

数据处理工具包
Data Processing Toolkit
Version 0.4

梁汉普

2019 年 7 月 25 日

目录

1 声明	1
1.1 Licence	1
2 安装	2
2.1 获取软件	2
2.2 依赖环境	2
2.3 安装软件	2
2.4 运行样例	3
3 命令概述	3
3.1 处理能带数据	3
3.2 处理能态密度数据	3
3.3 处理胖能带数据	4
3.4 命令加不加-c 的区别	4
4 输入输出文件概述	4
4.1 输入文件	4
4.2 能带的输出文件	4
4.3 态密度的输出文件	5
4.4 胖能带的输出文件	5
5 DPT 文件夹内包含内容	5

1 声明

Data Processing Toolkit (DPT) 是一个用来处理计算结束后数据的一个工具包。目前版本为测试版 0.4, 尚未成为正式版, 因此如有出现 bug、或者想要添加某些功能, 可以联系邮箱 hanpu-liang@cumt.edu.cn 进行反馈。

1.1 Licence

MIT License

Copyright (c) 2019 Hanpu Liang

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

2 安装

2.1 获取软件

本软件存放于 Github 中, 可以访问网址

<https://github.com/HanpuLiang/Data-Processing-Toolkit>

进行访问与下载。同样的, 你也可以通过在 Linux 系统的终端下输入

```
$ wget -N --no-check-certificate https://github.com/
HanpuLiang/Data-Processing-Toolkit/tree/master/public/DPT
-0.4.tar.gz
```

即可下载得到安装包 DPT-0.4.tar.gz。如果上面两种方法都不能够操作的话, 我们还提供了百度网盘用来存放当前版本的安装包。

百度网盘: <https://pan.baidu.com/s/1rHEtdflV86a4Af0-ElysyQ>。密码: km0r。

2.2 依赖环境

为了能够让本软件顺利的运行下去，你需要拥有以下必要的软件/库：

- Shell 的版本为 bash。
- python 2.7 及以上。并且拥有 numpy 库，后续版本中如果需要可视化作图，则还会需要 matplotlib 库。

2.3 安装软件

在获取了安装包之后，在终端命令行中输入解压命令，将压缩包解压

```
$ tar -zxcf DPT-0.4.tar.gz
```

压缩包解压后会生成一个文件夹 DPT-0.4/，这个文件夹中包含了所有的 DPT 的安装文件与运行文件。然后进入文件夹并运行安装文件

```
$ cd DPT-0.4
$ sh install.sh
```

如果没有报错的成功执行完了脚本的内容，重新开一个终端窗口，就可以尝试运行本软件了。

2.4 运行样例

解压的后出来的文件夹 DPT-0.4/ 中存在一个运行样例。可以用来测试软件是否成功安装。在当前目录下创建一个文件夹，进入后将样例复制到当前目录下

```
$ mkdir test
$ cp ../DPT-0.4/example/electronic_structure/* . -rf
```

然后你的目录下会多出 6 个文件，这 6 个文件就是我们在需要处理的数据的文件。一般情况下，这六个文件由 VASP 生成或者其他软件生成。

之后分别运行

```
$ DPT -b
$ DPT -d
$ DPT -f
```

这三条命令分别代表生成：能带数据文件 (K.dat, Energy.dat)、态密度数据文件 (ATOM_dos.dat) 和 能带数据文件 (all_orbits/)。所有命令在章节3中讲解，生成的数据文件中的格式具体在章节4再讲。如果成功生成了文件，那就说明该软件安装成功。

3 命令概述

3.1 处理能带数据

处理能带数据需要文件

EIGENVAL, OUTCAR, KPOINTS.3

所使用的命令为

```
$ DPT -b
```

或者

```
$ DPT -bc
```

两者的区别在于输出的文件格式不同。前者是将 k 点路径与能带分开保存为 K.dat 和 Energy.dat, 后者则将两者保存在同一个文件 band_data.dat 内。以此方便不同的处理工具进行处理。具体的文件格式则放在[章节4.2](#)中进行详细讲解。

3.2 处理能态密度数据

处理能态密度数据需要的文件有

POSCAR, DOSCAR

所使用的命令为

```
$ DPT -d
```

该命令会生成文件所有原子的态密度的数据文件 ATOM_dos.dat。如果一共有 3 个原子, 则有 3 个文件。具体的文件讲解在[章节4.3](#)中。

3.3 处理胖能带数据

处理胖能带的数据需要的文件有

EIGENVAL, OUTCAR, KPOINTS.3, PROCAR, POSCAR

所使用的命令为

```
$ DPT -f
```

或者

```
$ DPT -fc
```

这两者的区别在于输出的文件格式不同。前者会把每个原子的每个轨道单独生成一个文件, 与能带中各个数据点一一对应, 所有文件放在文件夹 all_orbits/中。而后者则是每个原子一个文件, 该原子的 10 个轨道数据为 10 列。以此方便不同的处理工具进行处理。具体的输出文件讲解在[章节4.4](#)。

3.4 命令加不加-c 的区别

如果加了 c, 则会生成另一种方便处理的文件格式。这种文件格式比较适合利用 Origin 直接导入进行作图。目前只有能带与胖能带可以加 c 处理。

4 输入输出文件概述

4.1 输入文件

不同的命令的输入文件已经在上文中提到过了。总的来说,算能带的话就需要 EIGENVAL, KPOINTS.3, OUTCAR 这三个数据。而胖能带中的原子轨道贡献数据主要放在 PROCAR 中, 能态密度 DOS 的数据主要放在 DOSCAR 中。后面这两个因为是程序自动分隔原子, 所以需要读取 POSCAR 的原子数据。

目前处理的所有数据不支持考虑了自旋等其他内容。

4.2 能带的输出文件

在正常的没有加参数-c 的输出文件为 K.dat 和 Energy.dat。其中 K.dat 为一列数据, 是高对称点的路径, 一共 m 行。Energy.dat 为 $m \times n$ 大小的数据。 m 行与高对称点路径一一对应, n 列代表 n 条能带。

而加了参数-c 后, 输出文件 band_data.dat 则只有两列, 第一列为高对称点路径 k , 第二列为能带。一共是 $m \times n$ 行。每一条能带之间用空行间隔开。

```
# band_data.dat
k    E
```

4.3 态密度的输出文件

态密度只有一种输出文件格式 ATOM_dos.dat。体系中有多少个原子, 就会输出多少个文件, 每个文件是一个原子的所有轨道的态密度。文件中每一列代表如下

```
# ATOM_dos.dat
E    TOTAL    ATOM_s    ATOM_p    ATOM_d
```

其中这个 total 是所有原子的总态密度。

4.4 胖能带的输出文件

胖能带也有两种输出文件格式。

在不加-c 时, 会产生一个 all_orbits/目录。目录中包含着所有原子的所有轨道。每种原子每种轨道单独为一个文件, 譬如: Zn_s.dat 与 Zn_py.dat。大小为 $m \times n$ 。文件中, 每个数据点与能带的不加 c 输出的 Energy.dat 位置一一对应, 都是以一列代表一个能带。第 1 行第 3 个数据点代表 Energy.dat 中第 1 行第 3 个能量值对应的 Zn 原子的 s 轨道的贡献 (例子)。

加了参数-c 后, 则只会生成原子数目的文件 fat_band_ATOM.dat。每个原子一个文件, 其中是每一个轨道为一列, 共为 $m \times n$ 行。每条能带之间用空行隔开。

```
# fat_band_ATOM.dat
k    E    s    py    pz    px    dxy    dyz    dz2    dxz    dx2    tot
```

5 DPT 文件夹内包含内容

- src (源代码)
- doc (用户使用手册)
- example (运行样例)