PHY321: Introduction to Classical Mechanics and plans for Spring 2021

Morten Hjorth-Jensen[1,2]

Scott Pratt[1]

[1]Department of Physics and Astronomy and Facility for Rare Ion Beams (FRIB), Michigan State University, USA
[2]Department of Physics, University of Oslo, Norway

Jan 10, 2021

Aims and Overview

The first week starts on Monday January 11. This week is dedicated to a review of learning material and reminder on programming aspects, useful tools, where to find information and much more.

- Monday: Introduction to the course and reminder on vectors, space, time and motion.
- Wednesday: Python programming reminder, elements from CS201 and how they are used in this course. Installing software (anaconda). See slides at `https://mhjensen.github.io/Physics321/doc/pub/week2/html/week2.html`).
- Friday: Introduction to Git and GitHub and getting started with numerical exercises.

Recommended reading: John R. Taylor, Classical Mechanics (Univ. Sci. Books 2005), `https://www.uscibooks.com/taylor2.htm`, see also `https://github.com/mhjensen/Physics321/tree/master/doc/Literature`. Chapters 1.2 and 1.3 of Taylor.

Recommenended review: Brush up your Python programming knowledge, in particular how you used **numpy** to handle arrays. Look at some of the programs in the slides here for Wednesday and Friday.

Monday

Overview video

Introduction

Classical mechanics is a topic which has been taught intensively over several centuries. It is, with its many variants and ways of presenting the educational material, normally the first **real** physics course many of us meet and it lays the foundation for further physics studies. Many of the equations and ways of reasoning about the underlying laws of motion and pertinent forces, shape our approaches and understanding of the scientific method and discourse, as well as the way we develop our insights and deeper understanding about physical systems.

There is a wealth of well-tested (from both a physics point of view and a pedagogical standpoint) exercises and problems which can be solved analytically. However, many of these problems represent idealized and less realistic situations. The large majority of these problems are solved by paper and pencil and are traditionally aimed at what we normally refer to as continuous models from which we may find an analytical solution. As a consequence, when teaching mechanics, it implies that we can seldomly venture beyond an idealized case in order to develop our understandings and insights about the underlying forces and laws of motion.

Space, Time, Motion, Reference Frames and Reminder on vectors and other mathematical quantities

Our studies will start with the motion of different types of objects such as a falling ball, a runner, a bicycle etc etc. It means that an object's position in space varies with time. In order to study such systems we need to define

- choice of origin
- choice of the direction of the axes
- choice of positive direction (left-handed or right-handed system of reference)
- choice of units and dimensions

These choices lead to some important questions such as

- is the physics of a system independent of the origin of the axes?
- is the physics independent of the directions of the axes, that is are there privileged axes?
- is the physics independent of the orientation of system?
- is the physics independent of the scale of the length?

Dimension, units and labels. Throughout this course we will use the standardized SI units. The standard unit for length is thus one meter 1m, for mass one kilogram 1kg, for time one second 1s, for force one Newton 1kgm/s^2 and for energy 1 Joule $1\text{kgm}^2\text{s}^{-2}$.

We will use the following notations for various variables (vectors are always boldfaced in these lecture notes):

- position $\boldsymbol{r}$, in one dimention we will normally just use x,
- mass m,
- time t,
- velocity $\boldsymbol{v}$ or just v in one dimension,
- acceleration $\boldsymbol{a}$ or just a in one dimension,
- momentum $\boldsymbol{p}$ or just p in one dimension,
- kinetic energy K,
- potential energy V and
- frequency ω.

More variables will be defined as we need them.

Dimensions and Units

It is also important to keep track of dimensionalities. Don't mix this up with a chosen unit for a given variable. We mark the dimensionality in these lectures as $[a]$, where a is the quantity we are interested in. Thus

- $[\boldsymbol{r}] =$ length
- $[m] =$ mass
- $[K] =$ energy
- $[t] =$ time
- $[\boldsymbol{v}] =$ length over time
- $[\boldsymbol{a}] =$ length over time squared
- $[\boldsymbol{p}] =$ mass times length over time
- $[\omega] = 1/$time

Elements of Vector Algebra

In these lectures we will use boldfaced lower-case letters to label a vector. A vector $\boldsymbol{a}$ in three dimensions is thus defined as

$$\boldsymbol{a} = (a_x, a_y, a_z),$$

and using the unit vectors in a cartesian system we have

$$\boldsymbol{a} = a_x\boldsymbol{e}_x + a_y\boldsymbol{e}_y + a_z\boldsymbol{e}_z,$$

where the unit vectors have magnitude $|\boldsymbol{e}_i| = 1$ with $i = x, y, z$.

Using the fact that multiplication of reals is distributive we can show that

$$\boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a}\boldsymbol{b} + \boldsymbol{a}\boldsymbol{c},$$

Similarly we can also show that (using product rule for differentiating reals)

$$\frac{d}{dt}(\boldsymbol{a}\boldsymbol{b}) = \boldsymbol{a}\frac{d\boldsymbol{b}}{dt} + \boldsymbol{b}\frac{d\boldsymbol{a}}{dt}.$$

We can repeat these operations for the cross products and show that they are distribuitive

$$\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}.$$

We have also that

$$\frac{d}{dt}(\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a} \times \frac{d\boldsymbol{b}}{dt} + \boldsymbol{b} \times \frac{d\boldsymbol{a}}{dt}.$$

The rotation of a three-dimensional vector $\boldsymbol{a} = (a_x, a_y, a_z)$ in the xy plane around an angle ϕ results in a new vector $\boldsymbol{b} = (b_x, b_y, b_z)$. This operation can be expressed in terms of linear algebra as a matrix (the rotation matrix) multiplied with a vector. We can write this as

$$\begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}.$$

We can write this in a more compact form as $\boldsymbol{b} = \boldsymbol{R}\boldsymbol{a}$, where the rotation matrix is defined as

$$\boldsymbol{R} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Scalars, Vectors and Matrices

A scalar is something with a value that is independent of coordinate system. Examples are mass, or the relative time between events. A vector has magnitude and direction. Under rotation, the magnitude stays the same but the direction changes. Scalars have no spatial index, whereas a three-dimensional vector has

3 indices, e.g. the position $\boldsymbol{r}$ has components r_1, r_2, r_3, which are often referred to as x, y, z.

There are several categories of changes of coordinate system. The observer can translate the origin, might move with a different velocity, or might rotate her/his coordinate axes. For instance, a particle's position vector changes when the origin is translated, but its velocity does not. When you study relativity you will find that quantities you thought of as scalars, such as time or an electric potential, are actually parts of four-dimensional vectors and that changes of the velocity of the reference frame act in a similar way to rotations.

In addition to vectors and scalars, there are matrices, which have two indices. One also has objects with 3 or four indices. These are called tensors of rank n, where n is the number of indices. A matrix is a rank-two tensor. The Levi-Civita symbol, ϵ_{ijk} used for cross products of vectors, is a tensor of rank three.

Unit Vectors

Also known as basis vectors, unit vectors point in the direction of the coordinate axes, have unit norm, and are orthogonal to one another. Sometimes this is referred to as an orthonormal basis,

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1}$$

Here, δ_{ij} is unity when $i = j$ and is zero otherwise. This is called the unit matrix, because you can multiply it with any other matrix and not change the matrix. The **dot** denotes the dot product, $\boldsymbol{A} \cdot \boldsymbol{B} = A_1 B_1 + A_2 B_2 + A_3 B_3 = |A||B| \cos\theta_{AB}$. Sometimes the unit vectors are called $\hat{x}$, $\hat{y}$ and $\hat{z}$. Vectors can be decomposed in terms of unit vectors,

$$\boldsymbol{r} = r_1 \hat{e}_1 + r_2 \hat{e}_2 + r_3 \hat{e}_3. \tag{2}$$

The vector components r_1, r_2 and r_3 might be called x, y and z for a displacement of v_x, v_y and v_z for a velocity.

From here on in, we may use bold face to denote vectors, e.g. $\mathbf{v}$ instead of $\boldsymbol{v}$. We also might use dots over quantities to represent time derivatives, e.g. $\dot{\mathbf{v}} = d\mathbf{v}/dt$. As mentioned above, repeated indices infer sums, e.g.,

$$\begin{aligned} x_i y_i &= \sum_i x_i y_i = \boldsymbol{x} \cdot \boldsymbol{y}, \\ x_i A_{ij} y_j &= \sum_{ij} x_i A_{ij} y_j, \end{aligned} \tag{3}$$

Vector Operations, Scalar Product (or dot product)

For vectors $\boldsymbol{A}$ and $\boldsymbol{B}$,

$$\begin{aligned} \boldsymbol{A}\cdot\boldsymbol{B} &= A_iB_i = |A||B|\cos\theta_{AB}, \\ |A| &\equiv \sqrt{\boldsymbol{A}\cdot\boldsymbol{A}}. \end{aligned}$$

Note that the summation sign is inferred any time there are repeated indices. For example,

$$A_iB_i = \sum A_iB_i. \tag{4}$$

Vector Product (or cross product) of vectors $\boldsymbol{A}$ and $\boldsymbol{B}$

$$\begin{aligned} \boldsymbol{C} &= \boldsymbol{A}\times\boldsymbol{B}, \\ C_i &= \epsilon_{ijk}A_jB_k. \end{aligned}$$

Here ϵ is the third-rank anti-symmetric tensor, also known as the Levi-Civita symbol. It is ±1 only if all three indices are different, and is zero otherwise. The choice of ±1 depends on whether the indices are an even or odd permutation of the original symbols. The permutation xyz or 123 is considered to be $+1$. For the 27 elements,

$$\begin{aligned} \epsilon_{ijk} &= -\epsilon_{ikj} = -\epsilon_{jik} = -\epsilon_{kji} \\ \epsilon_{123} &= \epsilon_{231} = \epsilon_{312} = 1, \\ \epsilon_{213} &= \epsilon_{132} = \epsilon_{321} = -1, \\ \epsilon_{iij} &= \epsilon_{iji} = \epsilon_{jii} = 0. \end{aligned} \tag{5}$$

You used cross products extensively when studying magnetic fields. Because the matrix is anti-symmetric, switching the x and y axes (or any two axes) flips the sign. If the coordinate system is right-handed, meaning the xyz axes satisfy $\hat{x}\times\hat{y}=\hat{z}$, where you can point along the x axis with your extended right index finger, the y axis with your contracted middle finger and the z axis with your extended thumb. Switching to a left-handed system flips the sign of the vector $\boldsymbol{C}=\boldsymbol{A}\times\boldsymbol{B}$. Note that $\boldsymbol{A}\times\boldsymbol{B}=-\boldsymbol{B}\times\boldsymbol{A}$. The vector $\boldsymbol{C}$ is perpendicular to both $\boldsymbol{A}$ and $\boldsymbol{B}$ and the magnitude of $\boldsymbol{C}$ is given by

$$|C| = |A||B|\sin\theta_{AB}.$$

Vectors obtained by the cross product of two real vectors are called pseudo-vectors because the assignment of their direction can be arbitrarily flipped by defining the Levi-Civita symbol to be based on left-handed rules. Examples are the magnetic field and angular momentum. If the direction of a real vector prefers the right-handed over the left-handed direction, that constitutes a violation of parity. For instance, one can polarize the spins (angular momentum) of nuclei

with a magnetic field so that the spins preferentially point along the direction of the magnetic field. This does not violate parity because both are pseudo-vectors. Now assume these polarized nuclei decay and that electrons are one of the products. If these electrons prefer to exit the decay parallel vs. antiparallel to the polarizing magnetic field, this constitutes parity violation because the direction of the outgoing electron momenta are a real vector. This is precisely what is observed in weak decays.

Differentiation of a vector with respect to a scalar

For example, the acceleration is $d\boldsymbol{v}/dt$:

$$(d\boldsymbol{v}/dt)_i = \frac{dv_i}{dt}.$$

Angular velocity

Choose the vector $\boldsymbol{\omega}$ so that the motion is like your fingers wrapping around your thumb on your right hand with $\boldsymbol{\omega}$ pointing along your thumb. The magnitude is $d|\phi|/dt$.

Gradient operator ∇

This is the derivative $\partial/\partial x$, $\partial/\partial y$ and $\partial/\partial z$, where ∂_x means ∂/∂_x. For taking the gradient of a scalar Φ,

$$\textbf{grad } \Phi, (\nabla\Phi(x,y,z,t))_i = \partial/\partial r_i \Phi(\boldsymbol{r},t) = \partial_i\Phi(\boldsymbol{r},t).$$

For taking the dot product of the gradient with a vector, sometimes called a divergence,

$$div\boldsymbol{A}, \nabla\cdot\boldsymbol{A} = \partial_i A_i.$$

For taking the vector product with another vector, sometimes called curl, $\nabla\times\boldsymbol{A}$,

$$\textbf{curl } \boldsymbol{A}, (\nabla\times\boldsymbol{A})_i = \epsilon_{ijk}\partial_j A_k(\boldsymbol{r},t).$$

The Laplacian is referred to as ∇^2 and is defined as

$$\nabla^2 = \nabla\cdot\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Some identities

Here we simply state these, but you may wish to prove a few. They are useful for this class and will be essential when you study E&M.

$$\begin{aligned} \boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C}) &= \boldsymbol{B}\cdot(\boldsymbol{C}\times\boldsymbol{A}) = \boldsymbol{C}\cdot(\boldsymbol{A}\times\boldsymbol{B}) \\ \boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C}) &= (\boldsymbol{A}\cdot\boldsymbol{C})\boldsymbol{B} - (\boldsymbol{A}\cdot\boldsymbol{B})\boldsymbol{C} \\ (\boldsymbol{A}\times\boldsymbol{B})\cdot(\boldsymbol{C}\times\boldsymbol{D}) &= (\boldsymbol{A}\cdot\boldsymbol{C})(\boldsymbol{B}\cdot\boldsymbol{D}) - (\boldsymbol{A}\cdot\boldsymbol{D})(\boldsymbol{B}\cdot\boldsymbol{C}) \end{aligned} \tag{6}$$

Example The height of a hill is given by the formula

$$z(x,y) = 2xy - 3x^2 - 4y^2 - 18x + 28y + 12.$$

Here z is the height in meters and x and y are the east-west and north-south coordinates. Find the position x, y where the hill is the highest, and give its height.

Solution: The maxima or minima, or inflection points, are given when $\partial_x z = 0$ and $\partial_y z = 0$.

$$\begin{aligned} \partial_x z(x,y) &= 2y - 6x - 18 = 0, \\ \partial_y z(x,y) &= 2x - 8y + 28 = 0. \end{aligned}$$

Solving for two equations and two unknowns gives one solution $x = -2, y = 3$. This then gives $z = 72$. Although this procedure could have given a minimum or an inflection point you can look at the form for z and see that for $x = y = 0$ the height is lower, therefore it is not a minimum. You can also look and see that the quadratic contributions are negative so the height falls off to $-\infty$ far away in any direction. Thus, there must be a maximum, and this must be it. Deciding between maximum or minimum or inflection point for a general problem would involve looking at the matrix $\partial_i \partial_j z$ at the specific point, then finding the eigenvalues of the matrix and seeing whether they were both positive (minimum) both negative (maximum) or one positive and one negative (saddle point).

Gauss's Theorem

For an integral over a volume V confined by a surface S, Gauss's theorem gives

$$\int_V dv\ \nabla\cdot\boldsymbol{A} = \int_S d\boldsymbol{S}\cdot\boldsymbol{A}.$$

For a closed path C which carves out some area S,

$$\int_C d\boldsymbol{\ell}\cdot\boldsymbol{A} = \int_S d\boldsymbol{s}\cdot(\nabla\times\boldsymbol{A})$$

and Stokes's Theorem

Stoke's law can be understood by considering a small rectangle, $-\Delta x < x < \Delta x$, $-\Delta y < y < \Delta y$. The path integral around the edges is

$$\begin{aligned}
\int_C d\boldsymbol{\ell}\cdot\boldsymbol{A} &= 2\Delta y[A_y(\Delta x,0)-A_y(-\Delta x,0)]-2\Delta x[A_x(0,\Delta y)-A_x(0,-\Delta y)] \quad (7)\\
&= 4\Delta x\Delta y\left\{\frac{A_y(\Delta x,0)-A_y(-\Delta x,0)}{2\Delta x}-\frac{A_x(0,\Delta y)-A_x(0,-\Delta y)}{2\Delta y}\right\}\\
&= 4\Delta x\Delta y\left\{\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right\}\\
&= \Delta S\cdot\nabla\times\boldsymbol{A}. \quad (8)
\end{aligned}$$

Here ΔS is the area of the surface element.

Rotations

Note: The material on matrices is optional and will not be used much (except for illustrations at the very end on garmonic oscillations) since most of you have not yet taken a course on linear algebra. The material is however included here for the sake of completeness.

Here, we use rotations as an example of matrices and their operations. One can consider a different orthonormal basis $\hat{e}'_1$, $\hat{e}'_2$ and $\hat{e}'_3$. The same vector $\boldsymbol{r}$ mentioned above can also be expressed in the new basis,

$$\boldsymbol{r} = r'_1\hat{e}'_1 + r'_2\hat{e}'_2 + r'_3\hat{e}'_3. \quad (9)$$

Even though it is the same vector, the components have changed. Each new unit vector $\hat{e}'_i$ can be expressed as a linear sum of the previous vectors,

$$\hat{e}'_i = \sum_j U_{ij}\hat{e}_j, \quad (10)$$

and the matrix U can be found by taking the dot product of both sides with $\hat{e}_k$,

$$\begin{aligned}
\hat{e}_k\cdot\hat{e}'_i &= \sum_j U_{ij}\hat{e}_k\cdot\hat{e}_j\\
\hat{e}_k\cdot\hat{e}'_i &= \sum_j U_{ij}\delta_{jk} = U_{ik}. \quad (11)
\end{aligned}$$

Thus, the matrix lambda has components U_{ij} that are equal to the cosine of the angle between new unit vector $\hat{e}'_i$ and the old unit vector $\hat{e}_j$.

$$U = \begin{pmatrix} \hat{e}'_1\cdot\hat{e}_1 & \hat{e}'_1\cdot\hat{e}_2 & \hat{e}'_1\cdot\hat{e}_3 \\ \hat{e}'_2\cdot\hat{e}_1 & \hat{e}'_2\cdot\hat{e}_2 & \hat{e}'_2\cdot\hat{e}_3 \\ \hat{e}'_3\cdot\hat{e}_1 & \hat{e}'_3\cdot\hat{e}_2 & \hat{e}'_3\cdot\hat{e}_3 \end{pmatrix}, \qquad U_{ij} = \hat{e}'_i\cdot\hat{e}_j = \cos\theta_{ij}. \quad (12)$$

Note that the matrix is not symmetric, $U_{ij} \neq U_{ji}$. One can also look at the inverse transformation, by switching the primed and unprimed coordinates,

$$\begin{aligned} \hat{e}_i &= \sum_j U^{-1}_{ij} \hat{e}'_j, \\ U^{-1}_{ij} &= \hat{e}_i \cdot \hat{e}'_j = U_{ji}. \end{aligned} \tag{13}$$

The definition of transpose of a matrix, $M^t_{ij} = M_{ji}$, allows one to state this as

$$U^{-1} = U^t. \tag{14}$$

A tensor obeying Eq. (14) defines what is known as a unitary, or orthogonal, transformation.

The matrix U can be used to transform any vector to the new basis. Consider a vector

$$\begin{aligned} \boldsymbol{r} &= r_1\hat{e}_1 + r_2\hat{e}_2 + r_3\hat{e}_3 \\ &= r'_1\hat{e}'_1 + r'_2\hat{e}'_2 + r'_3\hat{e}'_3. \end{aligned} \tag{15}$$

This is the same vector expressed as a sum over two different sets of basis vectors. The coefficients r_i and r'_i represent components of the same vector. The relation between them can be found by taking the dot product of each side with one of the unit vectors, $\hat{e}_i$, which gives

$$r_i = \sum_j \hat{e}_i \cdot \hat{e}'_j \; r'_j. \tag{16}$$

Using Eq. (13) one can see that the transformation of r can be also written in terms of U,

$$r_i = \sum_j U^{-1}_{ij} \; r'_j. \tag{17}$$

Thus, the matrix that transforms the coordinates of the unit vectors, Eq. (13) is the same one that transforms the coordinates of a vector, Eq. (17).

Find the rotation matrix U for finding the components in the primed coordinate system given from those in the unprimed system, given that the unit vectors in the new system are found by rotating the coordinate system by and angle ϕ about the z axis.

In this case

$$\begin{aligned} \hat{e}'_1 &= \cos\phi\hat{e}_1 - \sin\phi\hat{e}_2, \\ \hat{e}'_2 &= \sin\phi\hat{e}_1 + \cos\phi\hat{e}_2, \\ \hat{e}'_3 &= \hat{e}_3. \end{aligned}$$

By inspecting Eq. (11),

$$U = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Under a unitary transformation U (or basis transformation) scalars are unchanged, whereas vectors $\boldsymbol{r}$ and matrices M change as

$$\begin{aligned} r'_i &= U_{ij}\, r_j, \quad \text{(sum inferred)} \\ M'_{ij} &= U_{ik} M_{km} U^{-1}_{mj}. \end{aligned} \tag{18}$$

Physical quantities with no spatial indices are scalars (or pseudoscalars if they depend on right-handed vs. left-handed coordinate systems), and are unchanged by unitary transformations. This includes quantities like the trace of a matrix, the matrix itself had indices but none remain after performing the trace.

$$\mathrm{Tr}M \equiv M_{ii}. \tag{19}$$

Because there are no remaining indices, one expects it to be a scalar. Indeed one can see this,

$$\begin{aligned} \mathrm{Tr}M' &= U_{ij} M_{jm} U^{-1}_{mi} \\ &= M_{jm} U^{-1}_{mi} U_{ij} \\ &= M_{jm}\delta_{mj} \\ &= M_{jj} = \mathrm{Tr}M. \end{aligned} \tag{20}$$

A similar example is the determinant of a matrix, which is also a scalar.

Multiplying a matrix C times a vector $\boldsymbol{A}$

$$(CA)_i = C_{ij} A_j.$$

For the i^{th} element one takes the scalar product of the i^{th} row of C with the vector $\boldsymbol{A}$.

Mutiplying matrices C and D

$$(CD)_{ij} = C_{ik} D_{kj},$$

This means the one obtains the ij element by taking the scalar product of the i^{th} of C with the j^{th} column of D.

Wednesday: Python programming reminder, elements from CS201 and how they are used in this course

Overview video

Numerical Elements

Numerical algorithms call for approximate discrete models and much of the development of methods for continuous models are nowadays being replaced by methods for discrete models in science and industry, simply because **much larger classes of problems can be addressed** with discrete models, often by simpler and more generic methodologies.

As we will see throughout this course, when properly scaling the equations at hand, discrete models open up for more advanced abstractions and the possibility to study real life systems, with the added bonus that we can explore and deepen our basic understanding of various physical systems

Analytical solutions are as important as before. In addition, such solutions provide us with invaluable benchmarks and tests for our discrete models. Such benchmarks, as we will see below, allow us to discuss possible sources of errors and their behaviors. And finally, since most of our models are based on various algorithms from numerical mathematics, we have a unique oppotunity to gain a deeper understanding of the mathematical approaches we are using.

With computing and data science as important elements in essentially all aspects of a modern society, we could then try to define Computing as **solving scientific problems using all possible tools, including symbolic computing, computers and numerical algorithms, and analytical paper and pencil solutions**. Computing provides us with the tools to develope our own understanding of the scientific method by enhancing algorithmic thinking.

Computations and the Scientific Method

The way we will teach this course reflects this definition of computing. The course contains both classical paper and pencil exercises as well as computational projects and exercises. The hope is that this will allow you to explore the physics of systems governed by the degrees of freedom of classical mechanics at a deeper level, and that these insights about the scientific method will help you to develop a better understanding of how the underlying forces and equations of motion and how they impact a given system.

Furthermore, by introducing various numerical methods via computational projects and exercises, we aim at developing your competences and skills about these topics.

Computational Competences

These competences will enable you to

- understand how algorithms are used to solve mathematical problems,
- derive, verify, and implement algorithms,
- understand what can go wrong with algorithms,

- use these algorithms to construct reproducible scientific outcomes and to engage in science in ethical ways, and
- think algorithmically for the purposes of gaining deeper insights about scientific problems.

All these elements are central for maturing and gaining a better understanding of the modern scientific process *per se*.

The power of the scientific method lies in identifying a given problem as a special case of an abstract class of problems, identifying general solution methods for this class of problems, and applying a general method to the specific problem (applying means, in the case of computing, calculations by pen and paper, symbolic computing, or numerical computing by ready-made and/or self-written software). This generic view on problems and methods is particularly important for understanding how to apply available, generic software to solve a particular problem.

However, verification of algorithms and understanding their limitations requires much of the classical knowledge about continuous models.

A well-known examples to illustrate many of the above concepts

Before we venture into a reminder on Python and mechanics relevant applications, let us briefly outline some of the abovementioned topics using an example many of you may have seen before in for example CMSE201. A simple algorithm for integration is the Trapezoidal rule. Integration of a function $f(x)$ by the Trapezoidal Rule is given by following algorithm for an interval $x \in [a, b]$

$$\int_a^b (f(x)dx = \frac{1}{2}\left[f(a) + 2f(a+h) + \cdots + 2f(b-h) + f(b)\right] + O(h^2),$$

where h is the so-called stepsize defined by the number of integration points N as $h = (b-a)/(n)$. Python offers an extremely versatile programming environment, allowing for the inclusion of analytical studies in a numerical program. Here we show an example code with the **trapezoidal rule**. We use also **SymPy** to evaluate the exact value of the integral and compute the absolute error with respect to the numerically evaluated one of the integral $\int_0^1 dx x^2 = 1/3$. The following code for the trapezoidal rule allows you to plot the relative error by comparing with the exact result. By increasing to 10^8 points one arrives at a region where numerical errors start to accumulate.

```
from math import log10
import numpy as np
from sympy import Symbol, integrate
import matplotlib.pyplot as plt
# function for the trapezoidal rule
def Trapez(a,b,f,n):
   h = (b-a)/float(n)
   s = 0
```

```
    x = a
    for i in range(1,n,1):
        x = x+h
        s = s+ f(x)
    s = 0.5*(f(a)+f(b)) +s
    return h*s
#  function to compute pi
def function(x):
    return x*x
# define integration limits
a = 0.0;  b = 1.0;
# find result from sympy
# define x as a symbol to be used by sympy
x = Symbol('x')
exact = integrate(function(x), (x, a, b))
# set up the arrays for plotting the relative error
n = np.zeros(9); y = np.zeros(9);
# find the relative error as function of integration points
for i in range(1, 8, 1):
    npts = 10**i
    result = Trapez(a,b,function,npts)
    RelativeError = abs((exact-result)/exact)
    n[i] = log10(npts); y[i] = log10(RelativeError);
plt.plot(n,y, 'ro')
plt.xlabel('n')
plt.ylabel('Relative error')
plt.show()
```

Analyzing the above example

This example shows the potential of combining numerical algorithms with symbolic calculations, allowing us to

- Validate and verify their algorithms.
- Including concepts like unit testing, one has the possibility to test and test several or all parts of the code.
- Validation and verification are then included *naturally* and one can develop a better attitude to what is meant with an ethically sound scientific approach.
- The above example allows the student to also test the mathematical error of the algorithm for the trapezoidal rule by changing the number of integration points. The students get **trained from day one to think error analysis**.
- With a Jupyter notebook you can keep exploring similar examples and turn them in as your own notebooks.

Python practicalities, Software and needed installations

We will make extensive use of Python as programming language and its myriad of available libraries. You will find Jupyter notebooks invaluable in your work.

If you have Python installed (we strongly recommend Python3) and you feel pretty familiar with installing different packages, we recommend that you install the following Python packages via **pip** as

1. pip install numpy scipy matplotlib ipython scikit-learn mglearn sympy pandas pillow

For Python3, replace **pip** with **pip3**.

For OSX users we recommend, after having installed Xcode, to install **brew**. Brew allows for a seamless installation of additional software via for example

1. brew install python3

For Linux users, with its variety of distributions like for example the widely popular Ubuntu distribution, you can use **pip** as well and simply install Python as

1. sudo apt-get install python3 (or python for pyhton2.7)

etc etc.

Python installers

If you don't want to perform these operations separately and venture into the hassle of exploring how to set up dependencies and paths, we recommend two widely used distrubutions which set up all relevant dependencies for Python, namely

- Anaconda,

which is an open source distribution of the Python and R programming languages for large-scale data processing, predictive analytics, and scientific computing, that aims to simplify package management and deployment. Package versions are managed by the package management system **conda**.

- Enthought canopy

is a Python distribution for scientific and analytic computing distribution and analysis environment, available for free and under a commercial license.

Furthermore, Google's Colab is a free Jupyter notebook environment that requires no setup and runs entirely in the cloud. Try it out!

Useful Python libraries

Here we list several useful Python libraries we strongly recommend (if you use anaconda many of these are already there)

- NumPy is a highly popular library for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays

- The pandas library provides high-performance, easy-to-use data structures and data analysis tools
- Xarray is a Python package that makes working with labelled multi-dimensional arrays simple, efficient, and fun!
- Scipy (pronounced "Sigh Pie") is a Python-based ecosystem of open-source software for mathematics, science, and engineering.
- Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms.
- Autograd can automatically differentiate native Python and Numpy code. It can handle a large subset of Python's features, including loops, ifs, recursion and closures, and it can even take derivatives of derivatives of derivatives
- SymPy is a Python library for symbolic mathematics.
- scikit-learn has simple and efficient tools for machine learning, data mining and data analysis
- TensorFlow is a Python library for fast numerical computing created and released by Google
- Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano
- And many more such as pytorch, Theano etc

Your jupyter notebook can easily be converted into a nicely rendered **PDF** file or a Latex file for further processing. For example, convert to latex as

```
pycod jupyter nbconvert filename.ipynb --to latex
```

And to add more versatility, the Python package SymPy is a Python library for symbolic mathematics. It aims to become a full-featured computer algebra system (CAS) and is entirely written in Python.

Numpy examples and Important Matrix and vector handling packages

There are several central software libraries for linear algebra and eigenvalue problems. Several of the more popular ones have been wrapped into ofter software packages like those from the widely used text **Numerical Recipes**. The original source codes in many of the available packages are often taken from the widely used software package LAPACK, which follows two other popular packages developed in the 1970s, namely EISPACK and LINPACK. We describe them shortly here.

- LINPACK: package for linear equations and least square problems.
- LAPACK:package for solving symmetric, unsymmetric and generalized eigenvalue problems. From LAPACK's website http://www.netlib.org it is possible to download for free all source codes from this library. Both C/C++ and Fortran versions are available.
- BLAS (I, II and III): (Basic Linear Algebra Subprograms) are routines that provide standard building blocks for performing basic vector and matrix operations. Blas I is vector operations, II vector-matrix operations and III matrix-matrix operations. Highly parallelized and efficient codes, all available for download from http://www.netlib.org.

!spli

Basic Matrix Features

Matrix properties reminder.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \qquad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The inverse of a matrix is defined by

$$\mathbf{A}^{-1} \cdot \mathbf{A} = I$$

Relations	Name	matrix elements
$A = A^T$	symmetric	$a_{ij} = a_{ji}$
$A = \left(A^T\right)^{-1}$	real orthogonal	$\sum_k a_{ik}a_{jk} = \sum_k a_{ki}a_{kj} = \delta_{ij}$
$A = A^*$	real matrix	$a_{ij} = a^*_{ij}$
$A = A^\dagger$	hermitian	$a_{ij} = a^*_{ji}$
$A = \left(A^\dagger\right)^{-1}$	unitary	$\sum_k a_{ik}a^*_{jk} = \sum_k a^*_{ki}a_{kj} = \delta_{ij}$

Some famous Matrices.

- Diagonal if $a_{ij} = 0$ for $i \neq j$
- Upper triangular if $a_{ij} = 0$ for $i > j$
- Lower triangular if $a_{ij} = 0$ for $i < j$
- Upper Hessenberg if $a_{ij} = 0$ for $i > j + 1$
- Lower Hessenberg if $a_{ij} = 0$ for $i < j + 1$
- Tridiagonal if $a_{ij} = 0$ for $|i - j| > 1$

- Lower banded with bandwidth p: $a_{ij} = 0$ for $i > j + p$
- Upper banded with bandwidth p: $a_{ij} = 0$ for $i < j + p$
- Banded, block upper triangular, block lower triangular....

More Basic Matrix Features.

Some Equivalent Statements. For an $N \times N$ matrix $\mathbf{A}$ the following properties are all equivalent

- If the inverse of $\mathbf{A}$ exists, $\mathbf{A}$ is nonsingular.
- The equation $\mathbf{Ax} = 0$ implies $\mathbf{x} = 0$.
- The rows of $\mathbf{A}$ form a basis of R^N.
- The columns of $\mathbf{A}$ form a basis of R^N.
- $\mathbf{A}$ is a product of elementary matrices.
- 0 is not eigenvalue of $\mathbf{A}$.

Numpy and arrays

Numpy provides an easy way to handle arrays in Python. The standard way to import this library is as

```
import numpy as np
```

Here follows a simple example where we set up an array of ten elements, all determined by random numbers drawn according to the normal distribution,

```
n = 10
x = np.random.normal(size=n)
print(x)
```

We defined a vector x with $n = 10$ elements with its values given by the Normal distribution $N(0, 1)$. Another alternative is to declare a vector as follows

```
import numpy as np
x = np.array([1, 2, 3])
print(x)
```

Here we have defined a vector with three elements, with $x_0 = 1$, $x_1 = 2$ and $x_2 = 3$. Note that both Python and C++ start numbering array elements from 0 and on. This means that a vector with n elements has a sequence of entities $x_0, x_1, x_2, \ldots, x_{n-1}$. We could also let (recommended) Numpy to compute the logarithms of a specific array as

```
import numpy as np
x = np.log(np.array([4, 7, 8]))
print(x)
```

In the last example we used Numpy's unary function *np.log*. This function is highly tuned to compute array elements since the code is vectorized and does not require looping. We normaly recommend that you use the Numpy intrinsic functions instead of the corresponding **log** function from Python's **math** module. The looping is done explicitely by the **np.log** function. The alternative, and slower way to compute the logarithms of a vector would be to write

```
import numpy as np
from math import log
x = np.array([4, 7, 8])
for i in range(0, len(x)):
    x[i] = log(x[i])
print(x)
```

We note that our code is much longer already and we need to import the **log** function from the **math** module. The attentive reader will also notice that the output is $[1, 1, 2]$. Python interprets automagically our numbers as integers (like the **automatic** keyword in C++). To change this we could define our array elements to be double precision numbers as

```
import numpy as np
x = np.log(np.array([4, 7, 8], dtype = np.float64))
print(x)
```

or simply write them as double precision numbers (Python uses 64 bits as default for floating point type variables), that is

```
import numpy as np
x = np.log(np.array([4.0, 7.0, 8.0])
print(x)
```

To check the number of bytes (remember that one byte contains eight bits for double precision variables), you can use simple use the **itemsize** functionality (the array x is actually an object which inherits the functionalities defined in Numpy) as

```
import numpy as np
x = np.log(np.array([4.0, 7.0, 8.0])
print(x.itemsize)
```

Matrices in Python

Having defined vectors, we are now ready to try out matrices. We can define a 3×3 real matrix $\hat{A}$ as (recall that we user lowercase letters for vectors and uppercase letters for matrices)

```
import numpy as np
A = np.log(np.array([ [4.0, 7.0, 8.0], [3.0, 10.0, 11.0], [4.0, 5.0, 7.0] ]))
print(A)
```

If we use the **shape** function we would get $(3, 3)$ as output, that is verifying that our matrix is a 3×3 matrix. We can slice the matrix and print for example the first column (Python organized matrix elements in a row-major order, see below) as

```
import numpy as np
A = np.log(np.array([ [4.0, 7.0, 8.0], [3.0, 10.0, 11.0], [4.0, 5.0, 7.0] ]))
# print the first column, row-major order and elements start with 0
print(A[:,0])
```

We can continue this was by printing out other columns or rows. The example here prints out the second column

```
import numpy as np
A = np.log(np.array([ [4.0, 7.0, 8.0], [3.0, 10.0, 11.0], [4.0, 5.0, 7.0] ]))
# print the first column, row-major order and elements start with 0
print(A[1,:])
```

Numpy contains many other functionalities that allow us to slice, subdivide etc etc arrays. We strongly recommend that you look up the Numpy website for more details. Useful functions when defining a matrix are the **np.zeros** function which declares a matrix of a given dimension and sets all elements to zero

```
import numpy as np
n = 10
# define a matrix of dimension 10 x 10 and set all elements to zero
A = np.zeros( (n, n) )
print(A)
```

or initializing all elements to

```
import numpy as np
n = 10
# define a matrix of dimension 10 x 10 and set all elements to one
A = np.ones( (n, n) )
print(A)
```

or as unitarily distributed random numbers (see the material on random number generators in the statistics part)

```
import numpy as np
n = 10
# define a matrix of dimension 10 x 10 and set all elements to random numbers with x \in [0, 1]
A = np.random.rand(n, n)
print(A)
```

Meet the Pandas

Another useful Python package is pandas, which is an open source library providing high-performance, easy-to-use data structures and data analysis tools for Python. **pandas** stands for panel data, a term borrowed from econometrics and is an efficient library for data analysis with an emphasis on tabular data. **pandas** has two major classes, the **DataFrame** class with two-dimensional data objects and tabular data organized in columns and the class **Series** with a focus on one-dimensional data objects. Both classes allow you to index data easily as we will see in the examples below. **pandas** allows you also to perform mathematical operations on the data, spanning from simple reshapings of vectors and matrices to statistical operations.

The following simple example shows how we can, in an easy way make tables of our data. Here we define a data set which includes names, place of birth and date of birth, and displays the data in an easy to read way. We will see repeated use of **pandas**, in particular in connection with classification of data.

```
import pandas as pd
from IPython.display import display
data = {'First Name': ["Frodo", "Bilbo", "Aragorn II", "Samwise"],
        'Last Name': ["Baggins", "Baggins","Elessar","Gamgee"],
        'Place of birth': ["Shire", "Shire", "Eriador", "Shire"],
        'Date of Birth T.A.': [2968, 2890, 2931, 2980]
        }
data_pandas = pd.DataFrame(data)
display(data_pandas)
```

In the above we have imported **pandas** with the shorthand **pd**, the latter has become the standard way we import **pandas**. We make then a list of various variables and reorganize the above lists into a **DataFrame** and then print out a neat table with specific column labels as *Name*, *place of birth* and *date of birth*. Displaying these results, we see that the indices are given by the default numbers

from zero to three. **pandas** is extremely flexible and we can easily change the above indices by defining a new type of indexing as

```
data_pandas = pd.DataFrame(data,index=['Frodo','Bilbo','Aragorn','Sam'])
display(data_pandas)
```

Thereafter we display the content of the row which begins with the index **Aragorn**

```
display(data_pandas.loc['Aragorn'])
```

We can easily append data to this, for example

```
new_hobbit = {'First Name': ["Peregrin"],
              'Last Name': ["Took"],
              'Place of birth': ["Shire"],
              'Date of Birth T.A.': [2990]
              }
data_pandas=data_pandas.append(pd.DataFrame(new_hobbit, index=['Pippin']))
display(data_pandas)
```

Here are other examples where we use the **DataFrame** functionality to handle arrays, now with more interesting features for us, namely numbers. We set up a matrix of dimensionality 10×5 and compute the mean value and standard deviation of each column. Similarly, we can perform mathematial operations like squaring the matrix elements and many other operations.

```
import numpy as np
import pandas as pd
from IPython.display import display
np.random.seed(100)
# setting up a 10 x 5 matrix
rows = 10
cols = 5
a = np.random.randn(rows,cols)
df = pd.DataFrame(a)
display(df)
print(df.mean())
print(df.std())
display(df**2)
```

Thereafter we can select specific columns only and plot final results

```
df.columns = ['First', 'Second', 'Third', 'Fourth', 'Fifth']
df.index = np.arange(10)

display(df)
print(df['Second'].mean() )
print(df.info())
print(df.describe())

from pylab import plt, mpl
plt.style.use('seaborn')
mpl.rcParams['font.family'] = 'serif'

df.cumsum().plot(lw=2.0, figsize=(10,6))
plt.show()


df.plot.bar(figsize=(10,6), rot=15)
plt.show()
```

We can produce a 4×4 matrix

```
b = np.arange(16).reshape((4,4))
print(b)
df1 = pd.DataFrame(b)
print(df1)
```

and many other operations.

The **Series** class is another important class included in **pandas**. You can view it as a specialization of **DataFrame** but where we have just a single column of data. It shares many of the same features as $_{D}ataFrame.Aswith$**DataFrame**, $mostoperationsarevectorized, ac$

Friday: Introduction to Git and GitHub and getting started with numerical exercises