Scientific Python

A brief introduction

Trygve Eftestøl

August 13, 2018

Contents

1 Introduction **1**

1.1 Motivation for the course through an example 1

1.2 Installing Python 3

2 The Python user interface **3**

2.1 The Interpreter 5

2.2 The editor 6

2.3 Help 7

2.4 Exercises 7

3 Basic functions **8**

3.1 Creating matrices 8

3.2 Matrix operations 10

3.3 Matrix functions 11

3.4 Indexing matrices 12

3.5 Logical operators 14

3.6 Exercises 15

4 Visualisation **16**

4.1 Simple 2D plotting 16

4.2 3D plotting 18

4.3 Exercises 20

5 Programming **21**

5.1 Function files 21

5.2 Control structures 22

5.2.1 Conditional testing with *if* 22
5.2.2 Iterative structures using *while* or *for* 23
5.3 Exercises 24

6 Analysis **24**
6.1 The problem - computing the heart rate 25
6.2 Solving the problem 25
6.3 Exercises 29

These notes provides an introduction to using Python as a computational and programmingtool. It focus on typical areas of usage: mathematical computations, visualisation, data analysis and programming. Reading these notes, you will get to know Python's user interface, try out vector amd matrix manipulation, use some of the central computational functions, get impressed by Python's capabilities for and reach a level of expertise where you will be able to make your own small Python programs.

1 Introduction

Python is an interactive matrix based system for numeric computations and visualisation. The Python name refers to the *Monty Python's Flying Circus* which was a legendary British sketch comedy series airing in the period 1969 to 1974.

Python provides you with the possibility for computations, visualisation and analysis far exceeding the scope of these notes. Therefore you should get familiar with the use of the help and documentation system integrated in Python. By doing this you will be able to find the proper functionality when you need it.

1.1 Motivation for the course through an example

The course will soon dive into the world of number sequences: how to create sequences and tables of numbers and do mathematical and logic operations on the numbers as well as picking out parts of these. Therefore a real world problem will be decribed. Upon reaching the final section of the introduction, teh reader should be capable of solving this problem.

Determining heart rate from pulse beat measurements

For motivation, a realistic problem will be handled at the end of the course where the aim is to determine the heart rate. This is a technology resembling what can be found in modern wearable sports watches. During a cardiac cycle, the pressure pulse reaches the skin causing a change in volume. This change can be detected by illuminating the skin with light from a light-emitting diode (LED) and measuring the amount of light reflected or transmitted to a photodiode.

Figure 1 shows a text file where a photopletysmograpchic recording has been

stored after digitizing the measurements. The view is split in two, showing the lines at the start and end of the file. Let us consider an approach to determine

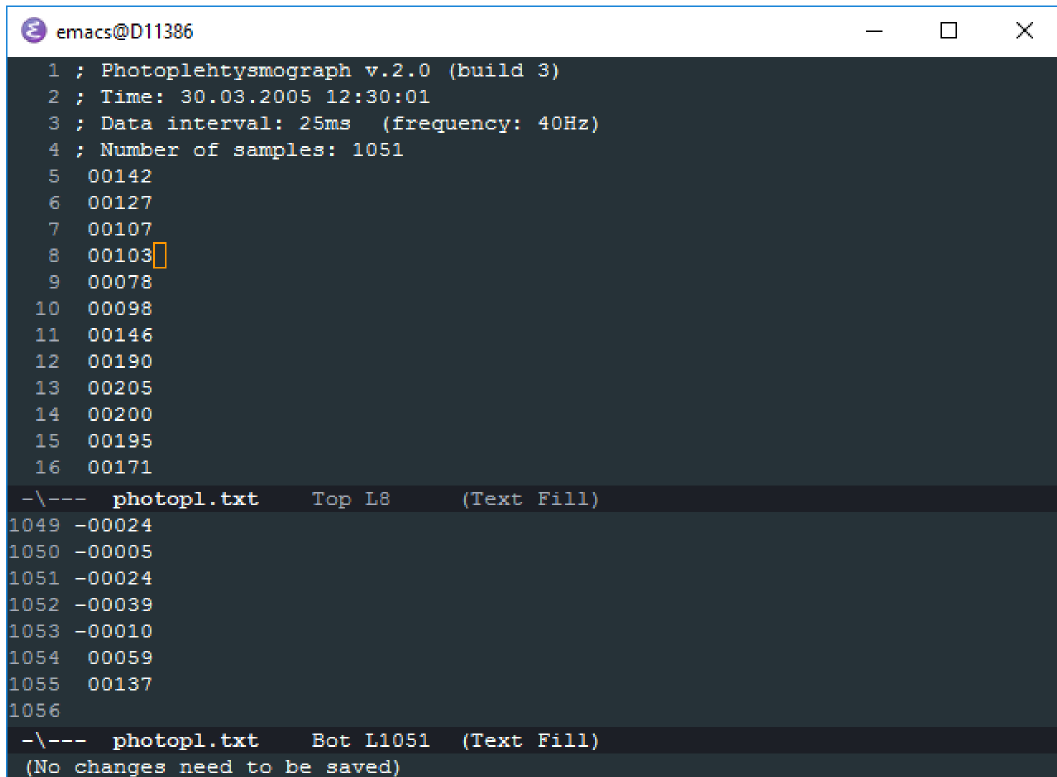

Figure 1: A textfile with photopletysmographic measurements.

the heart rate. For each pulsewave, the measurement values should go through a cycle where the values increase from some minimum value to a maximum and returning to a minimum value. If we can determine the timepoints for the maximum value, we can use this to represent the time for the corresponding heartbeats. If we find all the maximum points representing one heartbeat each throughout the recording and keep record of the corresponding time points we can calculate the cardiac cycle time intervals. We can use these to calculate the average time interval and thus the average heart rate.

Looking at the four first lines, the *header*, we see that the sampling frequency is 40 samples per second. This information is essential for us to be able to determine the time points in seconds. To find the maxima, we need to find the lines where the numbers starts to decreas after increasing. There is one example on line number 13 where the maximum value 205 is reached increasing from 190 before decreasing to 200. A similar maxima is found on line 1050. Now consider the burdensome task of finding all the maxima values among the 1051 values. As the recording corresponds to a total duration of 26.275 seconds one should expect to have to find approximately 30-40 such points. Unfortunately, the measurements contains many maxima that are not associated to

heartbeats. These false detections are due to measurement noise. They will typically have lower values that the heartbeat maxima. One could consider censoring these detections, but the task of handling this manually now seems to have grown unmanageable.

Section 6 will show how the numbers representing the measurements can be read into the Python workspace and further how the signal itself can be visualised, and the signal peaks detected. We will also introduce some signal processing concepts that will provide tools to handle the false detections. Before we can do this we need to introduce the basics on how to handle number arrays, visualisation and some basic programming concepts. This will be covered in sections 3, 4, and 5. But first, we will introduce the Python working environment in section 2.

1.2 Installing Python

The introduction requires that Python is installed on your computer. See https://www.python.org/ for instructions on how to install Python on your computer. If you are using Windows *pipwin* is recommended for installation of precompiled packages (https://pypi.python.org/pypi/pipwin/).

It is recommended to install scentific python by giving the command *pip install scipy* which can be replaced by *pipwin install scipy* on Windows.

2 The Python user interface

The procedure for starting Python varies according to your operating system. For *Windows* and *MAC* users the Python application will usually be available from the menu system, represented as an icon. Start Python by double clicking this icon. *UNIX* and *LINUX* users can start Python double clicking an icon or from a terminal window.

When working with Python, there are two important environments: The *editor* where programs can be written and the *interpreter* where commands can be given in interactive mode. Figure 2 shows an example where the text editor *Emacs* is used as a Python Integrated Development Environment (IDE). In the top window there is an editor for for writing code in a .py file. In the top window, you se an example of a Python interpreter, in this case *Ipython*.

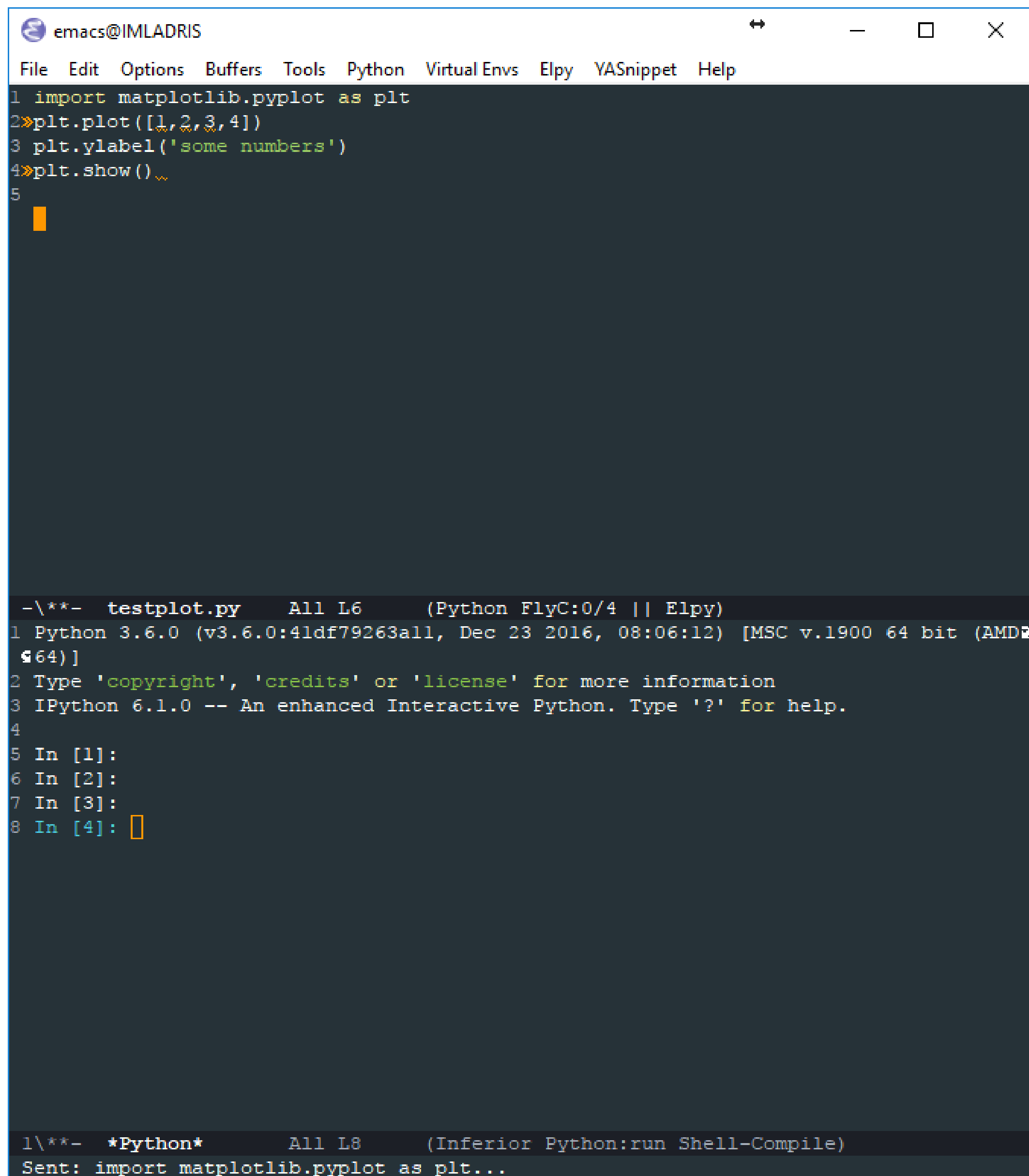

Figure 2: Emacs Python IDE: Editor (top), Interpreter (bottom).

2.1 The Interpreter

In the interpreter you can declare variables, perform simple operations and call functions. In figure 3 you see an interpreter window where the standard *CPython* which is an alternative to the Ipython interpreter which is used in notebbok solutions which offers solutions for rich media.

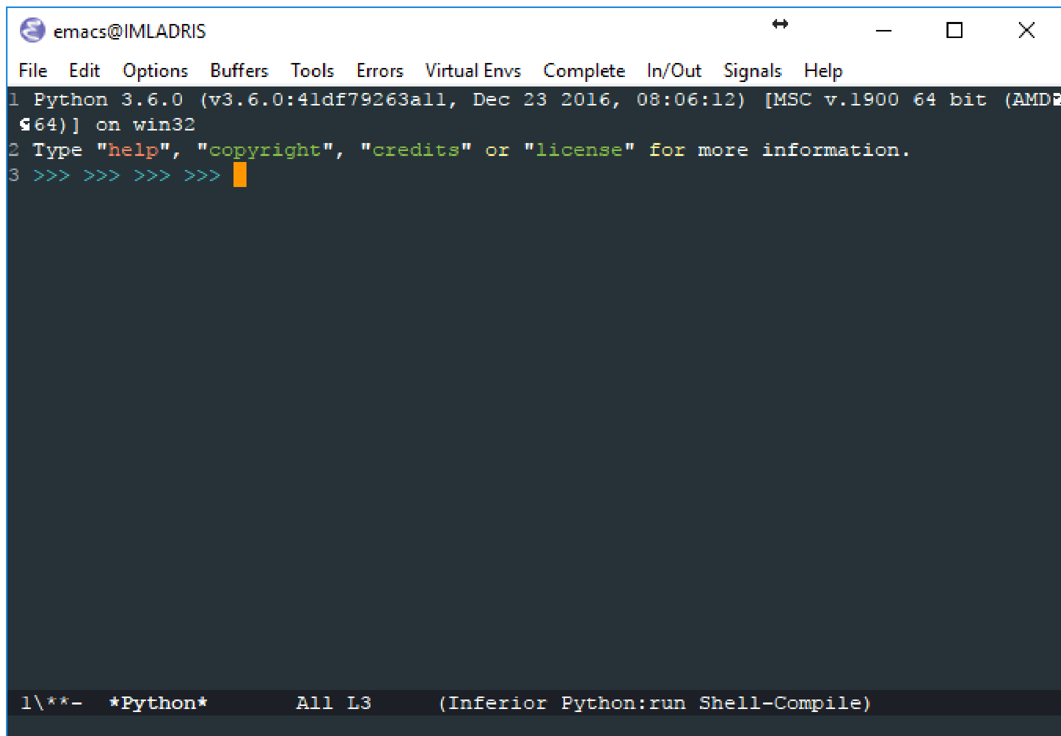

Figure 3: Python CPython interpreter.

Commands are given at the prompt, If you want to execute a command called *function*, this is usually done from a terminal window by giving command (executes at ENTER):
`C:\python function.py`
A simple example;

```
>>> 10
```

where Python will respond

```
10
```

Another important concept is the usage of *variables*. A variable can be created directly in the command window e.g. when we want to create the variable *a*

and store the value 10 in it. This is done quite simply by giving the command (executes at ENTER)

```
>>> a=10
```

Variables are often used to declare input and output parameters at function call. As an example,

```
>>> import numpy as np
>>> b=np.sin(a)
```

means that the command *sin* is called with the variable a as input parameter and the result from the operations executed is sent back through an output parameter and stored in the variable b. After the execution of these commands, the variables a og b can be displayed by calling them as shown below.

```
>>> a
10
```

```
>>> b
-0.54402111088936977
```

Note that the the *sine* function is part of the NumPy package. That is the reason for the `import numpy as np` command. The functions `foo` in this package can later be called by giving the command `np.foo()`

2.2 The editor

In the editor you can make your own Python-functions by writing your programs in *py-files*. The editor window is shown in the upper window in figure 2. The editor can be used for debugging of the programs you are writing. The two commands described above can be put into an py-file:

```
import numpy as np
a=10
b=np.sin(a)
```

If the py-file is named *instructions.py*, the two commands can be run by giving the command

```
C:\python instructions.py
```

2.3 Help

Figure 4 shows the *Documentation* page for Python 3.x which can be found at python.org. Here are links to tutorials and reference material. There is also a search page which can be used to search for and display documentation and demonstrations for Python functionality. Python has a large user community,

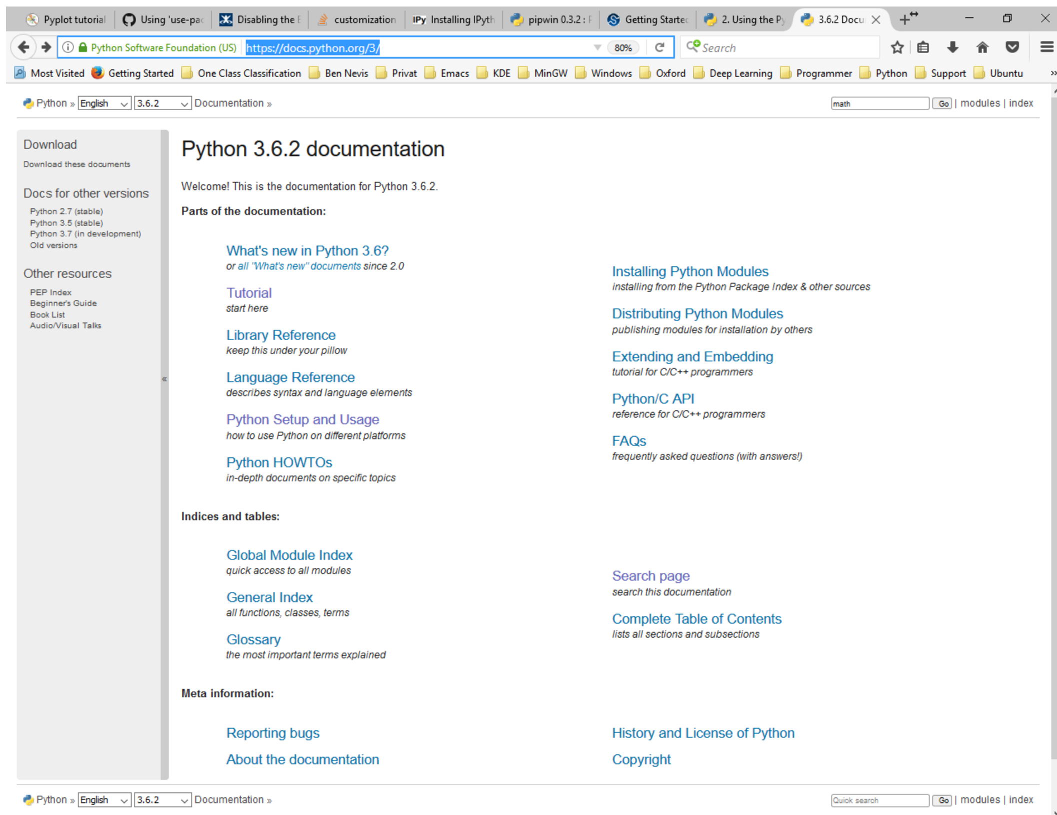

Figure 4: Python documentation.

so searching the web is also an option.

2.4 Exercises

a) Install Python.

b) Start a Python Interpreter.

c) Open an editor.

d) Find the *Tutorial* on the documentation page.

e) Use the search page find information on how the *sine* function is used.

3 Basic functions

Now it is time to get started with the main issue about Python - computations. It will be useful to keep in mind that Python, as the name indicates, is *matrix* based. A matrix is a set of elements organised in a rectangular array. These matrices can represent various types of data:

- Scalars, vectors and matrices which all are diferent types of matrices, but containing one single element, a single row of elements or several rows of elements, respectively.
- Tables which are organised as a list of elements. Python will not distinguish between a matrix and a table, but gives the user the opportunity to handle data as one or the other.

The concept of *elements* might seem a bit vague, but Python makes distinction between elements of different types:

- Numerical numbers which can take part in computations,
- Characters assembled in a vector denoted a *string*,
- Symbols,
- One might even define a matrix as an element.

The first three types of elements can be handled using matrices, while for the last one where the element definition may vary special *cell* structures are applicable. We will return to the various data stuctures of Python.

As the concepts of vectors and matrices might be vague to some, we give some hands-on examples:

- A *sound file* will be a vector. When the sound file is imported into Python, the sound samples will be represented as elements in a vector.
- A *grey tone image* will be a matrix. When the image file is imported into Python, the picture elements (pixels) will be represented as elements in a matrix.

3.1 Creating matrices

Matrices can be assigned in the following ways:

- The elements are explicitly put into a list,
- Generated using built-in functions,
- Created in an py-files called from the command window,
- Read from external data files or applications.

Here are some examples on how we can create matrices directly in the command window.

First we need to import MumPy which is a package with functions for scientific computtations and contains functions for handling vectors and matrices.

```
>>> import numpy as np
```

Note the *as np* part of the import statement. The functions in the package can thus be called by prepending *np.* as in `np.foo`.

Assigning scalars:

```
>>> a=5
>>> a=
5
```

Assigning a row vector:

```
>>> y=np.array([0, 5, 10, 15, 20])
>>> y
array([ 0,  5, 10, 15, 20])
```

The elements can be separated by a comma instead of a space. The result will be the same.

A column vector can be assigned correspondingly

```
>>> x=np.array([0, 5, 10, 15, 20])
>>> x=x.reshape(x.size,1)
>>> x
array([[ 0],
       [ 5],
       [10],
       [15],
       [20]])
```

A matrix can be assigned as follows

```
>>> A=np.array([[0,1,2,3,4],[5,6,7,8,9],[10,11,12,13,14],[15,16,17,18,19]])
>>> A
array([[ 0,  1,  2,  3,  4],
       [ 5,  6,  7,  8,  9],
       [10, 11, 12, 13, 14],
       [15, 16, 17, 18, 19]])
```

We can assign the same matrix from an M-file, *Amatrix.py*, where the following instructions are given

```
import numpy as np
A=np.array([[0,1,2,3,4],[5,6,7,8,9],[10,11,12,13,14],[15,16,17,18,19]])
```

A is then assigned and shown by

```
>>> run Amatrix
>>> A
array([[ 0,  1,  2,  3,  4],
       [ 5,  6,  7,  8,  9],
       [10, 11, 12, 13, 14],
       [15, 16, 17, 18, 19]])
```

3.2 Matrix operations

Some of the most fundamental mathematical matrix operators are addition, subtraction, multiplication og transponation. Another useful “house hold”-operator is the colon operator.

+	addition
-	subtraction
*	multiplication
/	division

```
>>> y+y
array([ 0, 10, 20, 30, 40])
```

`x+x` will give an equivalent results, but in column vector format.

for vectors, the operators will work element wise so that the resulting vector has the same dimension as the operand vectors, and the value in a specific position is the result of the operations performed on the elements in the corresponding positions in the operand matrices.

Operations between a scalar and a vector/matrix will always be element wise.

It is important to note the difference between *matrix* operations and *element* operations.

3.3 Matrix functions

Examples of matrix operations can be element-wise or operations well known from linear algebra.

Some basic mathematical matrix functions allow us to sum, diagonalise, find determinants etc.

power	power
transpose	transponation
sum	sums the elements in a vector/matrix matrix
dot	vector inner product, vector/matrix multiplication
diag	diagonalises a matrix
det	computes the determinant of a matrix

Note that all funcctions are either called `np.sum(A)` or `A.sum()`.

```
>>> b=np.dot(A,x)
>>> b
array([[150],
       [400],
       [650],
       [900]])
```

There are also some functions useful for “building” matrices.

eye	generates an identity matrix
ones	generates a matrix of ones
rand	generates a matrix of random numbers

size	determines the dimension of a matrix
arange	create an increasing vector
:	slicing

Note that the arguments to these vary eye(nr), ones((nr,nc)), matlib.rand(nr,nc), sise(foo,d), where nr and nc denotes the number of rows and columns. d denotes the dimension for which nr (d=0) or nc (d=1) can be determined.

As an example, a $2X2$-matrix can be generated by

```
>>> I=np.eye(2)
>>> I
array([[ 1.,  0.],
       [ 0.,  1.]])
```

The arange operator is used to generate number sequences.

```
>>> n=np.arange(1,6)
>>>n =
array([1, 2, 3, 4, 5])
```

As we can see, the operator fills in all the integers between 1 and 5. The default increment value is 1, but we can decide its size by giving its size. For example we can set the increment value in the previous example to 2.

```
>>> n=np.arange(1,6)
>>>n =
array([1, 2, 3, 4, 5])
```

As an example, some of the vectors and matrices we generated previously can be described more efficiently

```
>>> y=np.arange(0,21,5)
>>>y =
array([0, 5, 10, 15, 20])
```

3.4 Indexing matrices

By indexing one might work with excerpts of matrices. The colon operator is important in this context. The element in row i, column j in a matrix,

A, can be retrieved by giving the command `A[i-1,j-1]`. The first element indicates the row while the second element gives the column. Note that the first elements in rows and columns are indexed with 0. As an example

```
>>> A[1,2]
>>> A
7
```

retrives the element in row number 2, column number 3 in the matrix A.

We do not need to limit our attention to single elements, but can use the colon operator to indicate which part of the matrix we want to access. First we determine the number of rows and columns

```
>>> nr=np.size(A,0)
>>> nc=np.size(A,1)
```

```
>>> A[1,0:nc]
array([ 5,  6,  7,  8,  9])
```

retrives the second row from A. This command can be given an even more convenient short form

```
>>> A[1,:]
array([ 5,  6,  7,  8,  9])
```

Also, reference to last element to be indexed can be referenced by number of elements to last element.

```
>>> A[1,:-1]
array([ 5,  6,  7,  8])
```

Finally, it is useful to be able to expand a matrix by connecting it to another matrix.

```
>>> np.hstack((a,y))
array([ 5, 0,  5, 10, 15, 20])
```

3.5 Logical operators

Python has several logical operators which are applied on matrices element wise

<	less than
>	larger than
<=	less than or equal to
>=	less than or equal to
==	equal
!=	not equal

Note that `==` is used as a logical operator while `=` is used for assignment.

In addidion we have the well known logical operators

and	logical AND
or	logical OR
not	logical NOT

We can determine which elements in y are larger than or equal to 10 by giving the command

```
>>> y>=10
array([False, False,  True,  True,  True], dtype=bool)
```

which gives a boolean result indicating for each element position the test being FALSE or TRUE respectively.

Furthermore, we can use the function *where* to identify the elements in the vector fulfilling the test:

```
>>> idx=np.where(y>=10)
>>> idx
(array([2, 3, 4], dtype=int64),)
```

idx indexes the positions of the elements fulfilling the test.

3.6 Exercises

Some of the operations that you will performwill return an error message. Why? a) Generate the variables *a,y,x* and *A* corresponding to how it was shown earlier in this section. Use the colon operator to generate the vektor x and the matrix A efficiently.[1]

b) Execute the operations `a+a, a+y, a*y, a*A, x*A, x*x, A*A`.

c) Execute the operation `det(A), diag(A), diag(diag(A))`.

d) Establish a 3X3 identity matrix.

e) Generate a vector corresponding to $\mathbf{x} = (1\ 2\ \ldots 100)$.

f) Fetch every second element from column number 2 in A.

g) Fetch 6 elements from the middle of A.

h) Expand A by appending y as an extra row.

i) Write the elements of y larger than or equal to 15 to the command window.

j) Crudely put, genes are the "cook book" of an organism. A DNA-string contains genes, areas that regulate the expression of genes, and areas that do not have any function that we know of.

DNA is coded with four exchangeable "building blocks", called *bases*, abbreviated to A, T, C eller G, corresponding to their chemical names: Adenine, Tymine, Guanine og Cytosine. The variable *nucseq*[2] represents a DNA-string. Upon analysis of DNA-strings it is convenient to remove the areas without function. To do this we can take advantage of the periodicity of 3 of the areas of interest. These areas can thus be identified using frequency analysis. To do this, the sequence has to be reformated to a analysable sequence. One way to do this is to make a binary sequence for each of the bases. For instance, the binary sequence for 'a' for the DNA string 'atcgacgta' will be '100010001' while the binary sequence for 'c' will be '001001000'. Make a binary sequence for each of the four bases for the DNA string in nucseq.

[1] Hint: Use transponation for efficient generation of x.

[2] *nucseq.mat* is a MATLAB .mat file which can be downloaded from www.ux.uis.no/˜trygve-e. You need to find out how a .mat file can be read into Python.

4 Visualisation

Python has an extensive library of routines for visualisation of vectors and matrices. We will concentrate on a small selection of the most important functions for generating 2D and 3D graphics. In addidion, there are good opportunity for putting text to and printing graphics.

4.1 Simple 2D plotting

In this context we want to plot function values to their corresponding argument values. Consider the function $y = f(x)$. We would like to plot values of y against corresponding values of x.

To do this one might generate a vector, $\mathbf{x}$ with argument values, or *points of computation*, for the function. The next step is to perform the computation corresponding to $y = f(x)$ for each single point of computation in $\mathbf{x}$.

As a simple introductory example of plotting we consider the function $y = f(x)$ der $f(x) = x^2$. We want to plot the function in the interval $x \in [-2, 2]$.

The first thing we need to do is to generate points of computation in the interval $x \in [-2, 2]$.

```
>>> x=np.arange(-2,2,0.01)
```

Notice how we employ an increment step of 0.01 so that we get 400 points of computation in total from -1.99 to 1.99. Furthermore, we compute the functional values for each point of computation:

```
>>> y=np.square(x)
```

Here we make use of the power operator in an element wise manner so that the operation is executed on each single element in $\mathbf{x}$.

Finally we generate the plot:

```
>>> plt.plot(x,y)
>>> plt.show()
```

Python responds by generating a figure window as shown in figure 5.

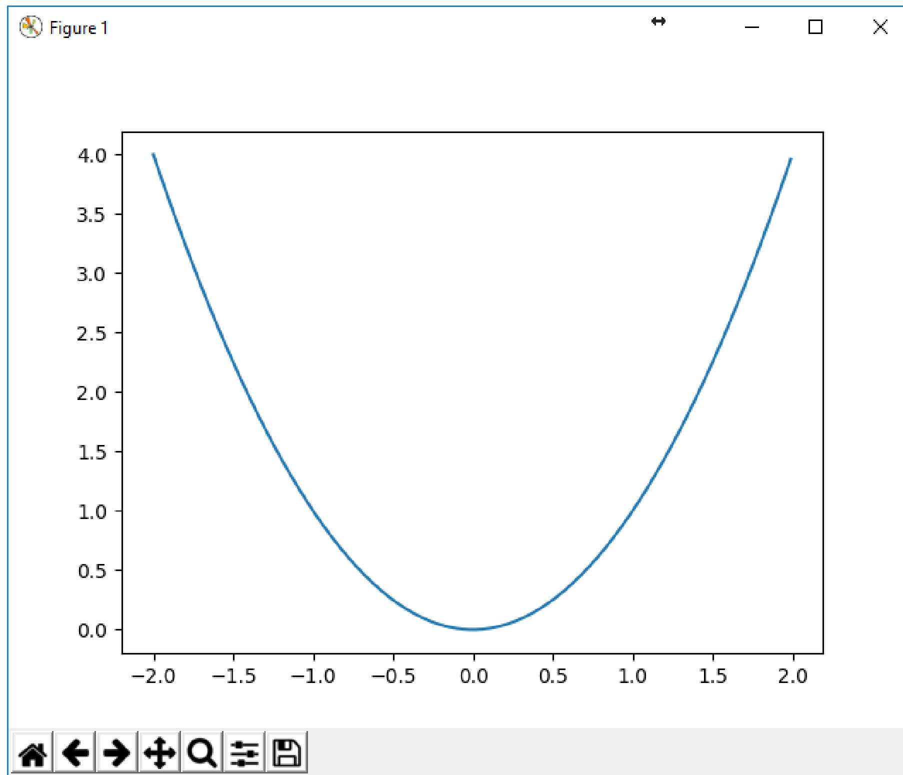

Figure 5: Figure window with plot of $y = x^2$

It is also possible to plot several functions simultaneously, define line types and colour, put text on the axes, make titles and print the graphics, just to mention some of the most useful functions.

```
>>> y2=2*abs(np.sin(x))
>>> plt.plot(x,y,'b',x,y2,'r:')
>>> plt.xlabel('x')
>>> plt.ylabel('y')
>>> plt.title('Demonstration of simple 2D plotting');
>>> plt.savefig('plot2D.png')
>>> plt.show()
```

Note that ' can not be copied to the command window. The results can be seen in figure 6. You should be able to find out the effect of the various commands by studying figure 6 which actually is a png-file, *plot2D.png*, generated by the *savefig* command.

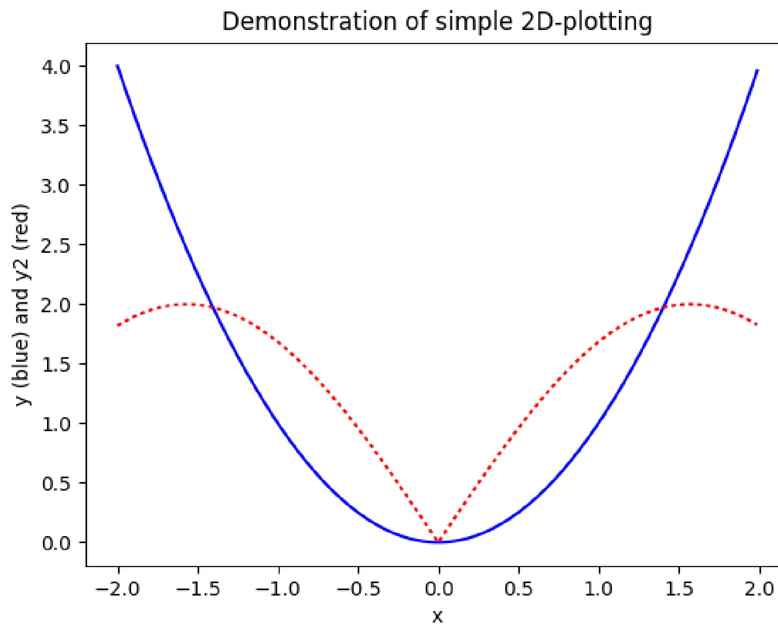

Figure 6: Figure window with plot of $y = x^2$

4.2 3D plotting

In the previous section we demonstrated how to plot functional values of y against values of x. For 3D plots we want to plot values of z as a function of x og y, $z = f(x, y)$.

3D line plot We can plot a line through (x, y, z) points defined parametrically using the following instructions:

```
>>> import matplotlib as mpl
>>> from mpl_toolkits.mplot3d import Axes3D
>>> fig = plt.figure()
>>> ax = fig.gca(projection='3d')
>>> t=np.arange(0,20*np.pi,0.01)
>>> x=np.cos(t)
>>> y=np.sin(t)
>>> z=np.power(t,3)
>>> ax.plot(x, y, z)
>>> plt.xticks(np.arange(-1,1.1,0.5))
>>> plt.yticks(np.arange(-1,1.1,0.5))
>>> plt.xlabel('x')
>>> title('Example of a 3D plot');
```

The result can be seen in figure 7.

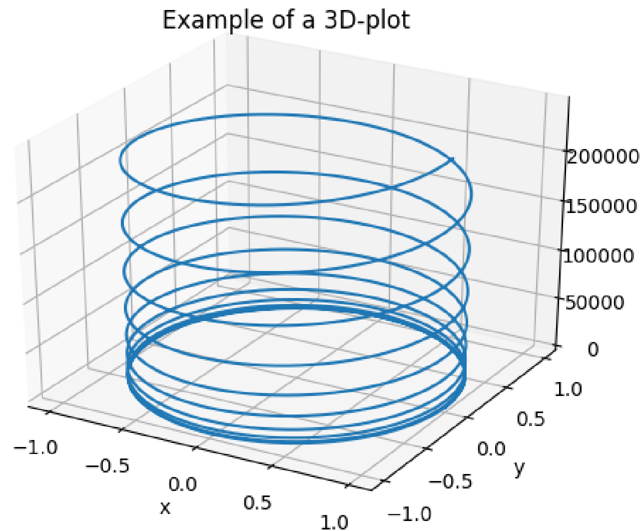

Figure 7: Figure window with 3D line plot

3D surface plot We can plot surface functions as for example the surface given as $z = e^{-x^2-y^2}$. We want to plot this function over the square $[-2, 2] \times [-2, 2]$. To do this we have to generate a grid of point of computations (computational grid) (x, y) within this square. This is done by first defining the computational grid along each axis and then to span the grid by using the function *meshgrid*.

```
>>> x=np.arange(-2,2,.2)
>>> y=np.arange(-2,2,.2)
>>> X, Y = np.meshgrid(x,y)
```

The computation of the function values and plotting of the surface using the function *mesh* renders the plot shown in figure 8.

```
>>> Z=np.exp(-np.square(X)-np.square(Y))
>>> fig = plt.figure()
>>> ax = fig.add_subplot(111, projection='3d')
>>> ax.plot_wireframe(X,Y,Z)
>>> plt.xlabel('x')
```

```
>>> plt.ylabel('y')
>>> plt.title('Example of a 3D mesh plot')
```

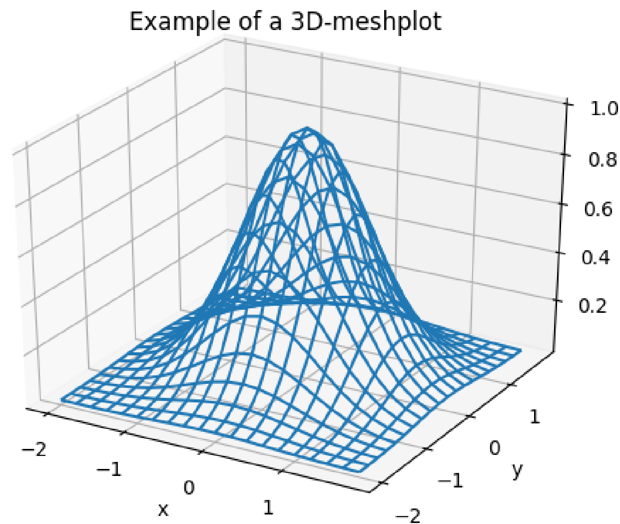

Figure 8: Figure window with 3D mesh plot

4.3 Exercises

a) Plot the signal $x = Acos(\omega_0 t + \theta_0)$, in the interval $t \in [0, 3\pi]$ for $\omega_0 = 2/3$, $A = 2$ and $\theta_0 = 0$. [3]

b) Display the function as a surface instead of mesh by using the function *plot_surface*.

c) Generate a 3D mesh plot of the two dimensional *sinc* function $z = sin(r)/r$, where $r = x^2 + y^2$ over the square $[-8, 8] \times [-8, 8]$. [4]

[3]Remember that $w_0 = 2\pi f_0$ and that the period is $T = 1/f_0$, so this corresponds to plotting one period of x.

[4]Remember that there will be a problem if you try to compute the function for $x = y = 0$. This can be handled by adding a value *eps* close to 0 to all points of computation.

5 Programming

As mentioned previously, you can make your own Python functions by writing programs in so called *M files*. The example in section 2.2 demonstrates how a set of indtructions can be run by calling an M file containing these instructions. This type of M files only containing a set of instructions are called *script* files.

In this section on programming, we shall focus on another type of M files that carries the term *function* files. This enables you to create your own Python functions with input and output parameters such as the sine function used as example of one of Python's built in functions in section 2.2.

5.1 Function files

As a simple example we will make a function for computing the normal probability density function,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}}. \tag{1}$$

Here we will use the variable x and the function parameters μ and σ as input parameters. p will be the output parameter. We will call the function *pdens.py*. The contents of the file might look like følger:

```
import numpy as np

def pdens(x, m, s):
    # PDENS Computes the probaaility density values
    # P=PDENS(X,M,S) computes the density value P for
    # X for a gaussian density function with
    # mean value M og standard deviation S.
    p = 1/(np.sqrt(2*np.pi)*s)* \
        np.exp(-1/2*np.square((x-m))/(2*np.square(s)))
    return p
```

Notice the help text where each comment line starts with a `#` so that Python won't interpret the line as an instruction to execute. Also, the function declaration by `def` ending with a `:`. The indentationn of code block within the finction body is also required. Note the input variables `x,m,s` and the output variable given by the `return` keyword.

To execute the function the following commands are given

```
>>> from pdens import pdens
>>> p=pdens(1,2,3)
>>> p
0.12933768267647971
```

Note that more than one function can be defined within a single .py file, so the filename does not need to be the same as the function name. Note the `import numpy as np` at the top of the file. Also note the way the function itself is imported from the file by `from pdens import pdens`.

5.2 Control structures

Python employs many of the control structures well known from other programming languages. We will look at the most important ones.

5.2.1 Conditional testing with *if*

One of the two important control strucures is *if*. In describing an algorithm, one often needs to choose between two different activities depending upon some test condition being true or false. In Python this can be expressed as

```
if condition:
    activity1
else:
    activity2
```

For example if one needs to make a procedure, *absolute*, that computes the absolute value $y = |x|$ of a number, x, so that $y = x$ when x is positive and $y = -x$ otherwise.

```
def absolute(x)

    # ABSOLUTE Computes the absolute value
    # Y=ABSOLUTE(X) computes the absolute value, Y, of X

    if x < 0:
        y=-x
    else:
        y=x
    return y
```

Note the *key words if* and *else*. These key words are reserved for this purpose and may not be used for other purposes.

5.2.2 Iterative structures using *while* or *for*

while The second of the two important control structures is *while*. This makes it possible to repeat the execution of an activity as long as som condition is fullfilled. Combined with conditional testing this gives the opportunity to describe algorithms in a powerful way. In Python this may be expressed as

```
while condition:
    activity(s)
```

As long as the condition is true, execute the activity. Return and execute the activity again. When the condition is false, the iterations terminate and the execution continues at the first instructions after the while structure.

As an example one might want to make a function, *divideby2*, which divides an integer n with 2 as many times as possible. The commands `fix` and `rem` are used to compute the integer quotient and the remainder respectively.

```
def divideby2(n)

    # DIVIDEBY2 Divide by 2 as long as the remainder is zero
    # Q=DIVIDEBY2(N) computes the quotient, Q,
    # the maximum number of divisions by 2

    q=n
    while np.remainder(q,2) == 0:
        q=np.fix(q/2)
    return q
```

As we can see `q` is a local variable that changes and is a part of the test condidion. This is an essential part of the control part of the while structure which consists of the components

- Initialisation: q is set equal to the integer to be divided by 2.
- Test : it is tested if q can be divided by 2.
- Modification: q is changed to the next quotient which will be attempted to be divided by 2. This modification ensures that the terminal condidion will be reached.

for It is worth noting that the *while* loop repeats an undetermined number of times. If you want to repeat a number of instructions a fixed, predetermined number of times, it is convenient to use *for* loops.

As an example we want to modify the function *pdens* so that you can compute the density values for more than one point of computation at a time. A possible way to do this:

```
def p=pdens2(x,m,s)

    # PDENS Computes the probaaility density values
    # P=PDENS(X,M,S) computes the density value P for
    # the values in a vector X for a gaussian density function
    # with mean value M og standard deviation S.

    N = np.shape(x)[0]
    p = np.zeros(np.shape(x))
    for n in np.arange(0, N - 1):
        p[n] = 1 / (np.sqrt(2 * np.pi) * s) * \
            np.exp(-1 / 2 * np.square((x[n] - m)) / (2 * np.square(s)))
    return p
```

As you can see very few modifications are needed to make the function handle vectors. The number of iterations are determined usint the *length* function, and the loop counter `n` is used for vector indexing of both `x` og `p`.

5.3 Exercises

a) Make a function to determine whether a number is odd or even. Return 1 if the number is odd, otherwise return 0.

b) Make a function that generates a vector of random numbers (integers in the range $[0, 9]$ which can be generated by use of *random.randint*). Further to this, the function shall distribute the numbers to two vectors, one with the odd numbers and the other with the even mumbers.

6 Analysis

w To finish this Python introduction, we present an example where Python is used for analysis of realistic data. This example will illustrate *reading* data

from a file, *analysis* of the signals decoded from the data. This analysis will involve *spectral analysis*, a central concept in signal analysis, and furthermore *filtering* of the signal. This filtering will make it possible for you to solve the *problem*. In the example the step by step procedure for reaching the will be illustrated. In the final part of the exercise you will have to do the calculations leading to the solution of the problem.

6.1 The problem - computing the heart rate

In this exercise you will be working with a photoplethysmographic dataset measured on one of the employees at the institute's laboratory of *medical engineering*. The data file is named *photopl.txt*. You will also need the function *findpeaks.m*[5]. This signal carries information on the pulse beats measured from the employee.

Using the signal, you shall compute the number of beats per minute (the heart rate) and how much the time between the individual beats vary (the heart rate variability). In brief, the method for doing this will be based on finding the time instants for each individual beat.

6.2 Solving the problem

When working with this kind of problem, it will obviously be useful to study the measured signal to consider a possible method for solving the given problem. The first step will be to read the signal from the data file.

Reading the input If the data file is opened in an editor e.g. *Notepad* you will see that the data are ASCII-coded and structured into one channel. The four first lines in the file give information on what kind of measurements the file contains in addition to date of recording, sample frequency and the number of recorded samples. To be able to analyse and perform arithmetic operations on the samples, the data has to be read into Python where the meaured sampled are organised in a vector. This can be done by giving the following commands:

```
>>> x=np.genfromtxt('photopl.txt',dtype=None,delimiter=None,skip_header=4)
>>> x=x/np.std(x)
```

The function *csvread* is used to decode the ASCII coded data from line 5 and further on, while dividing by the standard deviation is done to scale the signal

[5]The files *photopl.txt* and *findpeaks.m* can be downloaded from www.ux.uis.no/˜trygve-e

to unity variance which will be convenient in the further analysis.

Plotting the signal When we want to plot the signal, it will also be convenient to generate a time vector for the signal. Considering the four first lines in the file, we know that the sampling frequency is 40 Hz. The commands for plotting the necessary information can be like

```
>>> fs = 40.0
>>> N = np.size(x)
>>> t = np.arange(0, N) / fs
>>> t = t.reshape(t.size, 1)
```

The detections can be done by using the function *findpeaks*. The folowing commans illustrate the usage of this function and the visualisation of the detected peaks.

```
>>> from analyse import findpeaks
>>> ind, peaks = findpeaks(x)
```

The plotting is done as shown below.

```
>>> fig1 = plt.figure(1)
>>> ax1 = plt.axes()
>>> ax1.plot(t, x, 'b', t[ind, 0], x[ind, 0], 'go')
>>> plt.xlabel('t [sek]';
>>> plt.ylabel('x(t)')
>>> plt.grid('on')
>>> fig1.set_size_inches(8, 3)
```

The blue curve shown in 9 show the recorded signal, and the individual pulse beats are clearly displayed as a periodic signal with approximately one beat per minute. One immediate idea is that the pulse rate can be determined by first detecting the beats representing the individual pulse beats.

The variable `n` contains the sample number of the local maxima detected in `x` using the *findpeaks* function. The command *hold on* makes it possible to plot the detected tops superimposed on the pulse curve plotted previously. The plotting of the peaks is doe by indexing the `t` and `x` and specifying their representation as green circles ('go'). The command *axis* is used to delimit the plot along the axes.

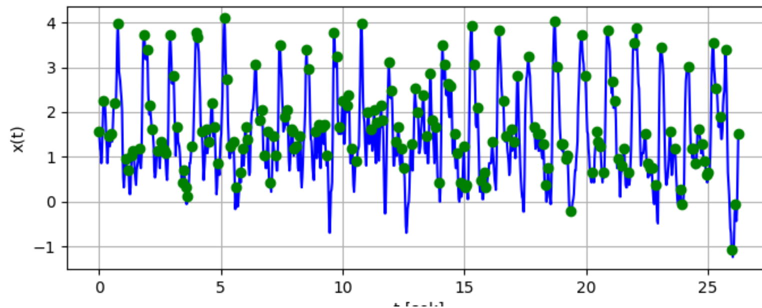

Figure 9: Pulse signal (blue curve) and detected peaks (green circles)

The detected tops are thus shown as green circles superimposed on the pulse curve as shown in figure 9. We can see that the pulse beats are detected according to our specification, but that the result includes a large number of false detections as well.

Spectral analysis of the signal When we study the signal, we can see that the pulse beats have a periodic nature. Using *filtering* techinques will possibly enable us to emphasise the periodic component corresponding to the pulse beats we want to detect. At the same time the filter should suppress the components corresponding to the false detections.

To design such a filter we need to know the exact frequency the pulse beats correspond to.

The spectral analysis is done by estimating the spectrum of the signal and then visualising it. This can be done by the following commands:

```
>>> f,Pxx = signal.welch(x.reshape(1,x.size),fs,nperseg=256,
>>>                      nfft=1024,detrend=False,scaling='density')
>>> Pxx = Pxx.reshape(Pxx.size, 1)
>>> f = f.reshape(f.size, 1)
>>> fig2 = plt.figure(2)
>>> ax21 = fig2.add_subplot(211)
>>> ax21.plot(f, 10 * np.log10(Pxx), 'b')
>>> plt.xlabel('f [Hz]')
>>> plt.ylabel('Magnitude [dB]')
>>> plt.grid('on')
>>> plt.axis([0, 20, -40, 20])
```

The spectral estimate is computed by the function *welch* which is one of a number of possible methods. The spectrum is shown in the upper part of figure 10. The command *add_subplot* is used to make a split figure window.

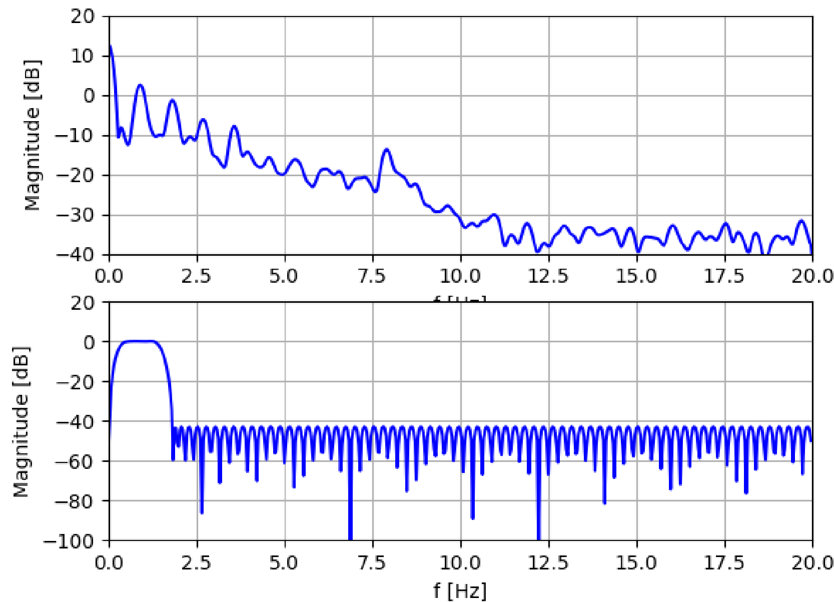

Figure 10: Spectrum (top) and designed band pass filter (bottom)

As can be seen from studying the spectrum, the signal has a dominant frequency component at 1 Hz. This corresponds to the periodic component of the pulse beats which we previously roughly estimated to be around 1 beat per second.

Filter design We want to design a band pass filter emphasising the frequency comonent at 1 Hz while suppressing other components.

```
>>> dt = 0.55
>>> pb = np.array([0,0.58-dt,0.58,1.27,1.27+dt,20])/(40)
>>> b = signal.remez(150, pb, [0, 1, 0], type='bandpass')
>>> w, h = signal.freqz(b)
>>> ax22 = fig2.add_subplot(212)
>>> ax22.plot(w / (2 * np.pi) * 40, 20 * np.log10(np.abs(h)), 'b')
>>> plt.xlabel('f [Hz]')
>>> plt.ylabel('Magnitude [dB]')
>>> plt.grid('on')
>>> plt.axis([0, 20, -40, 20])
```

The command *fir1* designs the filter where the first parameter states the order of the filter (the higher order, the steeper transition area). The other parameter indicates the lower and upper cutoff frequency for the pass band 0.58 og 1.27

Hz respectively (division by 20 corresponds to normalising to half the sampling frequency). In the frequency domain, the result of filtering will correspond to multiplication of the spectrum at the top with th efrequency response of the filter at the bottom of the figure.

Now we have come to the part where you perform the last steps to reach the solution of the problem. You will have to filter the signal and make a new detection attempt.a er vi kommet dithen at signalet kan filtreres og ny deteksjon kan utføres. But you will be on your own now (See the exercises given below).

6.3 Exercises

a) Use the function *signal.filtfilt* to filter the pulse signal x.[6]

b) Plot the spectrum of the filtered signal (red curve) in the same plot as the original signal (blue curve). What has been the effect of the filter?

c) Do the peak detection once more, but this time on the filtered signal.

d) Visualise the filtered signal and the recently detected peaks together in the same plot.

e) Determine the heart rate[7].

f) Determine the heart rate variability[8].

[6] In the function documentation the syntax is given as y=signal.filtfilt(b,a,x,...). Set a=1;

[7] It will be useful to apply the function *mean* to compute the average value.

[8] Express the variability using the standard deviation which can be computed by using the function *std* from Numpy.