

Python introduction

H.G. Vivien¹

¹*Laboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), UMR CNRS 7328 - Université d'Orléans, Orléans, France*

July 16, 2021

Introduction

The following document contains an introduction to the Python programming language, including some lesson about the basics, examples, and exercises. While it can form a useful basis I don't recommend solely relying on it to learn the language; using books and tutorial in parallel is the way to go to learn any language. I recommend particularly the website <https://www.learnpython.org/>, which proposes many interactive tutorials and example.

The following exercises will have a correction uploaded to <https://github.com/At0micBee/Exercises>, so that you can check your results and methods to mine. I will try to keep the answer as understandable to beginners as possible, rather than writing to be as efficient as can be.

This document is more centered towards scientific processes and data analysis, and thus might not be suited for other applications. The modules covered will mostly be used for science application, and the exercises require a solid basis in both math and physics.

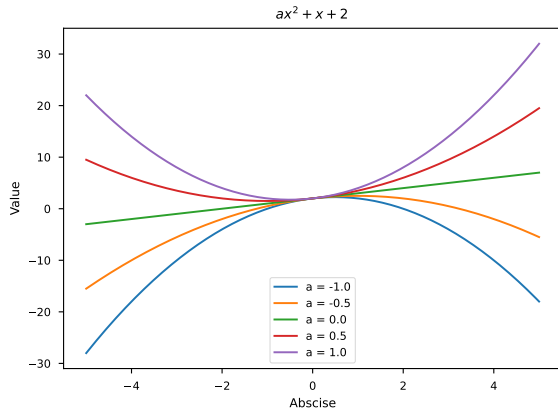


Figure 1: Example of solution to Ex. 1.

1 Polynomial coefficients

The definition of a second order polynomial is:

$$f(x) = ax^2 + bx + c \quad (1)$$

with a , b , c real numbers. We're going to look at the impact of changing these coefficients with x in the range $[-10, 10]$. Write a program that will modify each the parameters a , b or c independently. Each time we want to produce a figure that shows the resulting polynomial curve for each of the value taken by the variable parameter. Respectively, we want:

$$\begin{cases} a \in [-1, 1], b = 1, c = 2 \\ a = 1, b \in [-1, 1], c = 2 \\ a = 1, b = 2, c \in [-1, 1] \end{cases} \quad (2)$$

For example, the first case described here would produce Figure. 1

2 Polynomial roots

Create a program that computes the roots of a polynomial given its parameters a , b , c , knowing that:

$$\begin{cases} \Delta = b^2 - 4ac \\ r_{\pm} = \frac{-b \pm \sqrt{\Delta}}{2a} \end{cases} \quad (3)$$

1. Compute Δ , and figure out if the solution is real or complex
2. If it is real, solve it, otherwise print a message saying we won't compute it
3. As a bonus, add another if case to compute when $\Delta = 0$, and where $r = \frac{-b}{2a}$

3 Factorial function

Implement a function that takes in an **integer** and returns the result of the factorial of that number, knowing that the factorial function is:

$$n! = \prod_{i=1}^n i \quad (4)$$

Have it display the result for a few values. Once that is done, we want to look at the trend of this function. To achieve this, we want to compute the factorial of all integer up to 10, and plot the result.

4 Computing an average

We are given the following set of values:

```
1 v = [2.5, 3.2, 4.1, 0.8, 1.2, 2.4]
```

Copy this line to a program, and compute both the sum and the average value of this list, with the general form of the sum and average being:

$$s = \sum_{i=1}^n f(i) \quad (5)$$

$$a = \frac{1}{n} \sum_{i=1}^n f(i) \quad (6)$$

Where n is the number of values to take into account, i in the index of the values, and $f(i)$ is the value at index i .

5 Vector manipulation

Create two vectors of three elements \vec{v}_1 and \vec{v}_2 as follows:

$$\vec{v}_1 = \begin{pmatrix} 0.2 \\ 1.3 \\ 7.4 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} -7.1 \\ 2.4 \\ -0.7 \end{pmatrix} \quad (7)$$

Compute the following:

1. The addition $\vec{v}_{\text{add}} = \vec{v}_1 + \vec{v}_2$
2. Multiply them by a scalar: $2\vec{v}_1$ and $-3\vec{v}_2$
3. Combine them: $\vec{v}_{\text{tot}} = 2\vec{v}_1 - 3\vec{v}_2$
4. Compute the scalar product: $S = \vec{v}_1 \cdot \vec{v}_2$
5. Compute the rotational: $\vec{v}_{\text{rot}} = \vec{v}_1 \wedge \vec{v}_2$

Reminder about the \vec{curl} product between two vectors:

$$\vec{v}_1 \wedge \vec{v}_2 = \begin{pmatrix} v_{1y}v_{2z} - v_{2y}v_{1z} \\ v_{1z}v_{2x} - v_{2z}v_{1x} \\ v_{1x}v_{2y} - v_{2x}v_{1y} \end{pmatrix} \quad (8)$$

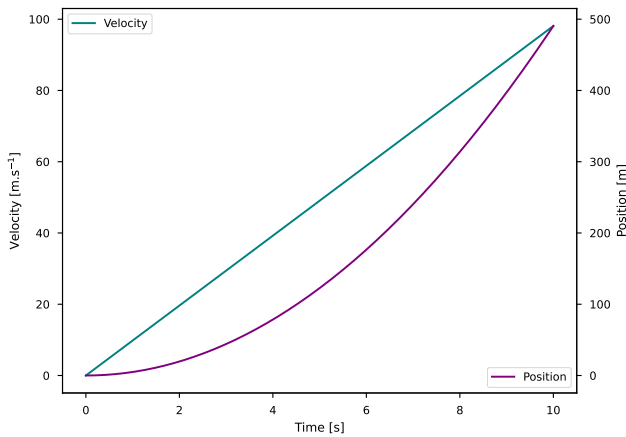


Figure 2: Plot showing the solution of Ex. 6. The velocity is read on the left and position on the right.

6 Gravity

Consider the force of gravity on a 1 dimensional axis, with $g = 9.81 \text{ m.s}^{-1}$. We want to compute the velocity and position of an object under this gravity field for 10 seconds. Assume the original position is the origin, and that it has no initial velocity.

Once that is done, we want to visualize the movement and speed on a plot to give us an idea of what is happening. Create a plot to represent both the velocity and position in function of time (see Figure. 2 for an example).

7 Surface areas

Create a function that returns the surface of a circle for a given radius r , and one that returns the surface of a square of side a . Compute the respective results for $a = r$ for a few values between 0 and 2.

Plot the results to visualize the difference of growth between the two.

8 Gravity, medium difficulty

Compute the force between the sun and an Earth mass planet for distances between 0.3 and 5AU. We have $M_{\odot} = 1.989 \cdot 10^{30} \text{ kg}$, $M_{\oplus} = 5.972 \cdot 10^{24} \text{ kg}$, $1 \text{ AU} = 150 \cdot 10^9 \text{ m}$, and with:

$$F_{\text{grav}} = G \frac{M_1 M_2}{r^2} \quad (9)$$

Where $G = 6.674 \cdot 10^{-11} \text{ m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$ is the gravitational constant, M_1 and M_2 are the masses of the objects and r is the distance separating them.

9 Studying infinite series

We are first going to look at a simple series, and then work our way towards something more compli-

cated.

$$s = \sum_{n=1}^{\infty} \frac{1}{n} \quad (10)$$

We would like to do two things:

1. Estimate if it converges, and if it does to what value
2. Visualize the impact of each term as i grows

Obviously, we cannot compute an infinity of terms, or it would literally take forever, but we can take advantage of the speed of the machine to compute a lot of them. We can start by only computing the first dozen terms, then we can push it to many more to study the series.

Once that is done, we want to try and apply it to something more useful, and so we are going to try and estimate pi. We use the sum presented in Eq. 11, and want to do the same data analysis as we did previously with the other series.

$$\pi = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \quad (11)$$

10 Chemistry reaction

We consider a chemical system with four solvents A, B, X and Y. A and B remain constant through time, but the concentration of X and Y are dependent on the current concentration of all solvents according to the following system:

$$\begin{cases} dX = A - (B+1)X + X^2Y \\ dY = BX - X^2Y \end{cases} \quad (12)$$

Where A, B, X and Y are the respective concentration of the solvent, and dX and dY the kinetic of the reaction at the given concentrations. We want to see how the concentrations of X and Y evolve through time; given $A = 1$, $X = A + \epsilon$, $Y = B/A$ and ϵ a small value. We want to be able to input the value of B from the terminal to easily change its value manually.

Compute the state of the system using the Euler scheme; which is defined as follows:

$$x_{i+1} = x_i + \frac{\partial x}{\partial t} \Delta t \quad (13)$$

Where x is the quantity we are trying to compute, i are the increment of the discretization, and $\frac{\partial x}{\partial t} \Delta t$ is the compute variation of parameter x during Δt .

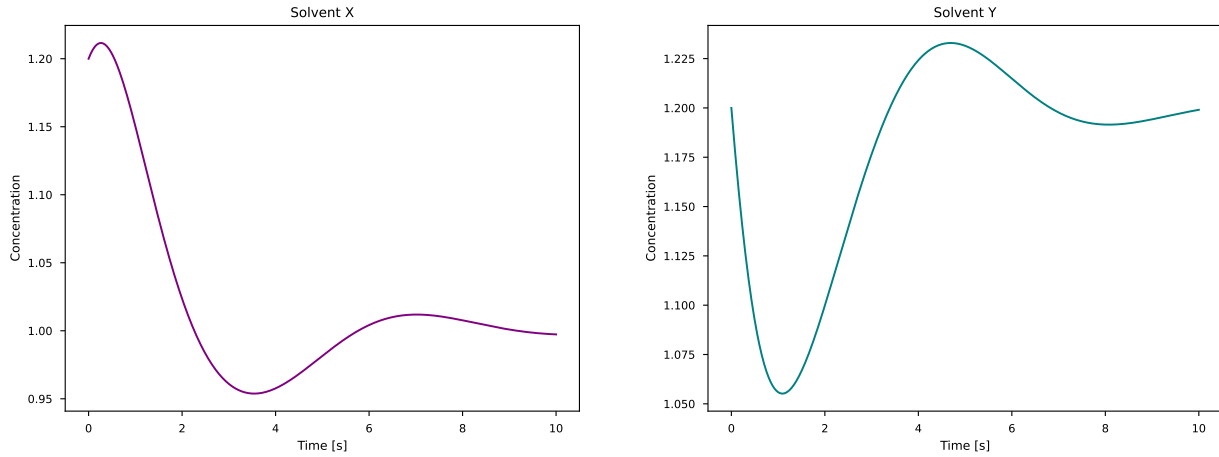


Figure 3: Evolution of the reaction through time using an Euler explicit integration scheme.

11 Gravity, but complicated this time

We're going to do a similar exercise to the Ex. 6, but more complicated this time. We're going to compute g at various altitude first, then use that value to compute the motion of a falling object. The general formula to compute gravity is:

$$g(r) = -\frac{GM(r)}{r^2} \quad (14)$$

Where r is the radius at which to estimate the gravity, $G = 6.674 \cdot 10^{-11} \text{m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$ is the gravitational constant and M is the mass under the given radius. Considering that the Earth is a perfect sphere of 6371km, that the mass under that radius is of $5.972 \cdot 10^{24} \text{kg}$ and that the mass of the atmosphere is negligible, compute the gravity for the altitude z between 0 and 100km. **Be careful, the altitude starts at the radius value of 6371km!**

We can see that even for a change of only 100km, there is a non-negligible change in gravity. Assuming an object starts to fall from an altitude of 100km with no initial velocity, and assuming there is no drag forces in the atmosphere, compute the following:

1. The velocity at each point in time during the fall
2. The position though time
3. The time it will take to reach the surface

And end the simulation when it hits the ground. To compute each point in time, we're going to use a simple explicit Euler scheme.

12 Loading data

Download the N-BK7.dat file on the Github page, load the data from this file. Plot the variation of the n , k index in relation to the wavelength.

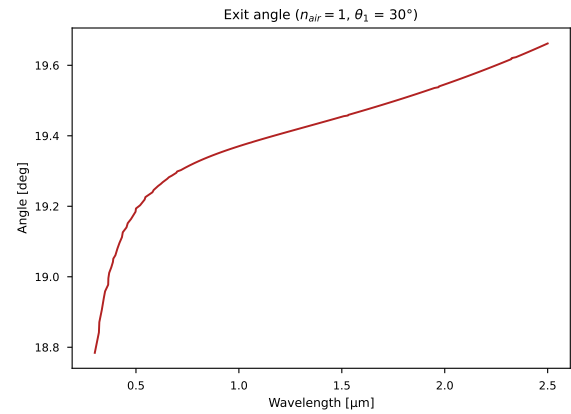


Figure 4: Illustration of the exit angle depending on the wavelength and n index of the medium.

Once that is done, we are going to see how a change in wavelength influences the medium. We are going to use the Snell-Descartes law (Eq. 15), and compute the exit angle of a beam of light interacting with the medium.

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \quad (15)$$

Where n_1 and n_2 are the indexes of the incident out outgoing medium, respectively, and θ_1 and θ_2 are the incident and exit angles. We are going to fix $n_1 = n_{\text{air}} \approx 1$ and $\theta_1 = 30^\circ$. We will use the n stored in the file to have a real value for the second medium.

Figure. 4 shows what should be obtained, we can see that the exit angle is greatly dependent on the n . This is a good illustration of the prism effect.

13 Planetary properties

We have three files from a simulation of planets with a varying amount of water on their surface

(0.01% to 5%). We want to first visualize the mass-radius relations of these planets on a plot. Then we want to study the evolution of the parameters in relations to others; namely:

- Mass & gravity
- Mass & pressure
- Mass & radius with atmosphere
- T_{eff} & T_{surf}
- Pressure & T_{surf}

Start by doing it for a single file, then implemented an automated method to do it to all files at the same time.

14 Spectral analysis

The file `all.tab` contains the emission lines from the spectrum from comet 122P/deVico. The file contains multiple columns, but we are only going to focus on the first two, which are respectively the wavelength of the line and its intensity.

First, we want to plot them together to observe the general trend. Then, we want to create a second plot representing the proper spectrum of the comet. To do that, we are going to model the rays with Gaussian functions. Assuming a width of 1, compute each emission lines and group them to create the full spectrum.

15 Working with date and time

We are going to go over the basics of the `datetime` module capabilities through the exercise.

First, we want to compute the time elapsed since your birthday. Second, we want to compute the time in six hours from now. Third, we want to compute the sum of two time deltas, to ensure that it is possible.

Finally, we want to load the file named `dates.dat`, and convert its content to proper dates using the `datetime` module. Then, using a delta, we want to check which dates, if any, are within 365 days from now. (Then test a few other time deltas to make sure it works as desired).

16 Optional: Database with Pandas

Numpy isn't the only module allowing to work with saved data, **Pandas** is another very useful and powerful tool. It is more centered toward very large databases, and has its own format to handle them called **dataframes**.

We are going to look at the Mie scattering from NB-K7 glass for multiple wavelength and particle

sizes. We want to plot the results of the scattering on a graph with the wavelength as abscissa and the intensity as ordinate. Each wavelength is in a different file, and can thus be easily plotted one by one.

The [pandas](https://pandas.pydata.org/)¹ documentation is very thorough and will help you understand the parallel between it and the way **numpy** works. The correction example given will also help show how the module works.

¹<https://pandas.pydata.org/>

A Terminology

In this section, we go over some of the terminology and notation related to Python and programming in general.

- *Data type*: Also sometimes referred to simply as *type*, it is the nature of some data in the code. For example 6 would be an **integer**, and [4.2, 6.5, 2.8] is a **list of float**.
- *Keyword*: In a programming language, a *keyword* is a word reserved for a specific use by the language. For example, **while** is a *keyword*.
- *Variable*: A *variable* is a container for a value within the program. It can store many different values of a *data type*.
- *Function*: As in mathematics, a *function* is a process that does something. Usually it acts on data within the code to either change it, or produce a result based on it.
- *Block of code*: Also referred to as *code block* or just *block* sometimes. It is used to describe a bit of code that does a specific thing on its own.
- *Declaration*: When we create a *function* or *variable*, we call that the *declaration*. It is an easy way to refer to the original instantiation of a bit of code. It is also under certain circumstances called *implementation*.

B Modules

There are a lot of modules available to do a multitude of things in Python. We are mostly focusing on math, physics and data analysis, and for that there are a few key modules that are likely to be used very often:

- **numpy**: Math and array / matrices
- **pyplot**: Data visualization
- **pandas**: Dataframe, easy data handling
- **json**: JSON format reader and writer

If a module isn't installed, it can easily be installed using the Python **pip** tool. To install any Python module, open a terminal and use the following command:

```
1 pip install <module>
```

Where **<module>** is the name of the name of the module to be installed. Once a module is installed, it can be used by the python interpreter, but it doesn't mean the interpreter *knows* that it has to use it when running a given program file. To tell it that it is the case, it is necessary to **import** the required module in the file. To do that, we use the **import** command:

```
1 import numpy as np
```

In the previous example, we **import** the **numpy** module to be used in the program. The **as** *keyword* allows us to create an alias name for a module, to simplify the notation later on.

C Conditional statements

Effectively, these are tests that the program looks at, and is able to execute a certain *block of code* in function of the result of the test. They are also called **if/else** statements. The semantics of a conditional statement is the call to the right *keyword*, then the Boolean test, then the *block of code* to execute if the result is **True**.

In the following example, we look at an arbitrary variable **a** and check to see what values it possesses.

```
1 if a < 5:
2     print("a is smaller than 5!")
3 elif a > 15:
4     print("a is larger than 15!")
5 else:
6     print("a is between 5 and 15!")
```

D Loops

Loops allow us to go over the same *block of code* and varying only a parameter at each iteration. The idea behind it is simple, we create some condition that will change iteratively, so that we don't have to write the same thing multiple times.

```
1 for i in range(0, 50, 1):
2     print("The square of", i, "is", i**2)
```

In the example above, we compute the square of the first 50 integers. If we didn't use a **for** loop, we would have needed to write manually the operation for each number. A very important thing to understand, is that the **for** loop requires an object that is *iterable*. That means that the object used must offer multiple values, so that each iteration can take place in a meaningful way.

There are other ways to create loops, we can for example use the **while** loop, which uses a condition to establish whether or not the next iteration will happen or not.

E Simple examples

E.1 Declaring a variable

A *variable* stores some data for us, so that we don't have to worry about it ourselves. It can hold multiple *types*, and can be *declared* in a number of ways. Usually, we will simply give it a name, and tell it what to store directly.

```
1 import numpy as np
2
3 # Two examples where we give a value manually
```

```

4 some_number = 3.548
5 some_text = "This is a text!"
6
7 # Giving a value through a function
8 an_array = np.linspace(0, 10, 2500)

```

E.2 Declaring a function

To *declare* a *function*, we need to use the `def` keyword. We then need to specify a name for our *function*, as well as explicitly tell whether or not it has some input values. The line is terminated with a column, and the lines belonging to the *function* process have to be indented once. At the end of the *function*, we have the option to `return` something out of this process.

For example, we can create a function that returns a polynomial based on input coefficients and point at which to estimate the function.

```

1 # Defining the polynomial function
2 def poly(a, b, c, x):
3     return a * x**2 + b * x + c
4
5 # Storing the return value to a variable
6 res = poly(2, -3, 7, 2.5)

```

In this example, the *function* `poly` computes the value of the polynomial $f(x) = 2x^2 - 3x + 7$ at $x = 2.5$. Because the function has a `return` value, we can store it to a *variable* to use it later on (in a *variable* called `res` in the example).

E.3 Append to list

To append a value at the end of a list, there is a *function* called `.append()` that let's us do just that. It can any list `ls` of length n , and create the $n + 1$ index and assign it a value.

```

1 ls = [42, 56, 84]    # List of length 3
2
3 ls.append(-10)       # Appending -10
4
5 print(ls)           # List is now length 4

```

In the previous example, we have a list of length 3 (max index = 2), and we wanted to create index 3 with a value of -10 . Using the `.append()` *function* allows us to do so very easily. Importantly, it also allows us to create a list of identical dimension to a reference when in a `for` loop. By creating an empty list, and appending a given value at each pass, we will obtain a list of identical length to our reference one.

```

1 # Creating a list of length 6, and an empty
  one
2 x = [0.0, 0.1, 0.2, 0.3, 0.4, 0.5]
3 y = []
4
5 # Going through the values of the first
6 # and appending to the second one
7 for value in x:
8     y.append(value**2)
9
10 # List y is now length 6
11 print(y)

```

In the previous example, we compute the square of some value in a list that has a fixed length (here 6). Because this list (`x`) could have unknown length, we create an empty list (`y`) to store the result at each point, and need to fill it accordingly. To do so, we go through our reference list `x`, which contains the values we want to use, and append the result to our second list `y`. Because we appended as many times as there are values in the `x` list, the `y` now has the same length as `x`.

E.4 Plotting some data

The following is a simple example showing how to use `pyplot` and `numpy` to visualize some data. We are going to plot the sin function for $x \in [-\pi, \pi]$.

```

1 # Calling the required modules
2 import numpy as np
3 import matplotlib.pyplot as plt
4
5 # Creating 1000 values for x in the range
6 # And taking their sinus values
7 x = np.linspace(-np.pi, np.pi, 1000)
8 y = np.sin(x)
9
10 # Creating a figure
11 plt.figure()
12
13 # Plotting our data
14 plt.plot(x, y)
15
16 # Useful functions for clarity of plot
17 plt.title("Sinus function")
18 plt.xlabel("Value of x")
19 plt.ylabel("Corresponding sinus value")
20
21 # Saving the figure
22 plt.savefig("sinus.pdf")
23
24 # Closing the figure
25 plt.close()

```