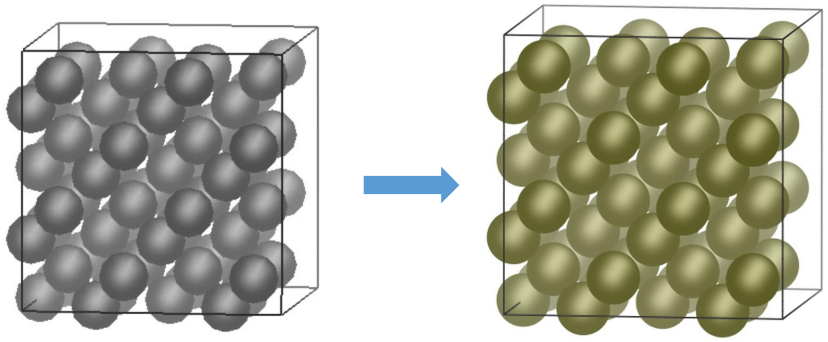
1. http://sobereva.com/545

VMD软件操作设置



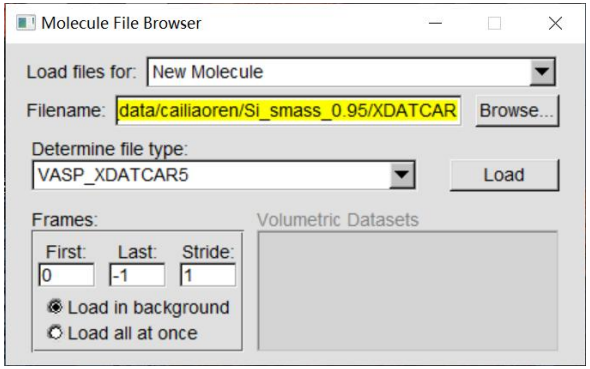
可在命令窗口快捷更改背景、设置盒子边界

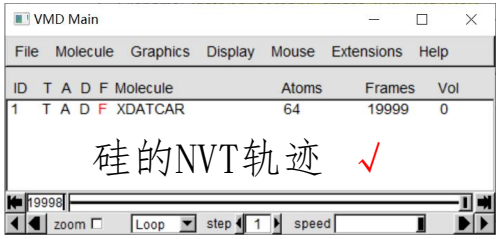
```
color Display Background white
pbc box -color black -width 2
```

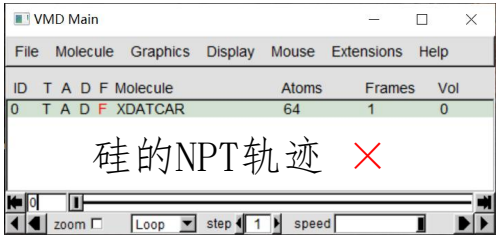


截图分辨率较低，可通过Tachyon渲染[1, 2]

如何导入模拟轨迹？







VMD对于XDATCAR的支持不完善！**不能读取晶格常数变化的轨迹。**

将VASP轨迹转化成其他轨迹格式！

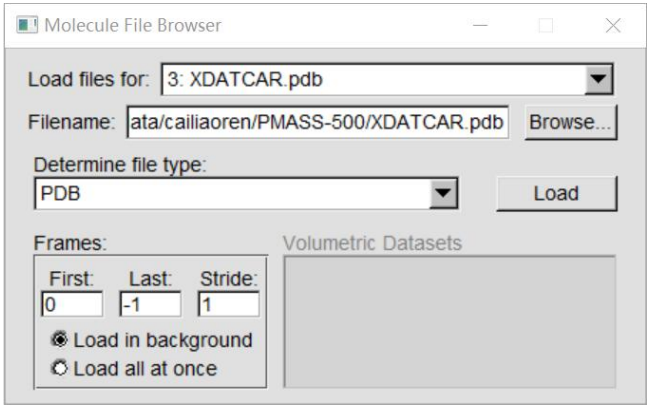
1. https://blog.csdn.net/mt811/article/details/119916295
2. http://bbs.keinsci.com/thread-9178-1-1.html
3. https://mp.weixin.qq.com/s/990wKlbngoUIZmoZWbGFog

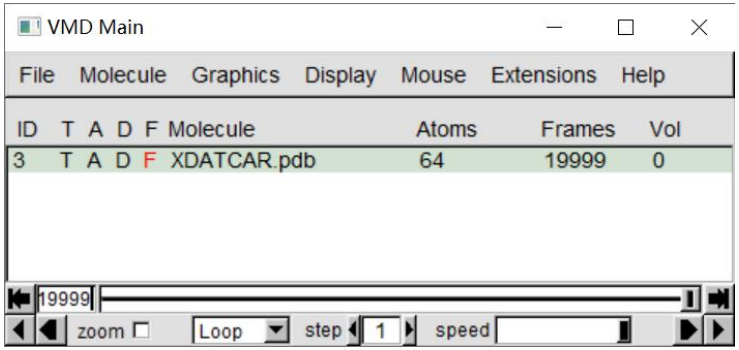
VMD软件操作设置

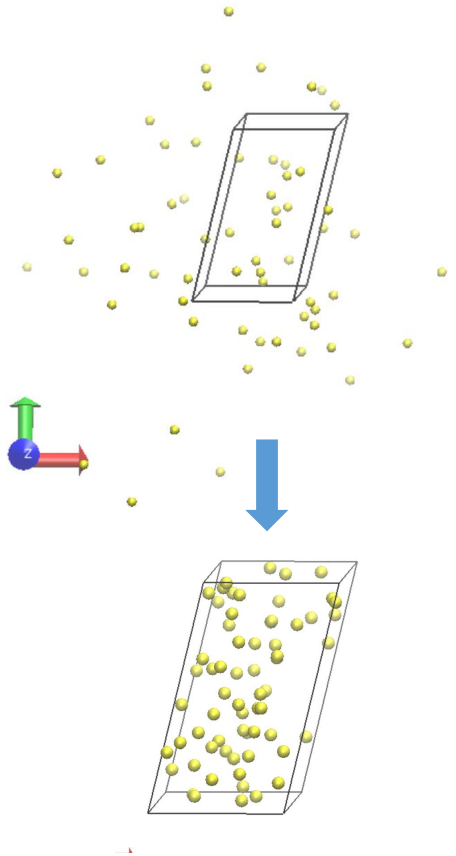
从我的github仓库[1]下载XDATCAR_toolkit.py程序，运行以下命令

python XDATCAR_toolkit.py -p -t 1.0

程序会生成XDATCAR.pdb文件，该文件里包含了多帧轨迹。-p表示开启结构转换，-t 表示时间步为1 fs







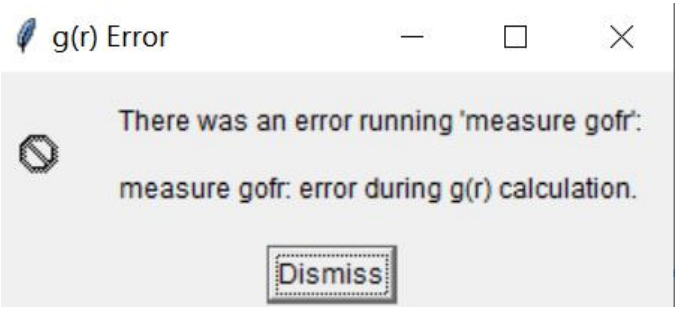
显示的是原子的真实坐标，可能会处于盒子外面，通过 pbc wrap -all 将所有原子“卷”进盒子里

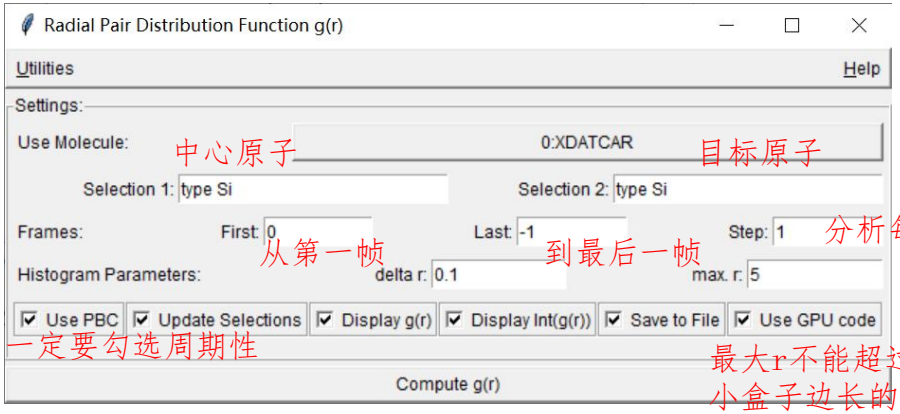
```
vmd > pbc box -color black -width 2
vmd > pbc wrap -all
Info) 0.0% complete (frame 0)
Info) 8.3% complete (frame 1667)
Info) 16.6% complete (frame 3322)
Info) 24.9% complete (frame 4974)
```

1. https://mp.weixin.qq.com/s/990wKlbngoUIZmoZWbGFog

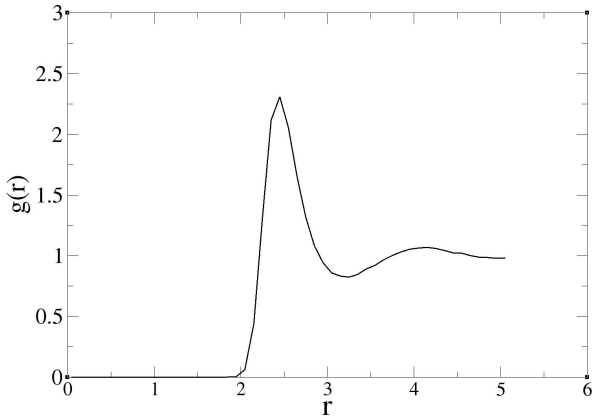
使用VMD计算RDF

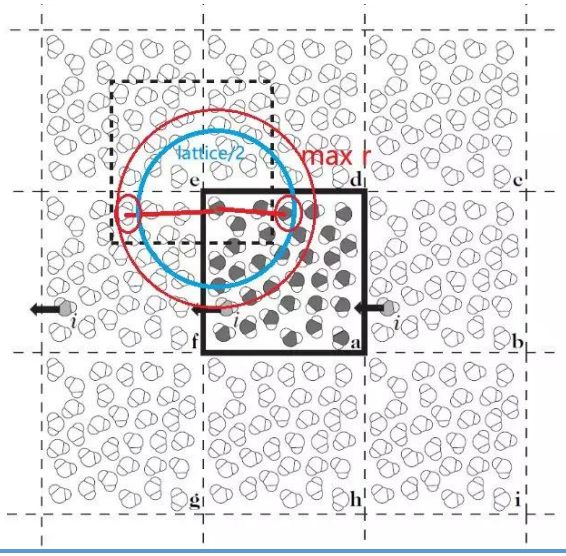
VMD Main→Extensions-Analysis-Radial Pair Distribution Function g(r)





目前VMD只支持分析正交盒子！无法用于NPT结果分析，分析NVT的结果试试





1. https://mp.weixin.qq.com/s/990wKlbngoUIZmoZWbGFog

基于VASPKIT进行后处理分析

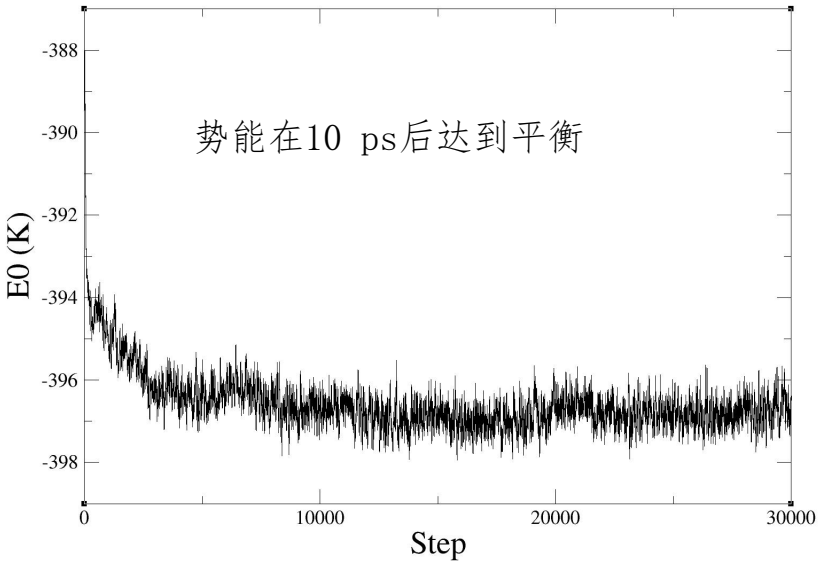
```
================ Molecular-Dynamics Operations ================
720) Energy and Temperature from OUTCAR
721) Average Mean Squared Displacement
722) Average Mean Squared Displacement Using FFT (Recommend)
723) Diffusion Coefficient and Ion Mobility from MSD.dat File
725) Pair Correlation Function from PCDAT File
726) Pair Correlation Function between Two Elements
727) Velocity Autocorrelation Function
728) Vibrational DOS from Velocity Autocorrelation Function
729) MD Trajectories of Selected Atoms with PDB Format

0)   Quit
9)   Back
------------>>
```

- 功能720：从OUTCAR中提取温度和能量。使用grep提取也很方便
- 功能721和722，计算均方根位移，后者使用快速傅里叶变化加速
- 功能723，基于722的结果计算扩散系数
- 功能725，基于VASP输出的PCDAT文件中计算RDF
- 功能726，计算两个元素的RDF
- 功能727，计算速度自相关函数
- 功能728，基于速递自相关函数的结果计算振动态密度
- 功能729，将选中原子的轨迹导出为PDB文件

1. https://zhuanlan.zhihu.com/p/367376130

基于VASPKIT计算RDF

以上一节课模拟的27个水分子的NVT轨迹为例,模拟时长30 ps，温度298 K，模拟在10ps 以后进入平衡




```
================ Molecular-Dynamics Operations ================
720) Energy and Temperature from OUTCAR
721) Average Mean Squared Displacement
722) Average Mean Squared Displacement Using FFT (Recommend)
723) Diffusion Coefficient and Ion Mobility from MSD.dat File
725) Pair Correlation Function from PCDAT File
726) Pair Correlation Function between Two Elements
727) Velocity Autocorrelation Function
728) Vibrational DOS from Velocity Autocorrelation Function
729) MD Trajectories of Selected Atoms with PDB Format

0)   Quit
9)   Back
------------>>
726
```

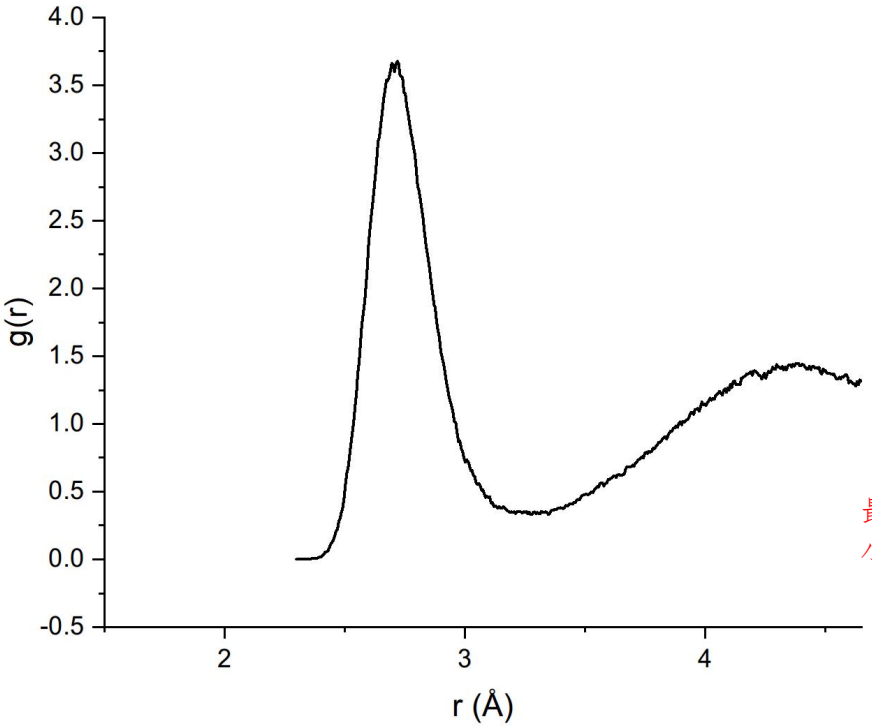
选择功能726，计算两个元素间的径向分布函数

```
Input TWO element-symbol [Elements in POSCAR: O H]

------------>>
O O
How many sampling frames do you want to calculate [1 <= N <=50000]

------------>>
20000
  -->> (03) Reading Input Parameters From INCAR File...
  -->> (04) Reading Atomic Relaxations From XDATCAR File...
  -->> (05) Written PCF.dat File!
```

选择计算O-O的RDF，此处未能选择从哪些帧开始

1. https://zhuanlan.zhihu.com/p/367376130

基于VASPKIT计算RDF



最大r不能超过最小盒子边长的一半

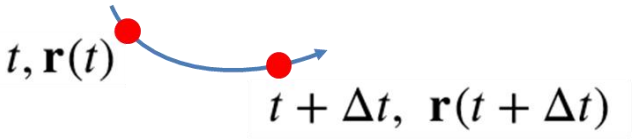
在0原子附近2.8 Å找到另一个0的概率最大

基于VASPKIT计算速度自相关函数

速度自时间关联函数：描述速度与接下来时刻的速度的关系函数

$$C(\tau) = \langle \mathbf{v}(t_0) \cdot \mathbf{v}(t_0 + \tau) \rangle$$

VASP只保存了不同时刻的位置，如何获得速度？



VASPKIT采用了向前差分的方法近似获得速度：

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t)}{\Delta t}$$

①

```
================ Molecular-Dynamics Operations ================
720) Energy and Temperature from OUTCAR
721) Average Mean Squared Displacement
722) Average Mean Squared Displacement Using FFT (Recommend)
723) Diffusion Coefficient and Ion Mobility from MSD.dat File
725) Pair Correlation Function from PCDAT File
726) Pair Correlation Function between Two Elements
727) Velocity Autocorrelation Function
728) Vibrational DOS from Velocity Autocorrelation Function
729) MD Trajectories of Selected Atoms with PDB Format

0)   Quit
9)   Back
------------>>
727
```

√

选择功能727，计算速度自相关函数

②

```
The SELECTED_ATOMS_LIST file will list which atoms are selected.
Input element-symbol and/or atom-index to choose [1 <= & <= 24]
(Free Format is Compatible, e.g., H O Au 1-4 7 18 all)

------------>>
O H
```

输入 O H 以计算O和H原子的VACF

③

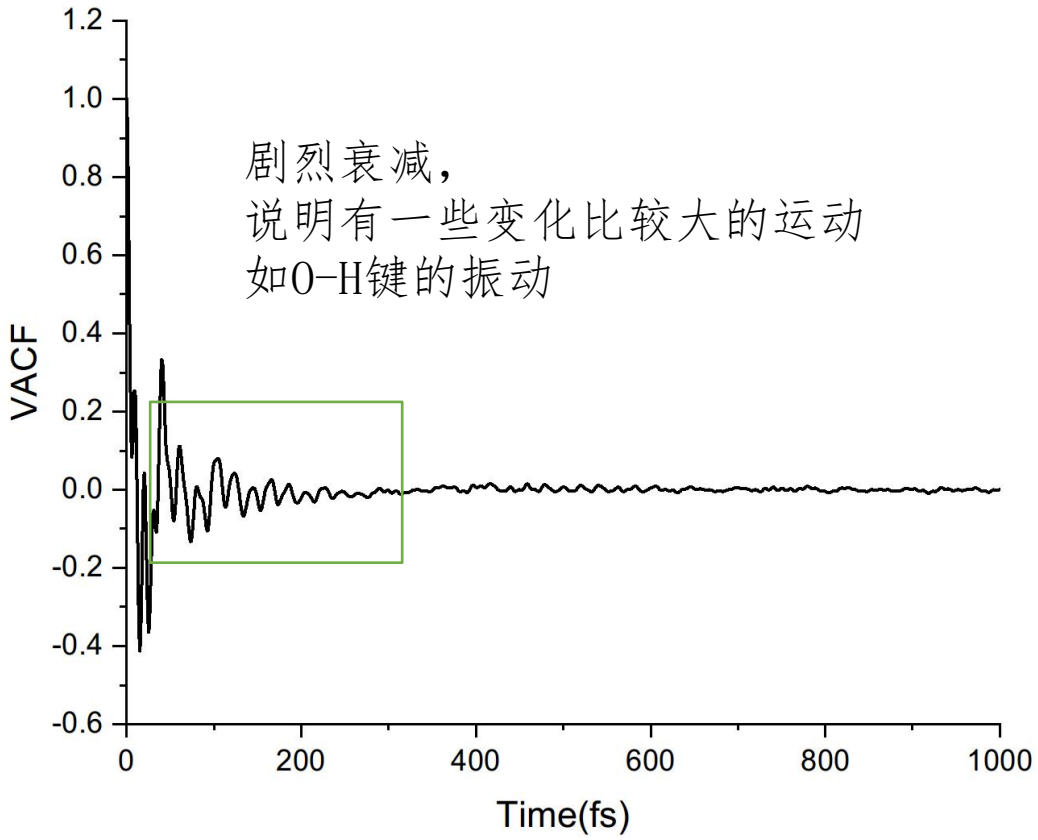
```
How many initial frames you want to skip? [Typical value: < 1000]
------------>>
10000
Which frames will be used to calculate? [Typical value: <=5]
(e.g., 10 means the 1st, 11st, 21st, ... frame will be used)
------------>>
1
 -->> (05) Written VACF.dat File!
```

舍去前10000帧

选择使用每一帧的轨迹，输出O和H的速度自相关函数VACF.dat

1. https://mp.weixin.qq.com/s/0xWx2eTy0_GTg2lXnX5dow

基于VASPKIT计算速度自相关函数



剧烈衰减，
说明有一些变化比较大的运动
如O-H键的振动

当 τ 为0时，函数值最大，τ 增大时，速度 $\mathbf{v}(t_0+\tau)$ 与 $\mathbf{v}(t_0)$ 的关联减弱，函数的数值减小并趋于0。

基于VASPKIT计算振动光谱

计算VDOS(振动态密度)

①

```
================ Molecular-Dynamics Operations ================
720) Energy and Temperature from OUTCAR
721) Average Mean Squared Displacement
722) Average Mean Squared Displacement Using FFT (Recommend)
723) Diffusion Coefficient and Ion Mobility from MSD.dat File
725) Pair Correlation Function from PCDAT File
726) Pair Correlation Function between Two Elements
727) Velocity Autocorrelation Function
728) Vibrational DOS from Velocity Autocorrelation Function
729) MD Trajectories of Selected Atoms with PDF Format

0)   Quit
9)   Back
------------>>
728
```

√

选择功能728，基于VACF计算振动态密度，对应振动光谱，信息更直观

②

```
==================== Energy Unit ================================
Which Energy Unit Do You Want to Adopt?
1) eV
2) nm
3) THz
4) cm^-1
------------>>
4
```

波数选择 cm-1

③

```
The SELECTED_ATOMS_LIST file will list which atoms are selected.
Input element-symbol and/or atom-index to choose [1 <= & <= 24]
(Free Format is Compatible, e.g., H O Au 1-4 7 18 all)

------------>>
O H
```

输入 O H 以计算O和H原子的VDOS

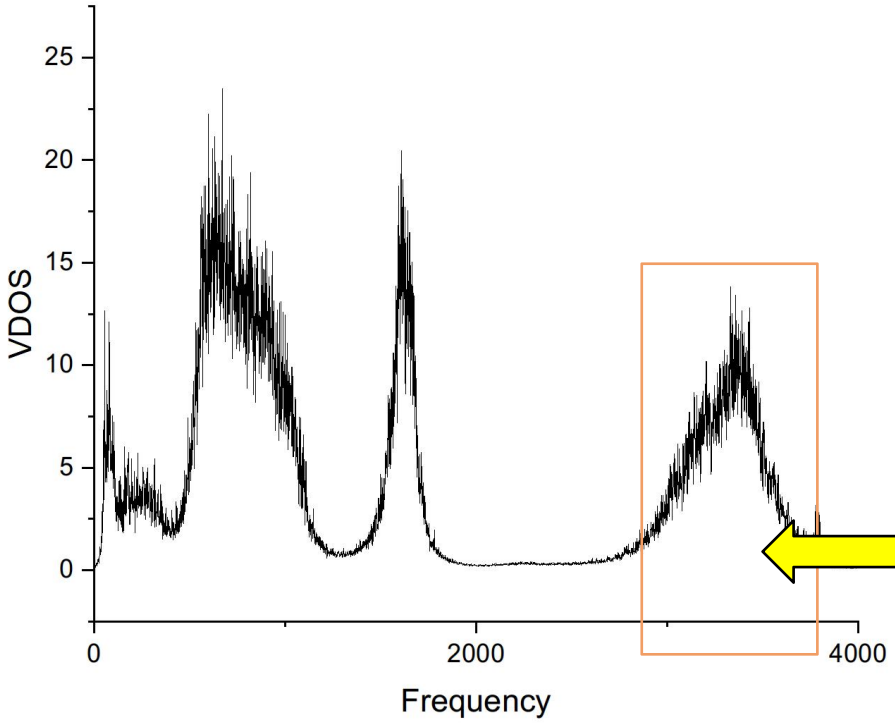
④

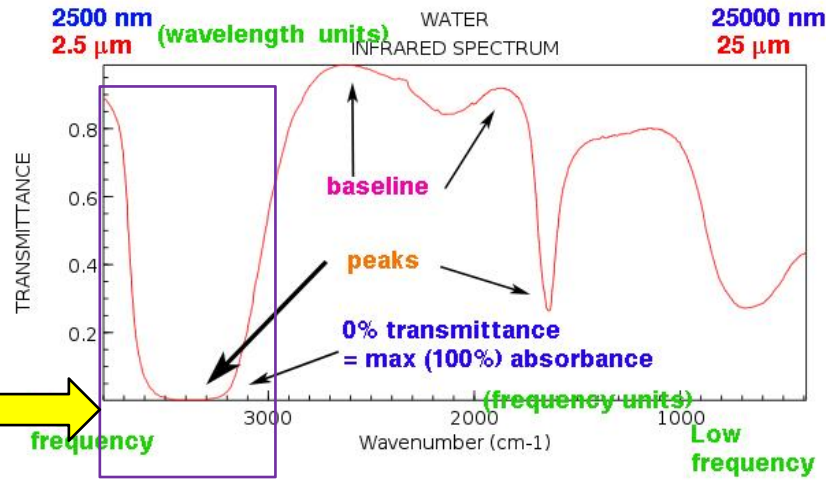
```
>> (04) Reading Atomic Relaxations from XDATCAR File...
How many initial frames do you want to skip? [Typical value: < 1000]

------------>>
10000
Which frames will be used to calculate? [Typical value: <=5]
(e.g., 10 means the 1st, 11st, 21st, ... frame will be used)

------------>>
1
 -->> (05) Written VDOS.dat File!
+----------------------------------------------------------------+
```

舍去前10000帧

选择使用每一帧的轨迹计算VDOS

基于VASPKIT计算红外光谱





羟基的
吸收峰

1. https://mp.weixin.qq.com/s/0xWx2eTy0_GTg2lXnX5dow

基于VASPKIT计算均方根位移

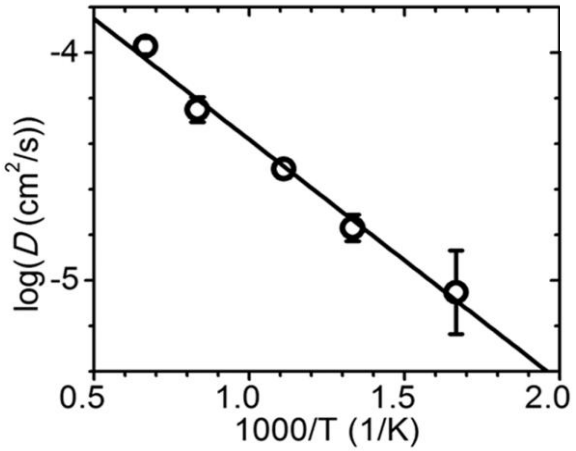
均方根位移(Mean Square Displacement)

$$\mathrm{MSD}(\tau) = \langle \Delta \mathbf{r}(\tau)^2 \rangle = \langle [\mathbf{r}(t_0 + \tau) - \mathbf{r}(t_0)]^2 \rangle$$

τ 为关联时间，t_0 为“时间原点”

$\langle \cdots \rangle$ 为时间平均（不同的时间原点）

常温下水的扩散极慢，在AIMD跑的几十ps范围内，水分子的移动距离很短，故MSD值很低，通常的做法是计算不同的高温下的扩散系数进行外推得到常温的扩散系数。**先以常温水的扩散为例计算MSD。**



①

```
================ Molecular-Dynamics Operations ================
720) Energy and Temperature from OUTCAR
721) Average Mean Squared Displacement
722) Average Mean Squared Displacement Using FFT (Recommend
723) Diffusion Coefficient and Ion Mobility from MSD.dat File
725) Pair Correlation Function from PCDAT File
726) Pair Correlation Function between Two Elements
727) Velocity Autocorrelation Function
728) Vibrational DOS from Velocity Autocorrelation Function
729) MD Trajectories of Selected Atoms with PDB Format

0)   Quit
9)   Back
```

√

选择功能722，利用FFT加速计算MSD

```
------------>>
722
 -->> (01) Reading Structural Parameters from POSCAR File...
 |-------------------------------------------------------------|
 The selected atoms will be listed in SELECTED_ATOMS_LIST file.
 Input element-symbol and/or atom-index to choose [1 <= & <= 81]
 (Free Format is Compatible, e.g., H O Au 1-4 7 18 all)

 ------------>>
0
 -->> (02) Written SELECTED_ATOMS_LIST File!
 -->> (03) Reading Input Parameters From INCAR File...
 -->> (04) Reading Atomic Relaxations From XDATCAR File...
 How many initial frames do you want to skip? [Typical value: < 1000]
 ------------>>
10000
 Which frames will be used to calculate? [Typical value: <=5]
 (e.g., 10 means the 1st, 11st, 21st, ... frame will be used)

 ------------>>
1
 -->> (05) Written MSD.dat and DIFFUSION_COEFFICIENT.dat Files!
```

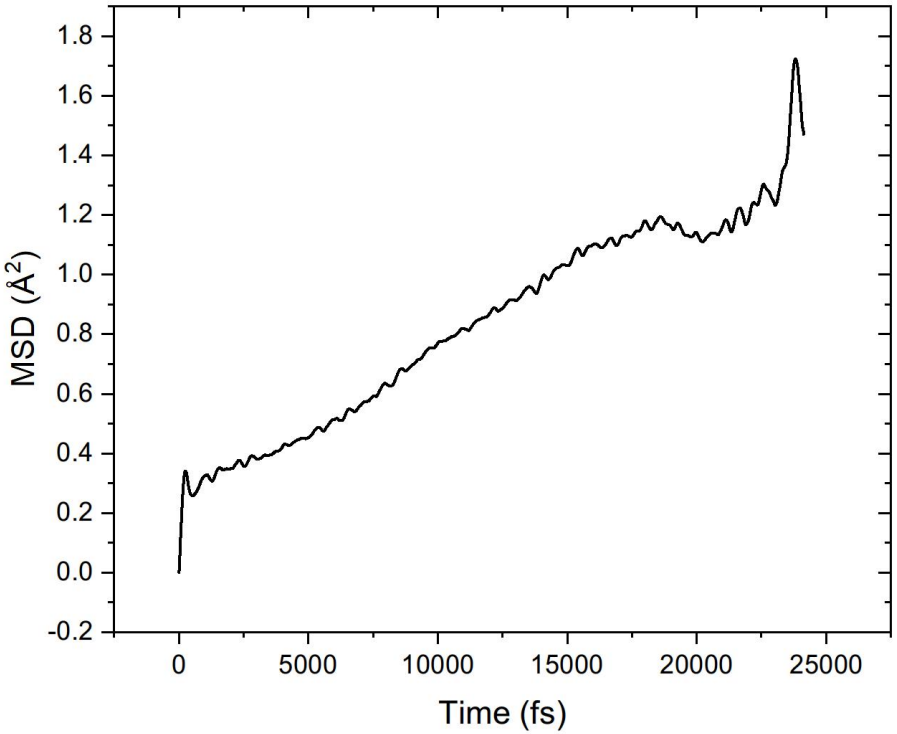
输入 0 以计算氧元素的MSD

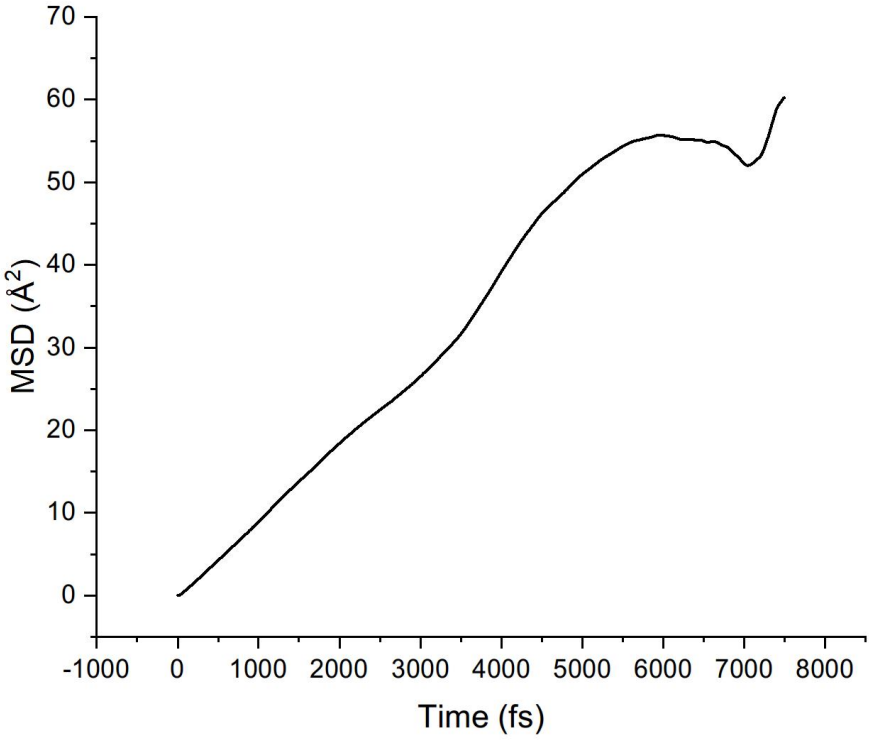
丢弃前10000帧。选择每一帧数据计算。

1. https://pubs.acs.org/doi/10.1021/acs.jctc.5b00574
2. https://pubs.acs.org/doi/10.1021/cm203303y

基于VASPKIT计算均方根位移

使用MSD.dat文件中的第1列（时间和第5列（总MSD）画图）





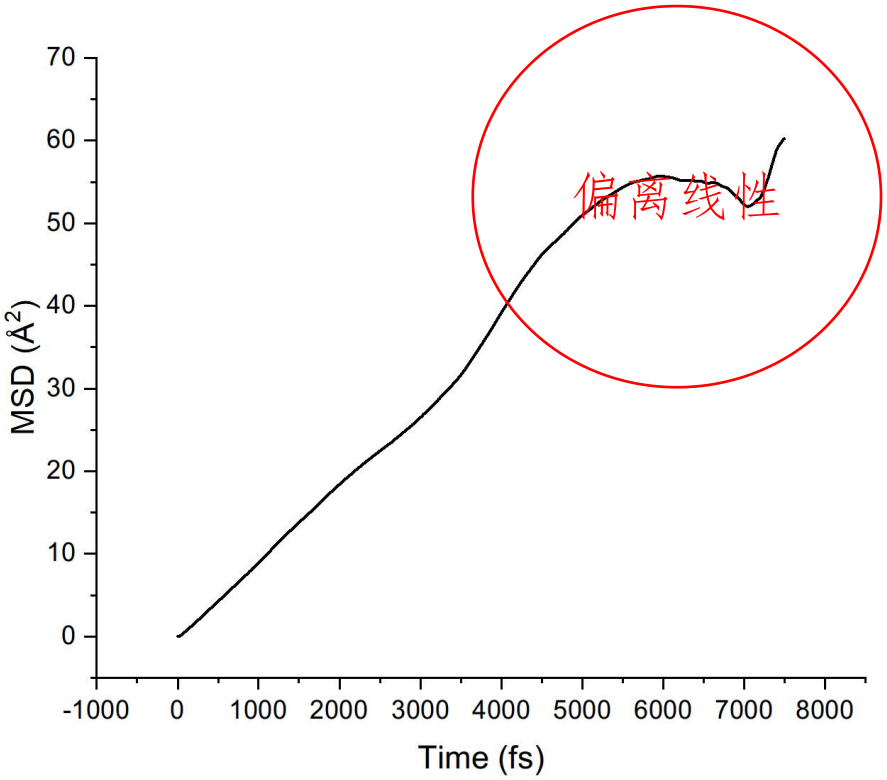
常温下水的扩散极慢，25 ps的MSD才1.6，拟合扩散系数误差极大

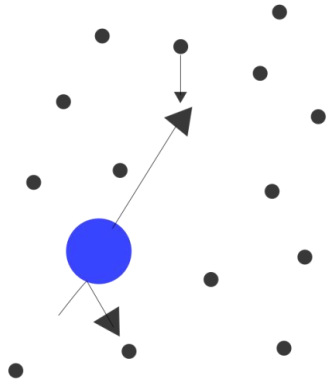
温度提高到800 K，在8 ps的MSD就达到了70

1. https://www.differencebetween.com/difference-between-brownian-motion-and-vs-diffusion/
2. https://jerkwin.github.io/2017/05/02/扩散模式的分类以及扩散系数的计算

基于VASPKIT计算扩散系数

故以800 K的模拟为例子计算扩散系数





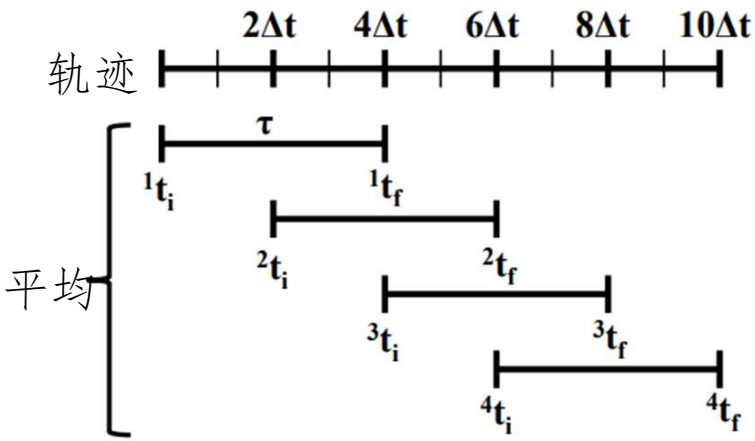
随机扩散
（布朗运动）

$$\mathrm{MSD}(\Delta t) = 6D\Delta t$$

水的扩散属于布朗运动，扩散系数恒定，MSD随关联时间应该是一条直线。为什么关联时间越长，MSD偏离线性？

1. https://www.differencebetween.com/difference-between-brownian-motion-and-vs-diffusion/
2. https://jerkwin.github.io/2017/05/02/扩散模式的分类以及扩散系数的计算

基于VASPKIT计算扩散系数

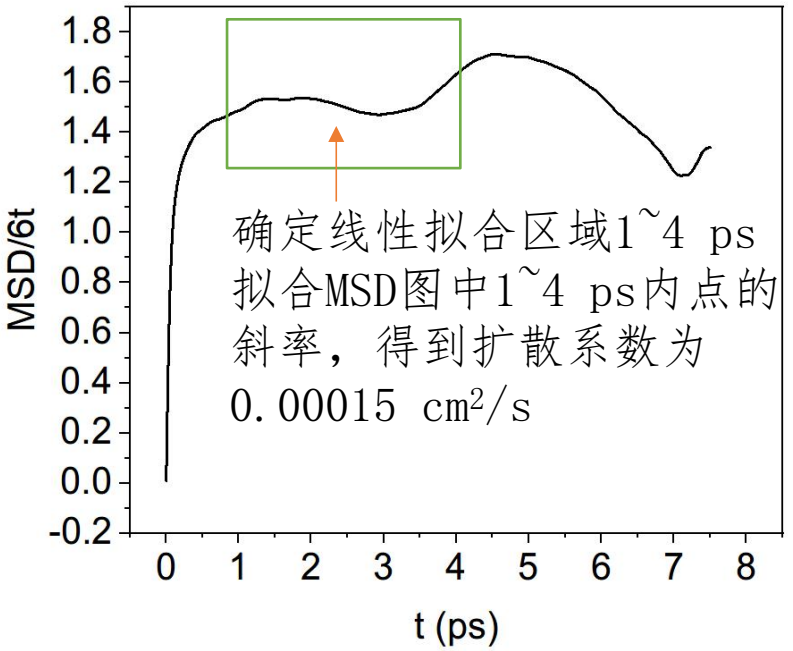


关联时间 τ 越长，数据点越少

$\tau = 4\Delta t$ 4个数据

$\tau = 10\Delta t$ 1个数据，误差大

将DIFFUSION_COEFFICIENT.dat选取第1列(t)和第5列(MSD/6t)作图，查看不同时间点的扩散系数与时间的关系



1. https://www.differencebetween.com/difference-between-brownian-motion-and-vs-diffusion/
2. https://jerkwin.github.io/2017/05/02/扩散模式的分类以及扩散系数的计算

基于Python进行后处理分析

VASPKIT能进行很多分析，但有些情况需要自定义分析，基于轨迹的分析可以使用MDAnalysis帮助分析

```
pip install MDAnalysis==2.0.0
```

但MDAnalysis不支持VASP的轨迹，故需要使用XDATCAR_toolkit.py[1]脚本将XDATCAR转成通用轨迹格式，如PDB。

运行以下命令将水的NVT轨迹转化为XDATCAR.pdb

```
python XDATCAR_toolkit.py -p -t 1.0
```

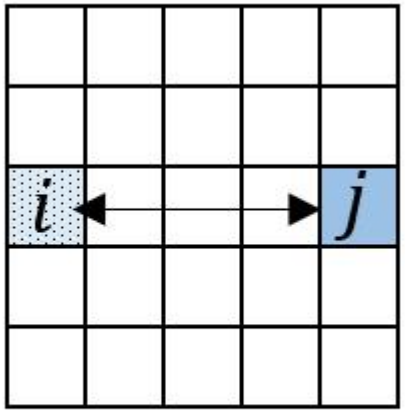
```
import MDAnalysis
u = MDAnalysis.Universe('XDATCAR.pdb', dt=1.0)
```

导入轨迹，设置时间步1 fs

1. https://mp.weixin.qq.com/s/990wKlbngoUIZmoZWbGFog

基于Python分析键长键角的变化

- 如何在计算距离时处理周期性？



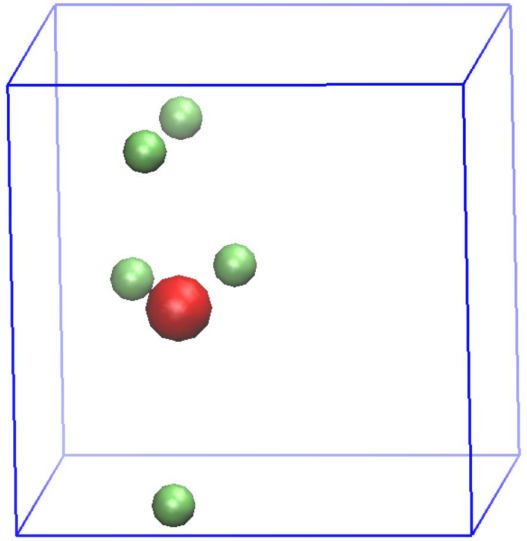
$r_{ij} = 4$
$L = 5$

处理前

$$r_{ij} = r_{ij} - round\left(\frac{r_{ij}}{L}\right) L$$
$$= 4 - 1.0 \times 5$$
$$= -1$$

周期性处理后

- 如何确定O所连接的H原子？

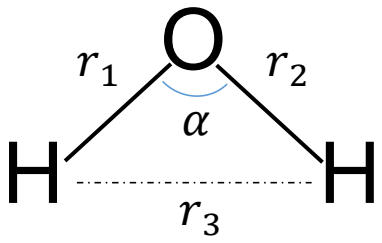


计算O与所有H的距离，选取最近的两个即为与该O连接的H

1. https://jerkwin.github.io/2013/08/13/取整截断函数及其在PBC中的使用

基于Python分析键长键角的变化

- 如何计算键角？



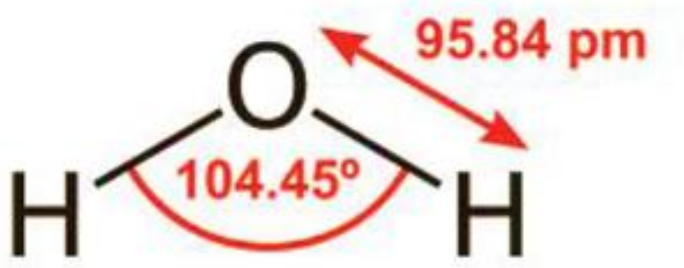
计算O与两个H之间距离以及两个H的距离，利用余弦定理计算键角：

$$\cos\alpha = \frac{r_1^2 + r_2^2 - r_3^2}{2r_1r_2}$$

- 使用Python脚本[1]计算平均键长键角

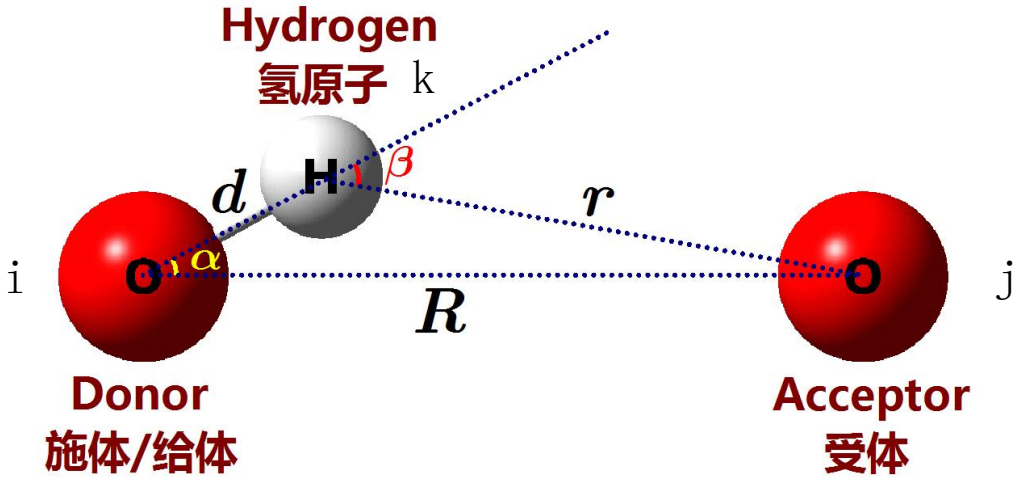
python bond_angle.py

```
Average bond length is 1.00.
Average angle is 106.28.
```



1. https://github.com/tamaswells/VASP_script/blob/master/bond_angle.py

基于Python计算氢键



$$d \approx 1 \quad R \leq 3.5 \quad \alpha \leq 30^{\circ}$$

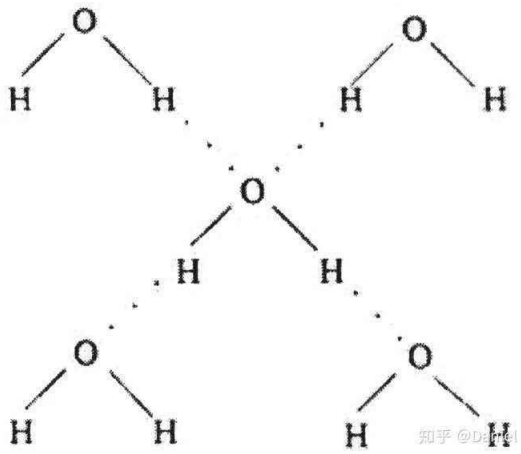
1. 对每一帧进行迭代，对所有的氧原子 j 进行迭代，以其为受体
2. 对所有的氧原子i进行迭代，计算 ij 距离，选择$0 < ij \leq 3.5$的氧原子作为施体。
3. 对所有的氢原子 k 迭代，计算 ik 距离，选择最近的两个原子作为 i 原子的两个H原子。
4. 如果夹角α不大于30°，则记为有效的氢键。

1. http://jerkwin.github.io/2016/12/31/GROMACS和VMD中的氢键判定标准/

基于Python计算氢键

- 使用Python脚本[3]计算平均氢键数

python hb.py
共27个氧原子
平均每帧含有52个氢键，故以每个O原子为受体统计的平均氢键为1.92个，若考虑它作为供体同样形成氢键，平均每个氧原子的氢键为3.85



1. http://jerkwin.github.io/2016/12/31/GROMACS和VMD中的氢键判定标准/
2. https://www.zhihu.com/question/311493740
3. https://github.com/tamaswells/VASP_script/blob/master/hb.py