Weekly Report

Siyu Liao

July 29, 2020 - August 5, 2020

1 Introduction

This week I mainly focus on how band structure forms when atoms get close to each other. Besides, I also try to utilize machine learning to uncover the reasons of indirect band gap.

2 Progress

2.1 Formation of band structure

In this chapter I will use GaP as an example for illustration. First, let's see how band structure would look like when the atoms are far away. Figure 1 shows GaP's band structure when lattice constant equals to 10Å, which is basically consisted of atomic band structure of Ga and P.

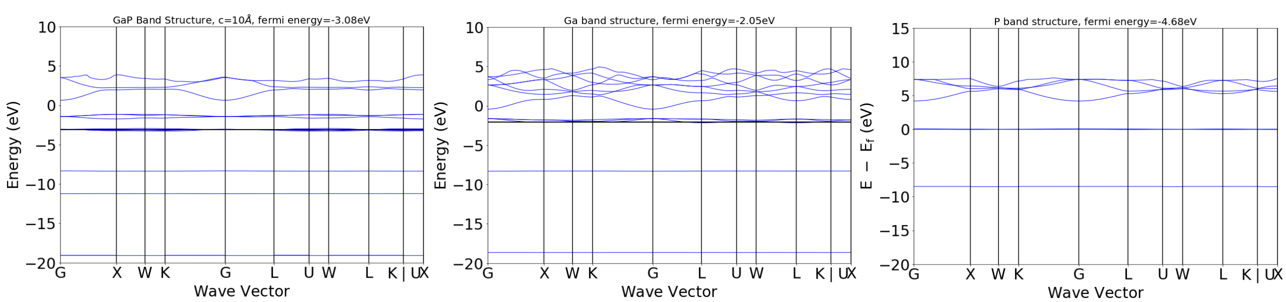

Figure 1: GaP's band structure when lattice constant equals to 10Å, the right two are atomic band structure

When we gradually lower the lattice constant, the orbitals start to couple. The following picture shows band structures and COHPs when lattice constant changes from 10Åto 3Å.

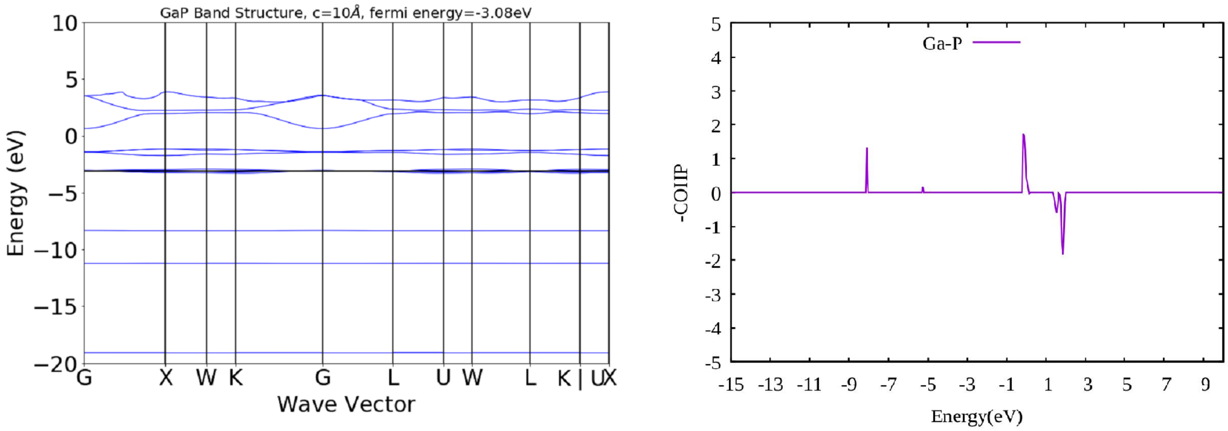

Figure 2: GaP's band structure and COHP when lattice constant equals to 10Å

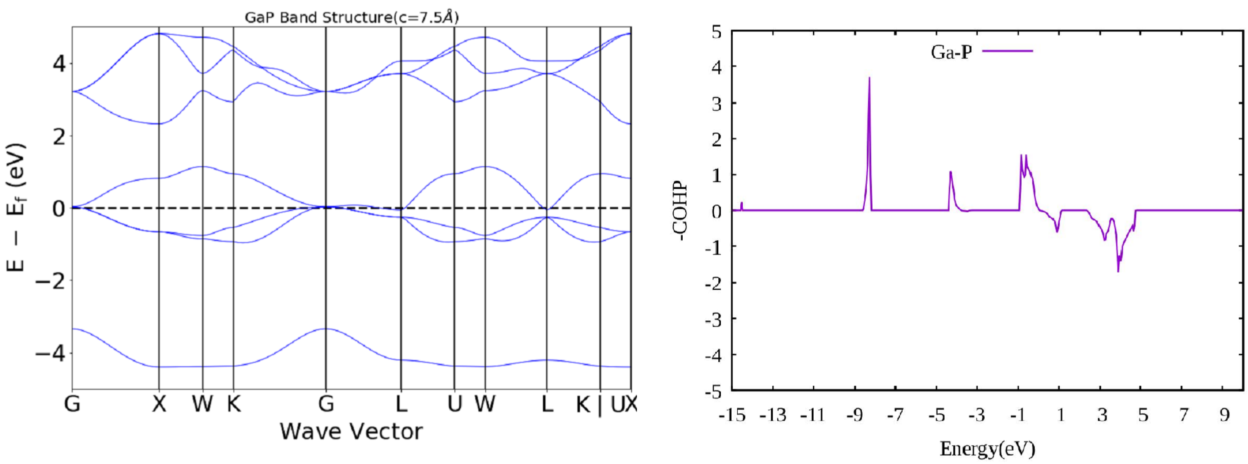

Figure 3: GaP's band structure and COHP when lattice constant equals to 7.5Å

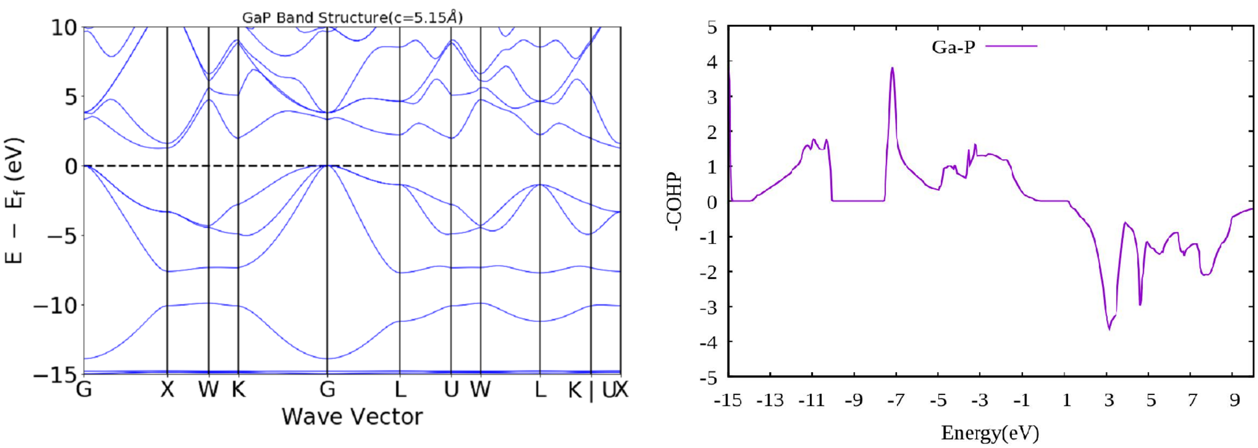

Figure 4: GaP's band structure and COHP when lattice constant equals to 5.15Å

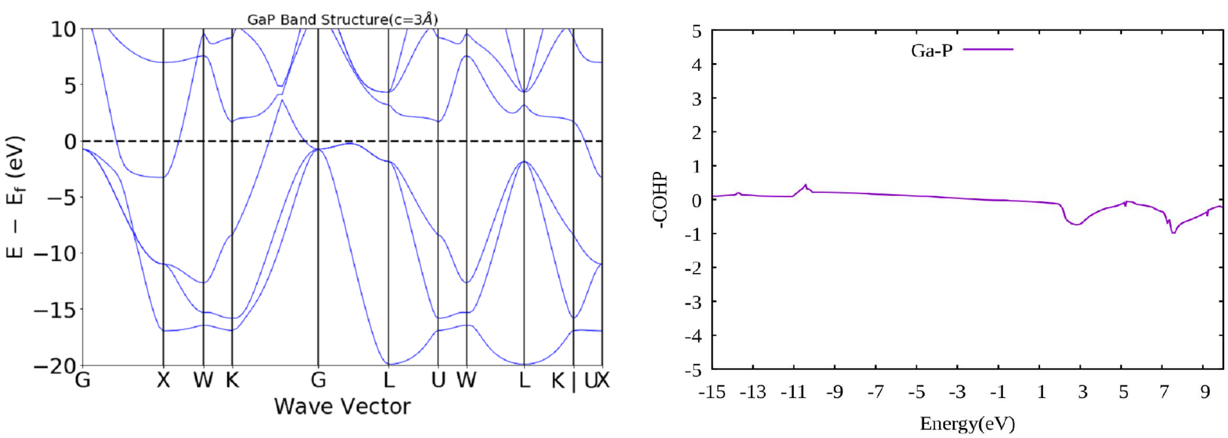

Figure 5: GaP's band structure and COHP when lattice constant equals to 3Å

We can find that when lattice constant decrease, the peak of COHP becomes wider, which suggests the orbital coupling becomes stronger. I hope to observe the intensity of orbital coupling, but they appears to be similiar by lobster. Meanwhile, the COHP of 3Åis weird since it has turned to metallic and the lobsterout file suggests that the charge spilling is up to 16%(usually a correct result is bellow 5%). I'm tring to figure it out.

2.2 Machine learning

Since now I have roughly handled how to analyze indirect band gap of a specific material, I suppose it high time that we could try machine learning to help our research. I know that bond length, eletronegativity, structure, occupied orbital and some other properties would affect the type of band gap, but it's hard for human to take those aspects into account at the time. However, machine learning, thanks to its multi-dimensional elaborative faculty, could make up this drawback as long as we choose suitable descriptors to convert those physical properties into numerical information. So I think maybe machine learning is also a good direction for furthur study.

This week I try to grap enough dataset from Material Project for indirect band gap but I face a big problem. I know that using MP API we should set criteria and properties, but I find that some criteria cannot be extracted by API, such as $'band\_structure'$ and $'band\_gap.is\_direct'$. Even if I put them in the criteria the result will be: $'band\_gap.is\_direct'$: None, which means results from MP API is inadequate for indirect band gap study. While these properties are reserved if we use some other inquiry way. So one solution I can think of is first inquiring materials using MP API and saving its MP-id, then for each materials in the API results re-inquire them using $'get\_bandstructure\_by\_material\_id'$ to get information about band gap.

However, this method has to pin MP thousands of times and I always get error like this: *MPRestError: ('Connection broken: IncompleteRead(0 bytes read)', IncompleteRead(0 bytes read)).* I suppose it's because MP doesn't allow too many inquries at the same time and I'm totally helpless about it. I'm wondering do you know how to dealing with this problem or is there any other ways to get information about indirect band gap?

And for practicing, I have tried predicting the bandgap using data from MP. It's turn out that with easy descriptors and hyper-paramater optimization, we can already get a good result:

R2:0.7912520168751132 RMSE:0.7872993733914431

That convinces me that machine learning is probably useful about some other questions like indirect band gap. I will attach the code in the report(I think .ipynb version is better for presentation since the results are saved, in .py version you may need to change somethings to run the code.)

3 Summary

I think this week my process is mainly about machine learning. The part of workflow is not satisfactory and I still don't know exactly about what to do with the workflow. Looking forward to next discussion.