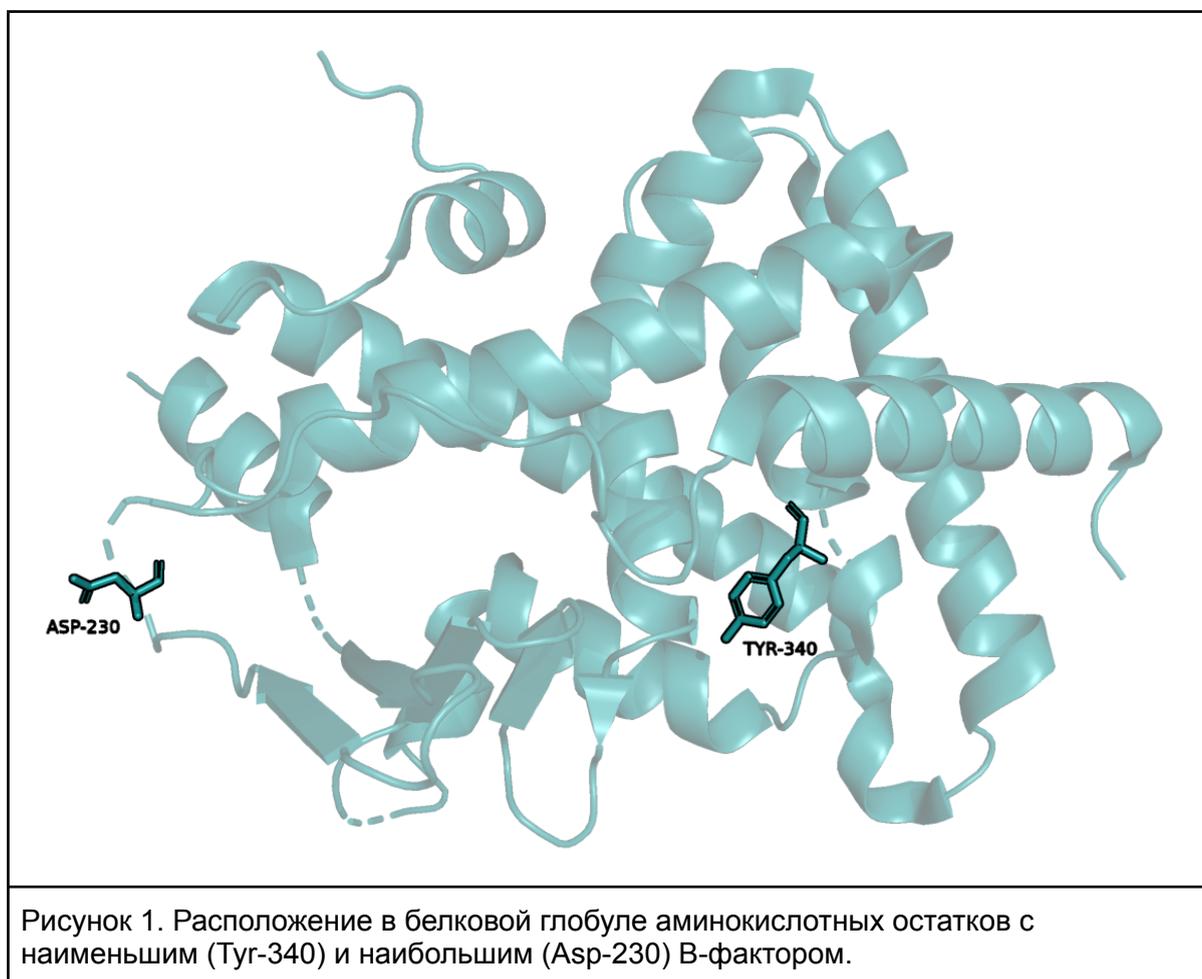


Введение

В рамках данной практической задачи необходимо было освоить некоторые функции python-пакета Prody, который позволяет значительно ускорить процесс анализа белковых структур, а также на полуигрушечном примере понять, как происходит расшифровка белковых молекул с помощью фаз и модулей структурных факторов, а также шум, наблюдаемый в экспериментах, может повлиять на качество расшифровки.

Задания 1 и 2

В рамках данного задания было необходимо найти и проанализировать зависимость В-фактора аминокислотных остатков от расстояния до центра масс белковой молекулы.



Каковы В-факторы соседних аминокислотных остатков? Покрасим всю белковую молекулу по величине В-фактора.

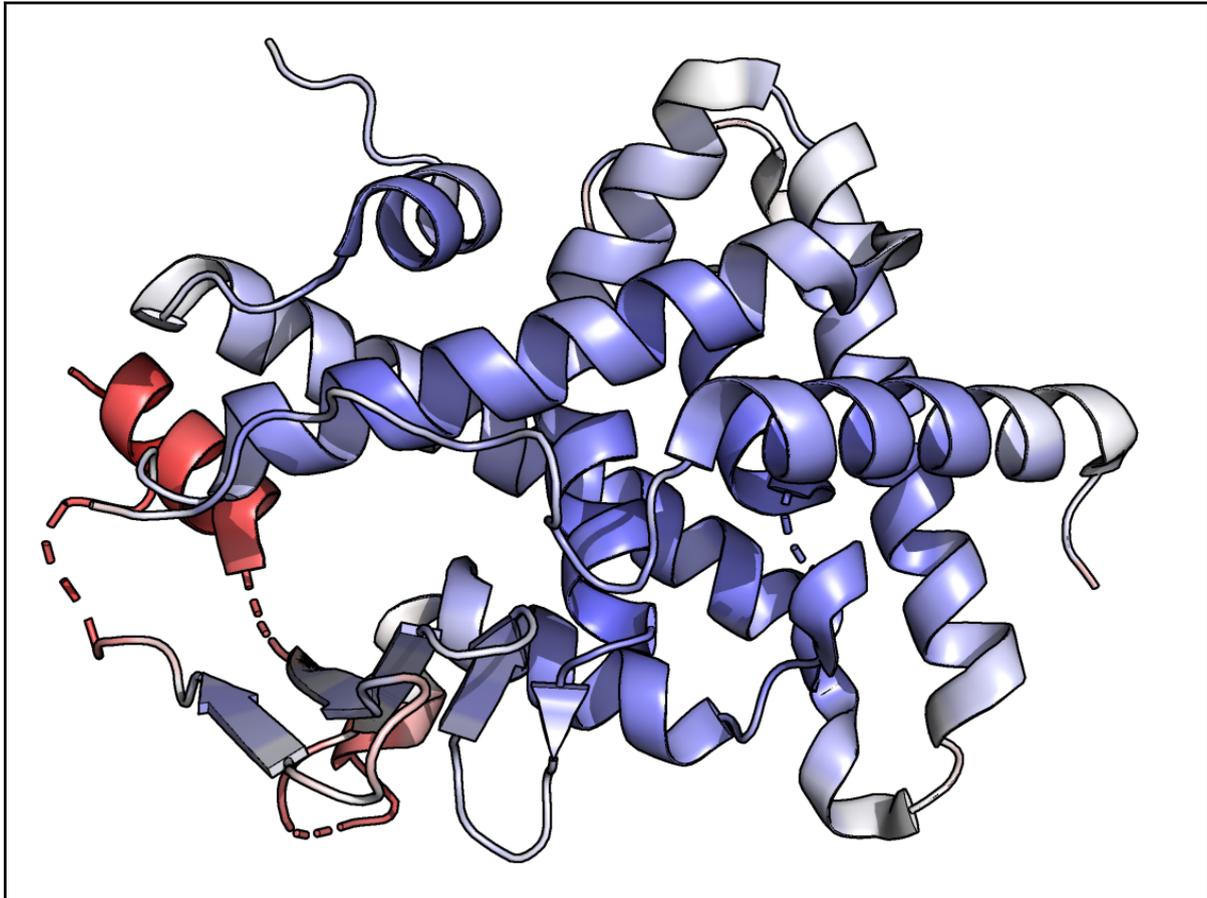


Рисунок 2. Белковая молекула, окрашенная по величине В-фактора.

Как мы видим, место, где располагается аминокислотный остаток с наибольшим средним В-фактором, также располагаются остатки с низким В-фактором (участок петли в целом плохо разрешен). Остаток с наименьшим В-фактором также лежит рядом с остатками с низким В-фактором в плотной зоне контакта α -петель.

Каково их окружение?

Такой высокое значение В-фактора для остатка Asp-230, а также для соседних с ними остатков в данном случае может быть объяснено очень высокой подвижностью петли, в состав которой он входит: подвижность ее может быть настолько высока, что это может объяснять то, что для части ее аминокислотных остатков структура вообще не разрешена (остатки 231-234).

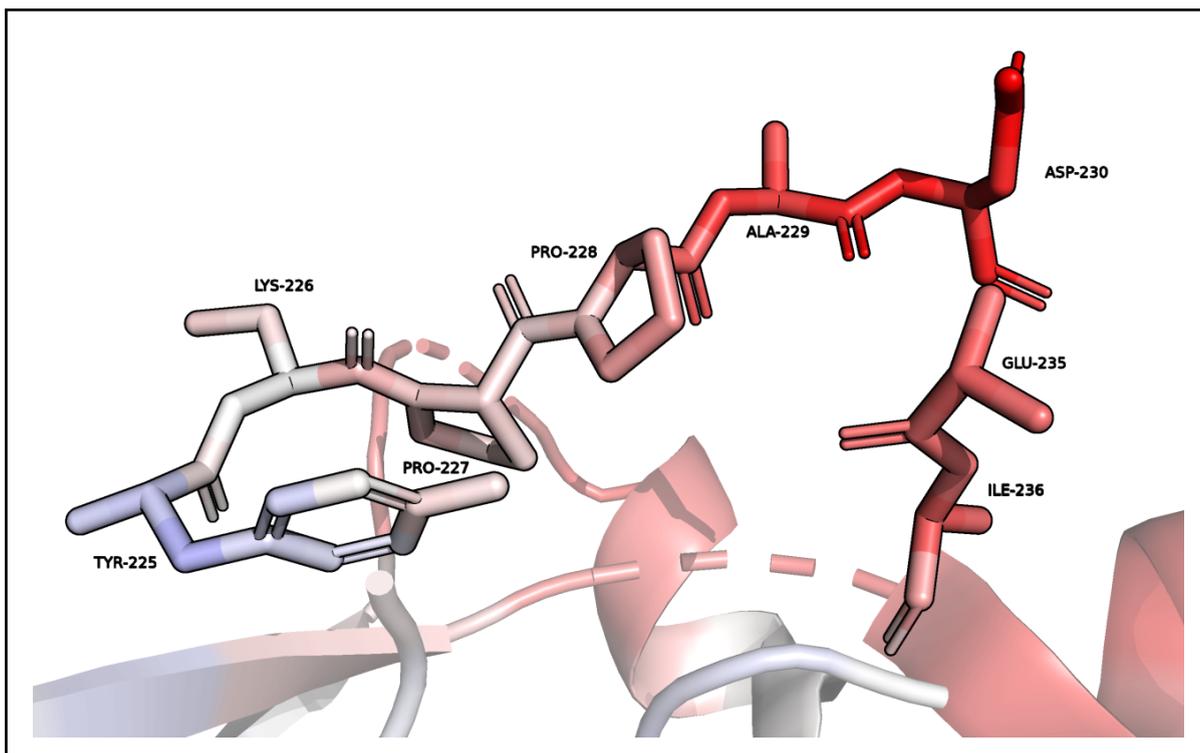


Рисунок 3. Окружение остатка Asp-230.

Для остатка Tyr-340 окружение полностью разрешено: остаток находится на поверхности взаимодействия α -петель в ярко-гидрофобном окружении.

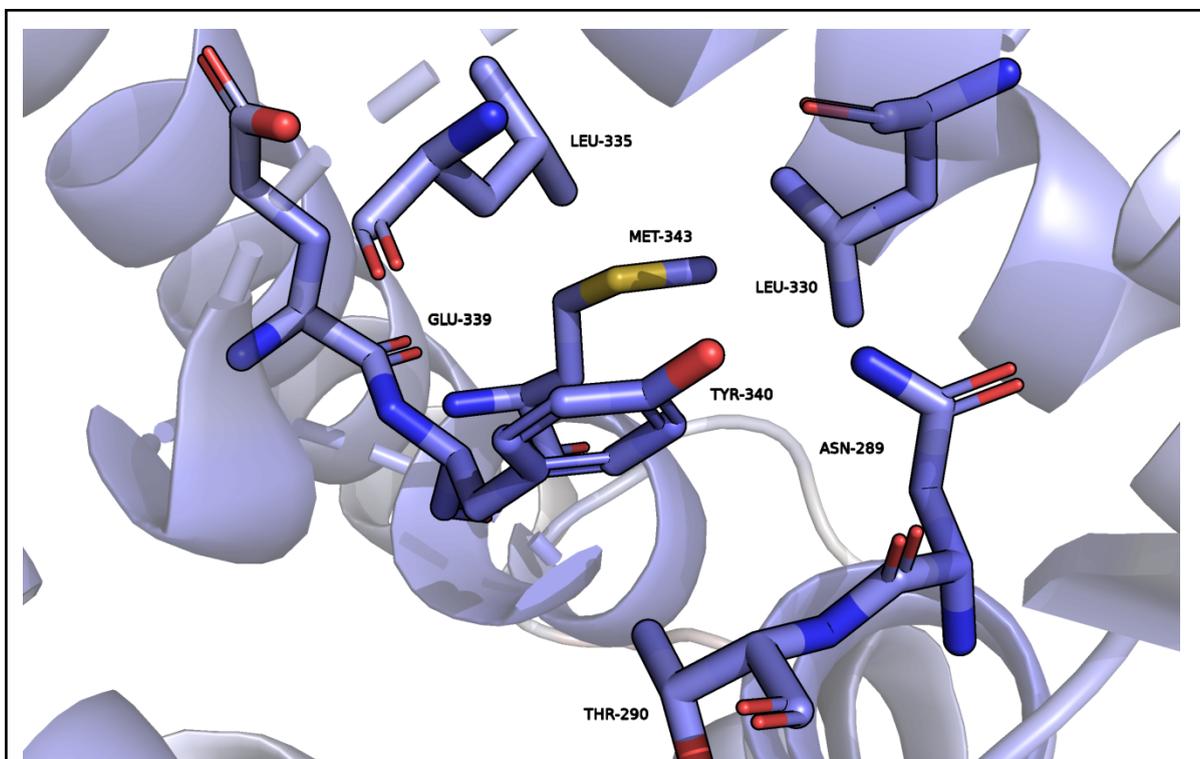


Рисунок 4. Окружение остатка Tyr-340.

Выполненное задание 2 я прилагаю в другом документе (т.к. нормально выгрузить jupyter notebook в pdf с русским языком мне не удалось).

Задание 3

В данном задании было необходимо “про моделировать” кристаллографический эксперимент на простом одномерном примере. Для этого было необходимо сгенерировать гауссову кривую, которая бы описывала расположение атомов.

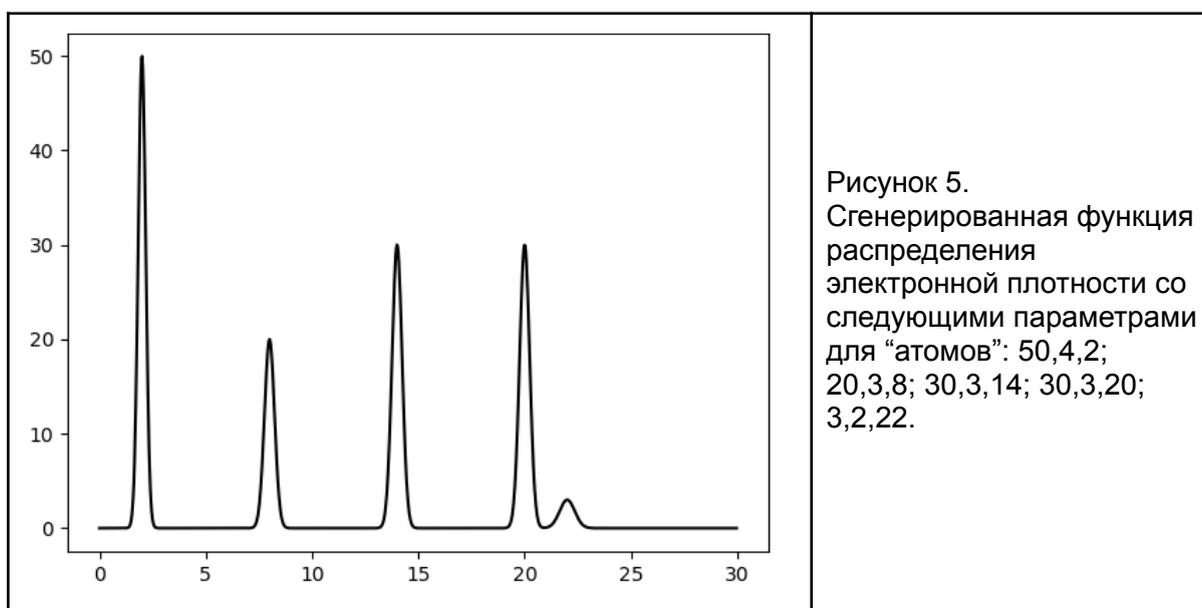


Рисунок 5.
Сгенерированная функция распределения электронной плотности со следующими параметрами для “атомов”: 50,4,2; 20,3,8; 30,3,14; 30,3,20; 3,2,22.

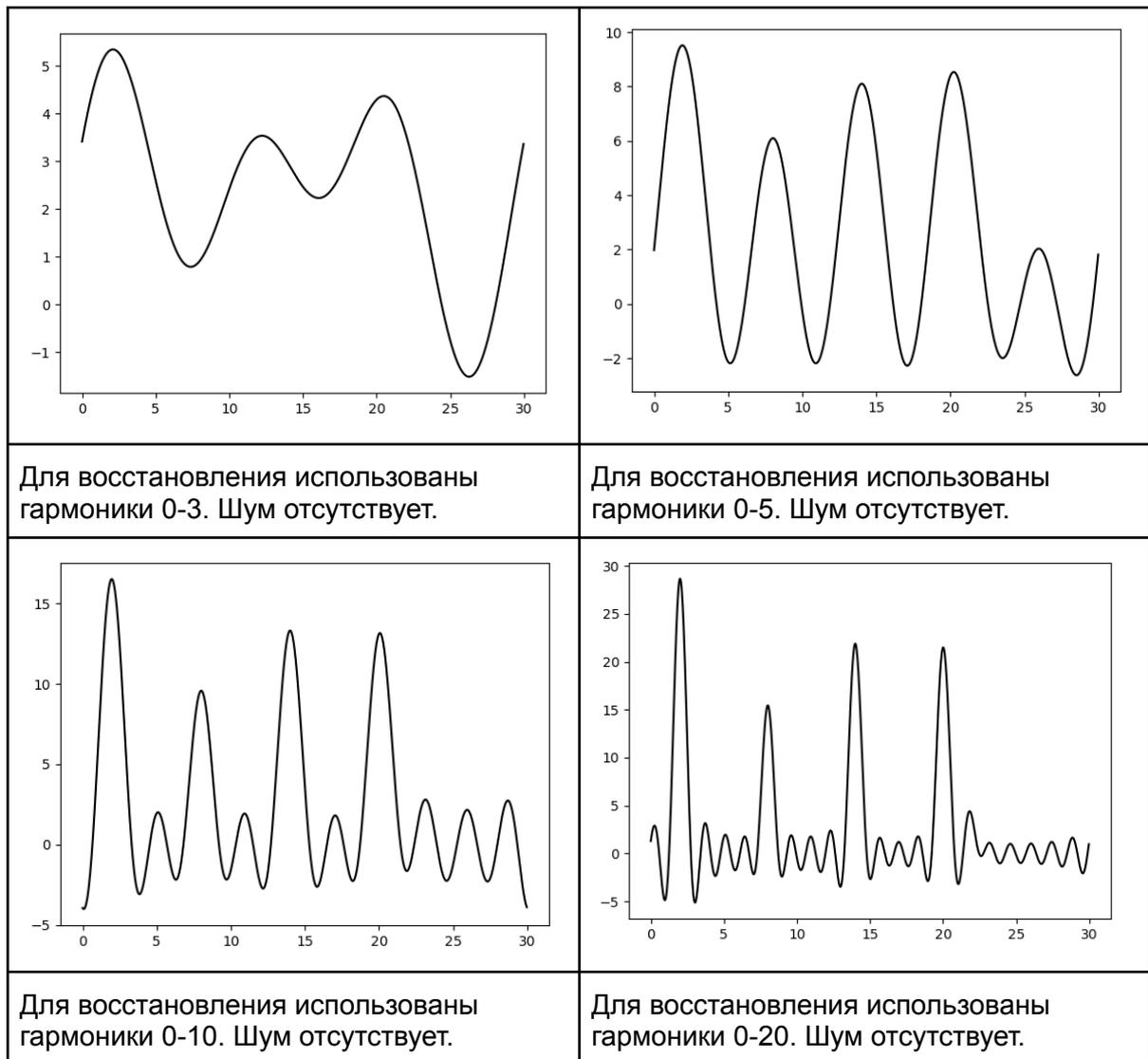
С использованием данной функции, сгенерируем модули и фазы структурных факторов, обрежем часть полученной информации и посмотрим, насколько хорошо по этой информации будет восстановлена исходная функция. Также модули и фазы были сгенерированы с присутствием шума амплитуды и фазы, а результаты всех экспериментов по восстановлению исходной функции представлены ниже.

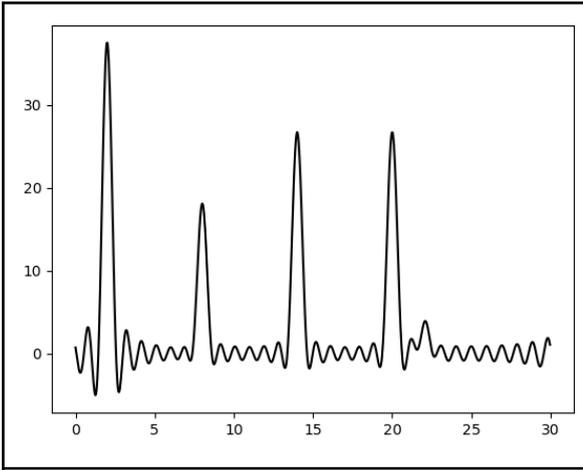
Набор гармоник	Разрешение (Å)	Полнота данных (%)	Шум амплитуды (% от величины F)	Шум фазы (% от величины phi)	Качество восстановления (отличное, хорошее, среднее, плохое)	Комментарии
Полный набор гармоник						
0-3	10	100	0	0	Плохое	Нельзя определить ни примерное положение атомов, ни их количество
0-5	6	100	0	0	Плохое/Среднее	Пики почти сливаются с шумом, но можно

						обозначить 4 пика
0-10	3	100	0	0	Среднее	Легкий атом совсем пропал
0-20	1.5	100	0	0	Хорошее	Легкий атом можно с натяжкой предсказать
0-30	1	100	0	0	Отличное	Легкий атом все еще видно, но он почти скрылся в шуме
0-50	0.6	100	0	0	Отличное	Прекрасно все видно!
Полный набор гармоник с шумом						
0-50	0.6	100	15	0	Среднее	Потерялся легкий атом
0-50	0.6	100	10	0	Хорошее	Пик угадывается, но только если знать, что его нужно искать
0-50	0.6	100	5	0	Отличное	А вот тут шум уже достаточно низкий, чтобы мы видели пик
0-50	0.6	100	0	15	Среднее	Потерялся легкий атом
0-50	0.6	100	0	10	Среднее	Потерялся легкий атом
0-50	0.6	100	0	5	Хорошее	Пик угадывается, особенно, если знать, что он там должен быть
0-50	0.6	100	10	10	Хорошее	Местоположение легкого атома можно предположить
0-30	1	100	10	10	Среднее	Легкий атом пропал и не различается
0-10	3	100	10	10	Среднее	Легкий атом пропал и не различается
Неполный набор гармоник						
3-30	1	86.6	0	0	Хорошее/Отличное	Можно догадаться о положении

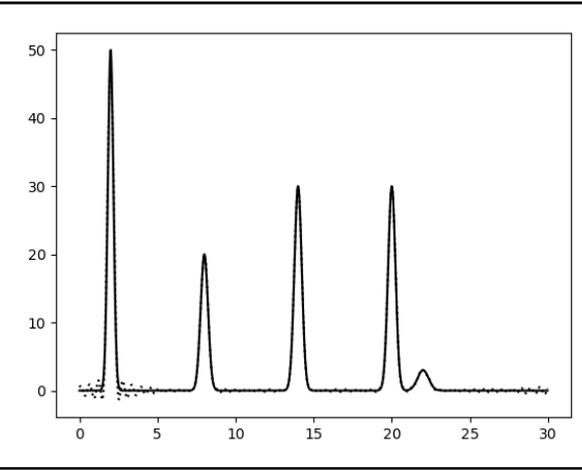
						легкого атома, но лучше знать о том, что он там есть. Baseline в виде синусоиды.
1-5, 20-40	0,75	62.5	0	0	Среднее	Легкий атом не различим при таком наборе данных
1-3, 10-15, 20	1,2	45	0	0	Среднее	Легкий атом не различим при таком наборе данных

Ниже на графиках представлены восстановленные функции.

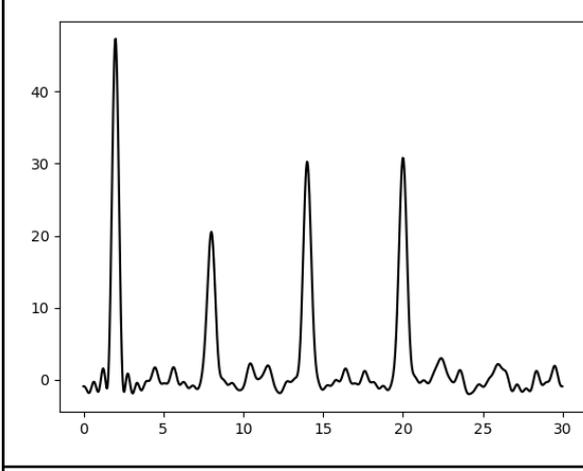




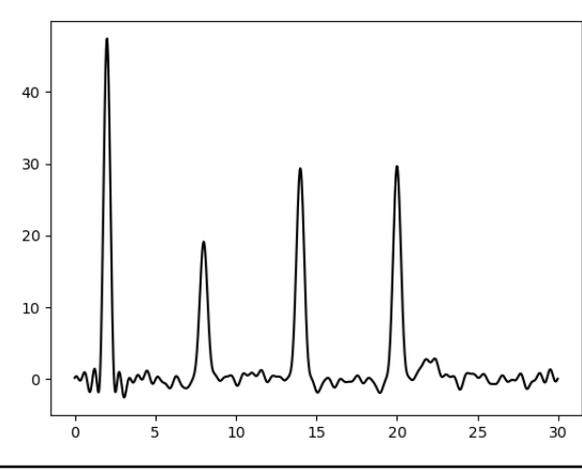
Для восстановления использованы гармоники 0-30. Шум отсутствует.



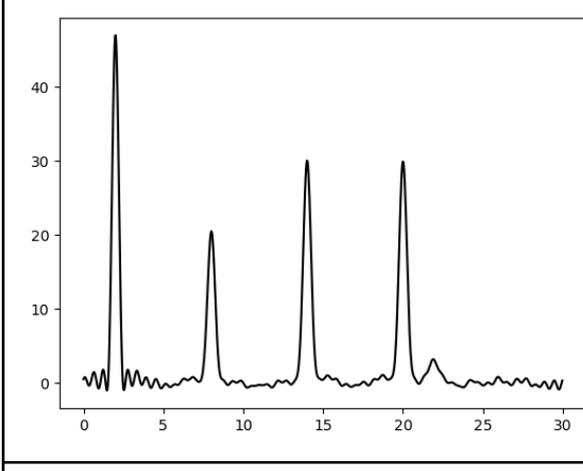
Для восстановления использованы гармоники 0-50. Шум отсутствует.



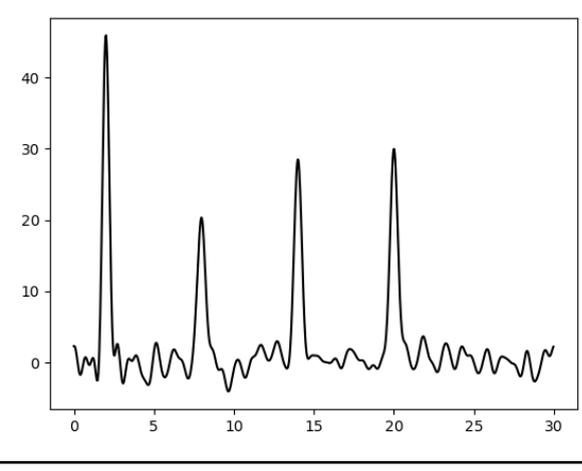
Для восстановления использованы гармоники 0-50. 15% амплитудного шума.



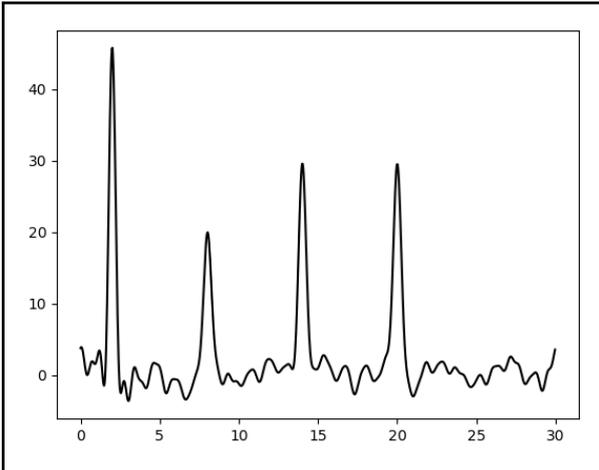
Для восстановления использованы гармоники 0-50. 10% амплитудного шума.



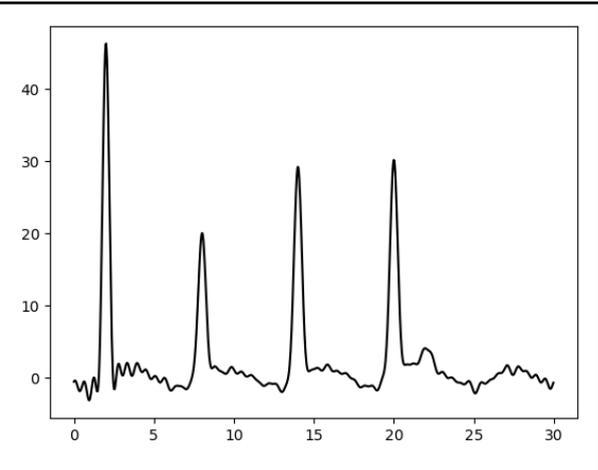
Для восстановления использованы гармоники 0-50. 5% амплитудного шума.



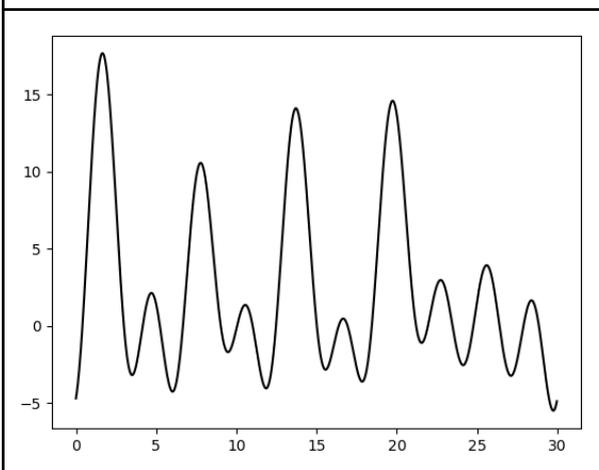
Для восстановления использованы гармоники 0-50. 15% фазового шума.



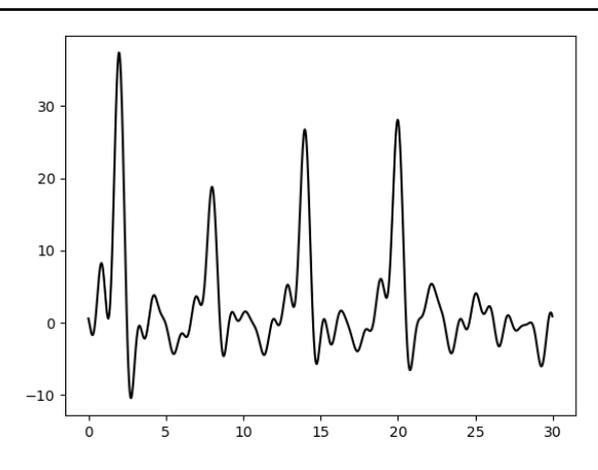
Для восстановления использованы гармоники 0-50. 10% фазового шума.



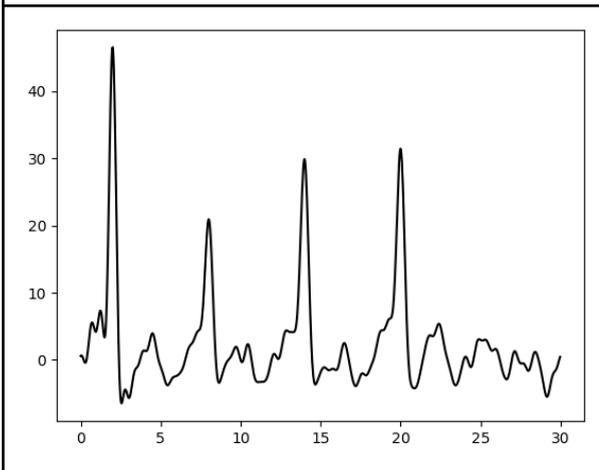
Для восстановления использованы гармоники 0-50. 5% фазового шума.



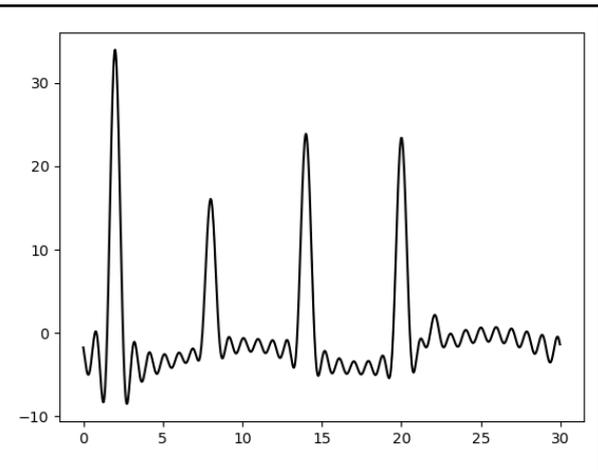
Для восстановления использованы гармоники 0-10. 10% фазового шума, 10% амплитудного шума.



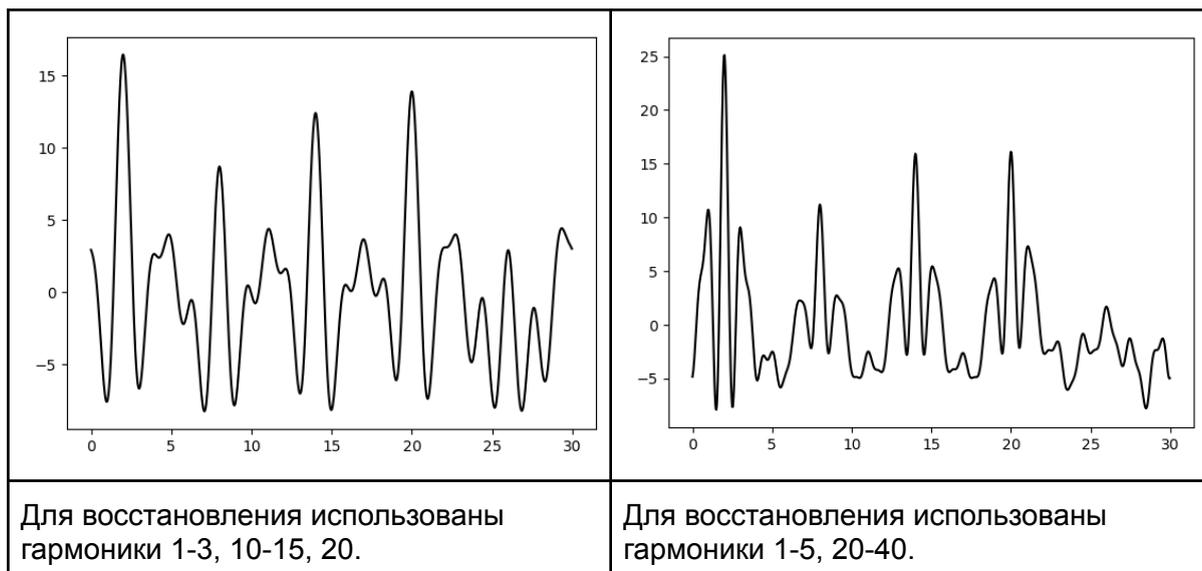
Для восстановления использованы гармоники 0-30. 10% фазового шума, 10% амплитудного шума.



Для восстановления использованы гармоники 0-50. 10% фазового шума, 10% амплитудного шума.



Для восстановления использованы гармоники 3-30.



Выводы по данной части задания:

1. РСА -- это сложно
2. При увеличении числа гармоник мы приближаемся к первоначальной функции
3. Гармоники с большим номером вносят несколько "уточняющий" вклад и позволяют определить положения более легких атомов
4. Фазовый шум является для процесса восстановления функции более "болезненным", чем амплитудный
5. Удаление из рассмотрения нескольких первых гармоник незначительно влияет на качество восстановления
6. Удаление из рассмотрения "средних" гармоник ухудшает качество восстановления и не позволяет восстановить местоположение легких атомов.