# MATHEMATICS FOR SCIENCE STUDENTS

**An open-source book**

Written, illustrated and typeset (mostly) by

### PELEG BAR SAPIR

with contributions from others

$$a^b = e^{b\log(a)}$$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$T\left(\alpha\vec{u} + \beta\vec{v}\right) = \alpha\,T\left(\vec{u}\right) + \beta\,T\left(\vec{v}\right)$$

$$A = Q\Lambda Q^{-1}$$

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t}\,dt$$

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

$$\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij}$$

$$\vec{v} = \sum_{i=1}^{n} \alpha_i \hat{e}_i$$

$$e^{\pi i} + 1 = 0$$

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$$

**PUBLISHED IN THE WILD**

! **To be written**: Rights, lefts, etc. will be written here in the future !

HERE BE TABLE

# CHAPTER

# 0



# INTRODUCTION

In this chapter we introduce key concepts that will be used in later chapters. For this reason, unlike other chapters it contains many statements, sometimes given without thorough explanations or reasoning. While all of these statements are grounded in deep ideas and can be formulated in a rigorous manner, it is advised to first get an intuitive understanding of the ideas before diving into their more formal construction.

> **Note 0.1   In case you are already familiar with the topics**
>
> It is recommended for readers who are familiar with the topics to at least gloss over this chapter and make sure they know and understand all the concepts presented here.

## 0.1   TAYLOR SERIES

### 0.1.1   Introduction

Say we wish to calculate the value of $\cos(x)$ at some non-trivial value, e.g. $x = \frac{\pi}{7}$, or $x = 1.0423$. Of course today we can simply use a computer or a calculator - but how do these even calculate such values? Moreover, a lot of these values are non-algebraic, meaning that we can't express them as decimal fractions to an infinite precision. What if for some important calculation we need to know the value of $e^{3.153}$ down to 15 digits, but our calculator only shows 10 digits after the period?

We can answer all these questions by using **approximations** instead of precise values. For example, we know values such as $\sin\left(\frac{\pi}{2}\right)$ with infinite precision (it's exactly 1), but in the case of less trivial values such as the ones discussed above, we can instead settle for an approximation, as long as it is *close enough* to the actual value for our needs. By "close enough" we essentially mean *within some given error range*, e.g. in the case of $e^{3.153}$ above, 15 digits of precision means that we want that the value we get is no more than $10^{-15}$ away from the actual value. Let us write this mathematically: if we denote $a$ as our approximation of $e^{3.153}$, then

$$|a - e^{3.153}| \leq 10^{-15}.$$

More generally, given a value $y$ and and we want to know it to within an **error range** $\Delta$, then for an approximation $a$ to be acceptable, the following must be true:

$$|a - y| \leq \Delta. \tag{0.1.1}$$

A good method to approximate functions would allow us to get as precise as we wish, given that we put enough "work" into finding an approximated value, e.g. we could use it to get an approximation of $e^{3.153}$ up to $\Delta = 10^{-15}$, but we could also use it to get an approximation up to $\Delta = 10^{-20}$ or $\Delta = 10^{-30}$ or any other value - we will simply have to do more calculations to reach such low error ranges. Generaly, the more precise we want our approximation to be, the more calculation we would need to carry out.

Unlike most real functions, polynomials are relatively easy to calculate for any real value $x$, since we simply need to perform the following operations: addition, subtraction, muliplication and raising by an integer power, all operations that are easy for both humans and computers to perform. Ideally, we would like to use polynomials to approximate all functions, e.g. given the function $f(x) = \cos(x)$ it would be great if we could find some polynomial $P(x)$ of a finite order $n$ for which $P(x) = \cos(x)$. Unfortunately such a polynomial does not exist, nor does such polynomials exist for $\sin(x), \sqrt{x}, \exp(x), a^x$ $(a > 0)$ or any other of the so-called elemntary functions and their compositions (except, of course, polynomials). The reason for this unfortunate reality is rather complicated, but in short it lies in the fact that these functions are *non-algebric*[1].

---

[1] technically $\sqrt{x}$ is algebraic, and there are indeed many methods to calculate its values - but non of these methods are as simple as calculating a polynomial... except the one we discuss in this section.

However, as mentioned before, for each of these functions, we know at least some values with inifinite precision. For example, we know that $\cos(0) = 1$ and $\cos\left(\frac{1}{2}\pi\right) = 0$. From symmetry we thus know that $\cos(\pi) = -1$ and $\cos\left(\frac{3}{4}\pi\right) = 0$. Since $\cos(x)$ is periodic, i.e. $\cos(x + 2\pi k) = \cos(x)$ for any $k \in \mathbb{Z}$, we actually know infinitely many values of the function. For $\exp(x)$ we know that $\exp(0) = 1$. Table 0.1 lists some known values of common elemntary functions.

**Table 0.1** Some known values of common elemntary functions. Each such value is known to an infinite precision.

| Function | $x$ | $f(x)$ |
|---|---|---|
| $\sqrt{x}$ | 4 | 2 |
| $\sin(x)$ | 0 | 0 |
| $\cos(x)$ | 0 | 1 |
| $\tan(x)$ | 0 | 0 |
| $\exp(x)$ | 0 | 1 |
| $\log(x)$ | 1 | 0 |

We can use this knowledge to constract a polynomial with $n$ terms, which approximates the function's value for any $x \in \mathbb{R}$ to whatever precision we wish, given that a specific condition is met. We will describe this condition later, but first let us use an example function to construct such a polynomial: $\exp(x)$. Since we only know $\exp(0)$ with infinite precision, we can use it as a simple approximation of $\exp(x)$, which we denote $T_0(x)$:

$$T_0(x) = 1. \tag{0.1.2}$$

(see Figure 0.1)

We call $T_0(x)$ the **0th-order approximation** of $\exp(x)$. Obviously, this is not a particularly good approximation: it's perhaps ok-ish for values *really* close to $x = 0$, but rapidly diverges from the actual value of $\exp(x)$ as we change $x$. More precisely, the error $\Delta_0(x)$ is

$$\Delta_0(x) = |\exp(x) - T_0(x)| = |\exp(x) - 1|, \tag{0.1.3}$$

i.e. the error essentialy grows as $\exp(x)$ which is... not great.

Ok, then perhaps we can add another term to the approximation? After all, our notation pretty much suggests that there are more terms we can use[2]. We know that close to a point $a$ a differential function $f(x)$ behaves pretty much like its *derivative* at that point, $f'(a)$. We also know the value of the derivative of $\exp(x)$ at $x = 0$, i.e. $\exp'(0) = 1$, so we can add a line of slope $m = 1$ to our approximation, yielding:

$$\exp(x) = 1 + x \cdot \exp'(0) = 1 + x. \tag{0.1.4}$$

Unsurprisingly, we call $T_1(x)$ the **1st-order approximation** of $\exp(x)$. Looking at Figure 0.2 we see that this approximation is better than the previous one, at least for values
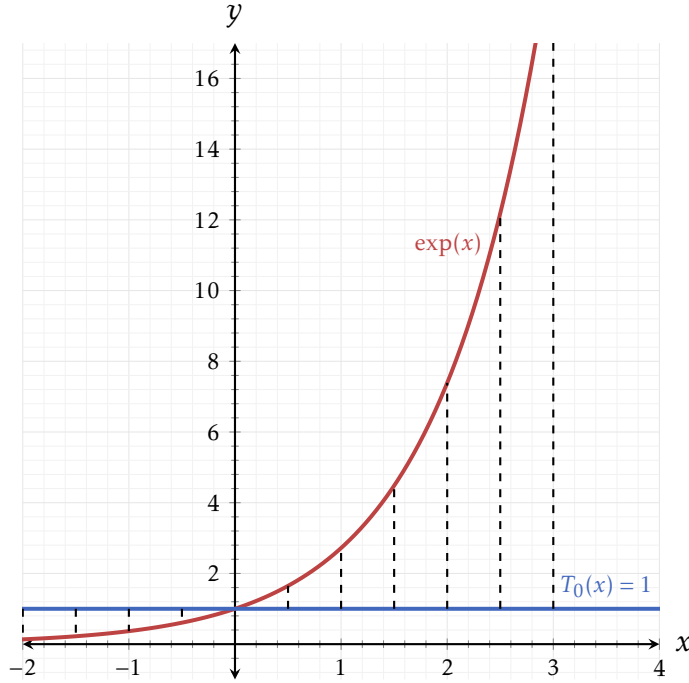
---

[2]more like *screams* it.

**Figure 0.1** The 0th-order approximation $\exp(x) = 1$. The dashed lines show $\Delta_0(x)$ as $x$ diverges from 0.

near $x = 0$. In other words, the error $\Delta_1$ behaves a bit better than $\Delta_0$:

$$\Delta_1(x) = |\exp(x) - T_1(x)| = |\exp(x) - x - 1|. \tag{0.1.5}$$

Similarly to $\Delta_0(x)$, $\Delta_1(x)$ also generally grows as $\exp(x)$. However, closer to $x = 0$ it behaves like $\exp(x) - x$, so still not great - but definitely an improvement over $\Delta_0(x)$. In practical terms, this means that by using $T_1(x)$ we would generally get better approximations than $T_0(x)$, at least close to $x = 0$. But we can do better!

The next step is of course to add an $x^2$ term. The second derivative of $\exp(x)$ is also $\exp(x)$, and at $x = 0$ also equal to 1. This time we will divide the term by 2 (for know just accept it as is, it would be explained when we formalize the method). Altogether, we get

$$T_2(x) = 1 + x + \frac{x^2}{2}. \tag{0.1.6}$$

(see Figure 0.3)

Definitely an improvement, but why stop here? We know all the derivatives of $\exp(x)$ at $x = 0$: they are all 1, since $\exp'(x) = \exp(x)$. We can continue adding power terms, yielding the following general approximation:

$$T_n(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots + \frac{x^n}{n!} = \sum_{k=0}^{n} \frac{x^k}{k!}. \tag{0.1.7}$$
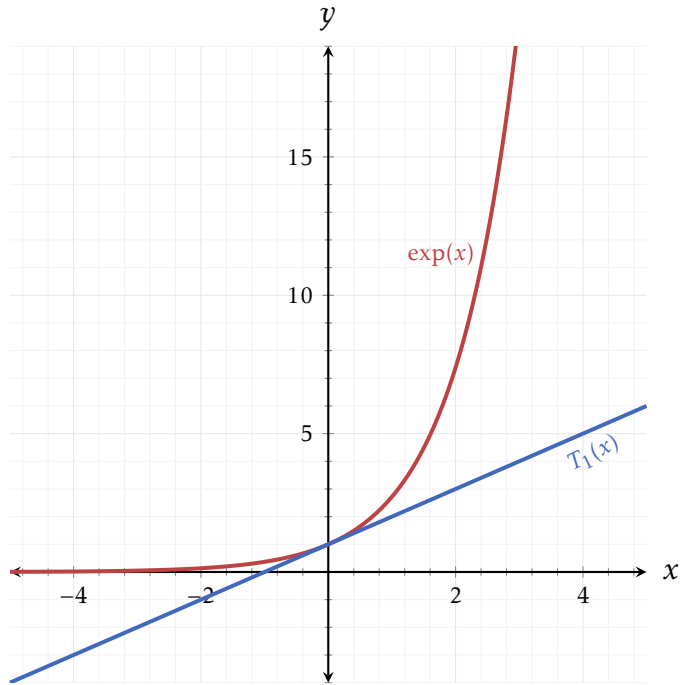
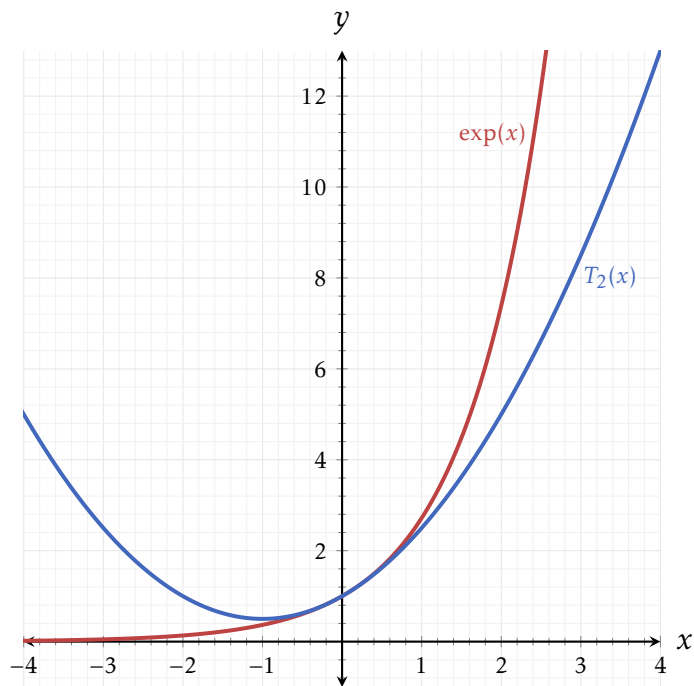**Figure 0.2** The 1st-order approximation $T_1(x) = 1 + x$.



**Figure 0.3** The 2nd-order approximation $T_2(x) = 1 + x + \frac{x^2}{2}$.

(again, for now you should just accept the coefficients $\frac{1}{n!}$, they would be explained soon)

The higher $n$ is, the more term are in the approximation and it gets more precise (see **??**). Of course this means that we need to carry out more calculations, as suggested earlier.

This kind of approximation, where we use a polynomial to approximate a real function, is called a **Taylor series** (sometimes also **Taylor expansion**). To calculate the Taylor series of a real function $f(x)$, we must first select a value of $a \in \mathbb{R}$ for which we know $f(a)$ with infinite precision (e.g. $x = 0$ in the case of $\exp(x)$). The function $f$ must also be infinitely differentiable at $a$, i.e. the $k$-th order derivative $f^{(k)}(a)$ must exist for any $k \in \mathbb{N}$, and we must know its value with inifinite precision. If these conditions are met, then the Taylor series of order $n$ of $f(x)$ at $x = a$ is given by

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k. \tag{0.1.8}$$

We say that a Taylor series as described above is *expanded about* the point $x = a$.

> **Note 0.2 Maclaurin series**
>
> A Taylor series with $a = 0$, i.e. where the series is expanded about $x = 0$, is called a **Maclaurin series**.
>
> !

> **Example 0.1 Taylor exansion of $\exp(x)$**
>
> Let us verify that using Equation 0.1.8 indeed yields Equation 0.1.7: since we're using $a = 0$, we need to know the value $\exp^{(k)}(0)$ for any $k \in \mathbb{N}$. This is pretty simple: the $k$-th derivative of $\exp(x)$ is $\exp(x)$, and therefore $\exp^{(k)}(0) = 1$ for any $k \in \mathbb{N}$. Substituting $a = 0$ and $f^{(k)}(x) = 1$ to Equation 0.1.8 gives back Equation 0.1.7.
>
> ★

> **Example 0.2 Taylor series of $\sin(x)$ and $\cos(x)$**
>
> Now let us calculate the Taylor series of $\sin(x)$ and $\cos(x)$, denoting them as $T_n^s(x)$ and $T_n^c(x)$, respectively. For both the functions we can use $a = 0$ since we know the values of the functions at these points: $\sin(0) = 0$, $\cos(0) = 1$. The $k$-th order derivative of $\sin(x)$ depends on the value of $k$ modulo 4, i.e. the derivatives are a repeating sequence of 4 element:
>
> $$\xrightarrow{\phantom{aa}} \sin(x) \xrightarrow{\frac{d}{dx}} \cos(x) \xrightarrow{\frac{d}{dx}} -\sin(x) \xrightarrow{\frac{d}{dx}} -\cos(x) \xrightarrow{\phantom{aa}}$$
> $$\underbrace{\phantom{aaaaaaaaaaaaaaaaaaaaaaaaaaa}}_{\frac{d}{dx}}$$
>
> Since each **even** order derivative of $\sin(x)$ is $\pm\sin(x)$, and $\pm\sin(0) = 0$, we can ignore these terms, since they will vanish from the sum. We therefore need to "run" the sum only for **odd** values, which we can write as $m = 2k+1$. It's important to notice that the values of the derivatives "jump" between positive and negative valiues, i.e. $+\cos(0), -\cos(0), +\cos(0), -\cos(0), \cdots = +1, -1, +1, -1, \ldots$. We therefore

set the derivative term to be $(-1)^k$. Altogether we get

$$T_n^s(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \pm \frac{x^{2n+1}}{(2n+1)!} = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1}.$$

Similarly, the Taylor series for $\cos(x)$ would have only half of its terms non-zero, but the parity is opposite in comparison to $\sin(x)$: this time, the odd terms disappear, and we're left with only the even terms. Therefore we use $m = 2k$, and altogether get

$$T_n^c(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \pm \frac{x^{2n}}{(2n)!} = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k}.$$

! **To be written**: When reaching the proof that in the limit $n \to \infty$ the Taylor series equals the functio
!

## 0.1.2   Remainder

Recall that one of the requirements we put on an approximation of a function $f(x)$ is that it can get *as close* to $f(x)$ as we want, for any point $x = a$ on some interval $(b, c)$ where $b < c \in \mathbb{R}$. This means that given an approximation $T_n(x)$ we can write the function to be approximated as

$$f(x) = T_n(x) + R_n(x), \tag{0.1.9}$$

i.e. that the function to be approximated can be written as the some of out approximation $T_n(x)$, and a **remainder** which we can make as small as we want. "As small as we want" means that

$$\lim_{n \to \infty} R_n(x) = 0, \tag{0.1.10}$$

or writing this in another way

$$\lim_{n \to \infty} T_n(x) = f(x). \tag{0.1.11}$$

The last equation shows exactly what we expect of $T_n(x)$: as we increase $n$ the approximation gets closer and closer to the function $f(x)$ which we are approximating.

In essence, Taylor series do exactly that for functions which are infinitely differentiable on some interval: at the limit $n \to \infty$ they give back the function itself (Equation 0.1.11). This is called **Taylor's theorem**.

> **Note 0.3   No proof**
>
> We won't prove Taylor's theorem here since it requires some more formal calculus. Instead, we would just take as granted that it is indeed true.
>
> !

The convergence of the Taylor series allows us to prove some theorems about several functions, one of which we will use soon. For now though we concentrate on the use

of Taylor series as approximations. Using approximations is always a balance between getting a more precise approximation on one hand, and not doing too much work on the other hand. For example, if we want to use the Taylor series of $\exp(x)$ to calculate $\exp(3.153)$ up to $10^{-15}$, we want to know the minimum value of $n$ such that the remainder $R_n(x)$ is smaller then or equal to $10^{-15}$. Any higher $n$ would also give us an approximation within that range, however it will require us to calculate more terms of $T_n(x)$. Therefore, we should find a way to estimate $R_n(x)$ as a function of $n$, and determine the minimum amount of $n$ needed for an approximation within the desired range.

### 0.1.3 Coefficients