Chapter 1

Fundamental concepts

1.1 Motivation

Differential equations are essential to most branches of mathematical physics and applied mathematics, as we live in a changing world. For example, it is an empirically observed fact, beloved of Physics interviewers asking questions about tea, milk, and mysterious absences, that the rate of change of the temperature of a body is proportional to the difference in temperature between the body and its surroundings. So, an obvious question is how does the body's temperature vary with time?

Let's set the problem up mathematically.

- Let us call the temperature of the body $\theta(t)$: "theta" for those without the benefits of a classical education. We say that θ is a **dependent** variable because it **depends** on (i.e. varies with) time t.
- We also say that θ is a **function** of its **argument** t, and we wish to determine the functional relationship between θ and t.
- We say that t is the **independent** variable.
- We also need to say what the temperature of the surroundings is, and for simplicity let us suppose that it is a constant θ_0.

Then we need a notation for "the rate of change", which we write as an (ordinary) derivative:

$$\frac{d}{dt}\theta \propto \theta - \theta_0 \to \frac{d}{dt}\theta = -k(\theta - \theta_0),$$

where $k > 0$ is a constant. The sign follows from the reasonable assumption that if $\theta > \theta_0$ we expect the body to cool down towards θ_0, while if $\theta < \theta_0$ we expect it to heat up.

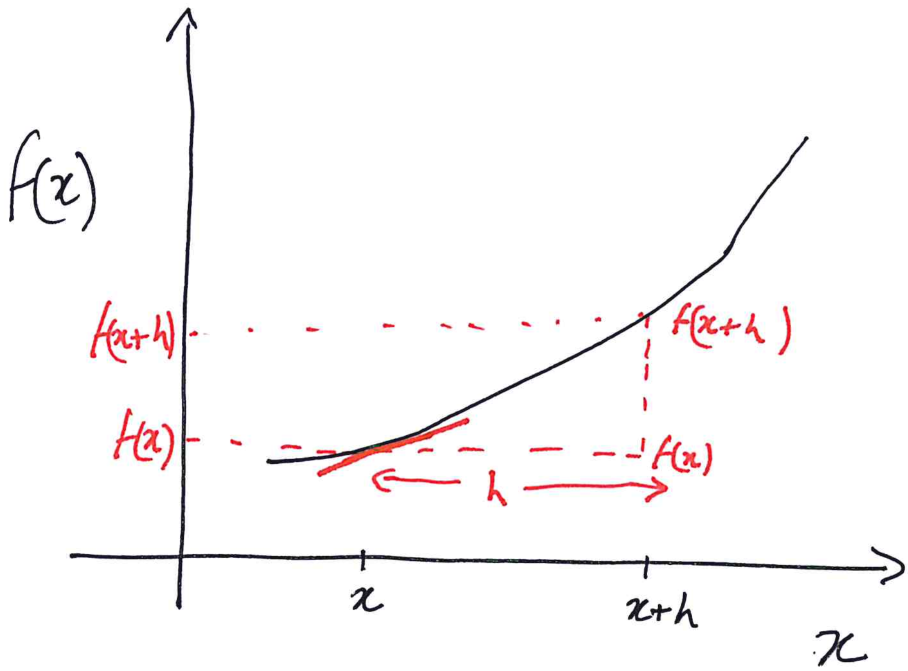

Figure 1.1: Schematic representation of the derivative of a function $f(x)$ at a point as the slope of the graph of the function at that point.

Now, how might we solve this **differential equation**? Well, from the name of the course, you hopefully have come to the right place. . . but there are a lot of open questions:

- how is a derivative defined?
- have we got enough information to solve the equation?
- how would we solve the equation?

These are the sort of questions which will be addressed in this course. More formal issues of existence/uniqueness and validity will be addressed in the Analysis courses. . . but there's still plenty in this course to keep us all busy!

1.2 Preliminary definitions

1.2.1 Definition of a derivative

We define the **derivative** of a **function** $f(x)$ with respect to x as the function defined by the **limit**:

$$\frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}. \tag{1.1}$$

We are not going to define functions formally in this course, but loosely they are mathematical expressions which take inputs and give well-defined outputs for those inputs. In particular, for a given input, the output is uniquely defined. So, if we know x, we can work out $f(x)$. Similarly, if we know x and $f(x)$ we can use the expression (1.1) to work out df/dx at x, assuming of course the limit exists.

Notationally, it is sometimes important to distinguish between the function, defined over all its possible input points x (i.e. over its **domain**) and the specific value of the function for a particular input value of $x = x_0$ say. Applying (1.1), we write

$$\left.\frac{df}{dx}\right|_{x=x_0} = \frac{df}{dx}(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h}. \tag{1.2}$$

As shown in the schematic figure 1.1, the value of df/dx at $x = x_0$ is the slope of the graph of $f(x)$ at the point $x = x_0$. It should always be clear by context or notation whether one is discussing the function, or its specific value for a given input argument.

1.2.2 Differentiability

For the function $f(x)$ to be **differentiable**, and so for the function df/dx to be well-defined at the point x_0, the left-hand limit (i.e. h is negative and approaches zero from below) and the right-hand limit (h is positive and approaches zero from above) must be defined and equal, i.e.

$$\lim_{h\to 0^-} \frac{f(x_0+h) - f(x_0)}{h} = \lim_{h\to 0^+} \frac{f(x_0+h) - f(x_0)}{h}.$$

This is actually quite a strong restriction on the "smoothness" of the function $f(x)$. As an example, $f(x) = |x|$ is not differentiable at $x = 0$. Deliberately in this course we do not define what is meant by a "function" or a "limit", but don't worry, these concepts will come up again...

Notation

Writing the derivative of a function $f(x)$ as df/dx is known as Leibniz notation. Notice how the denominator shows what the argument of the function is. Of course, other mathematicians had other notations, and two important ones are Lagrange's notation $f'(x)$ (as we shall see this notation often has the argument specifically quoted to avoid ambiguity) and Newton's notation of a superscript dot, $\dot{f}$, which is typically only used when the independent variable is time.

Furthermore, the definition of derivative can be applied recursively, defining derivatives of derivatives, i.e. second derivatives etc:

$$\frac{d}{dt}\left(\frac{df}{dt}\right) = \frac{d^2 f}{dt^2} = f''(t) = \ddot{f}.$$

1.3 Big O and little o notation

Two very useful concepts in applied mathematics are big O (pronounced "Oh") and little o, which is sometimes written $\underline{o}$ to distinguish clearly between the two symbols. These two concepts, sometimes called **order parameters**, are used to give comparative scalings between functions sufficiently close to some limiting point x_0. Conventionally, these concepts are used with an abuse of the equals notation, but let's all keep calm.

1. Definition "little oh": $f(x)$ is $o[g(x)]$ as $x \to x_0$ if

   $$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0,$$

   commonly written as $f(x) = o[g(x)]$.

2. Definition "big Oh": $F(x)$ is $O[G(x)]$ as $x \to x_0$ if:

   (a) for the case where x_0 is finite, there exist two positive finite constants M and δ such that for all x with $|x - x_0| < \delta$,

   $$|F(x)| \leq M|G(x)|;$$

   (b) for the case where x_0 is infinity, there exist two positive finite constants M and x_1 such that for all $x > x_1$,

   $$|F(x)| \leq M|G(x)|.$$

   Similarly, this is also commonly written as $F(x) = O[G(x)]$.

The equals notation is most commonly used, though really these concepts can be thought of as conditions for the functions f and F to belong to sets of functions with the required property of varying in a particular way as they approach the special limiting point $x = x_0$.

Note from the definitions, $f = o(g) \to f = O(g)$ but not vice versa. Very commonly, the functions $g(x)$ and $G(x)$ are powers of x but that is not necessary. Loosely:

1. If, in the definition "little oh", $g(x)$ tends towards infinity as $x \to x_0$, $f(x)$ is definitely growing more slowly.

2. If, in the definition "little oh", $g(x)$ tends towards zero as $x \to x_0$, $f(x)$ is definitely decaying to zero even more rapidly.

3. If, in the definition "big Oh", $G(x)$ tends towards infinity at some rate as $x \to x_0$, $F(x)$ can at most be going to infinity at a fixed multiple of that rate.

4. If, in the definition "big Oh", $G(x)$ tends towards zero at some rate as $x \to x_0$, $F(x)$ is definitely eventually decaying to zero at least as quickly as a fixed multiple of that rate.

Some examples:

- $x = o(\sqrt{x})$ as $x \to 0$.
- $x = O(\sqrt{x})$ as $x \to 0$.
- $\sin 2x = O(x)$ as $x \to 0$.
- $\sqrt{x} = o(x)$ as $x \to \infty$.
- $\cos(x) = O(1)$ for all x.

1.3.1 Tangent line at x_0

Armed with these parameters, we can now rewrite the (specific) derivative definition (1.2):

$$\left.\frac{df}{dx}\right|_{x_0} = \frac{f(x_0 + h) - f(x_0)}{h} + \frac{o(h)}{h}. \tag{1.3}$$

The equivalence of this definition and (1.2) can be proved by contradiction. Assume that

$$\left.\frac{df}{dx}\right|_{x_0} = \frac{f(x_0 + h) - f(x_0)}{h} + \frac{R(h)}{h},$$

for some remainder function $R(h)$ which is categorically **not** $o(h)$. Take limits as $h \to 0$ of both sides of the equation. The left-hand side is a constant which doesn't change in the limit. The first term on the right-hand side is by definition (1.2), and so is the same as the left-hand side. But by construction,

$$\lim_{h\to 0} \frac{R(h)}{h} \neq 0,$$

and we have a contradiction!

Multiplying (1.3) across by h, we obtain

$$f(x_0 + h) \quad = \quad f(x_0) + \left(\left. \frac{df}{dx} \right|_{x_0} \right) h + o(h). \tag{1.4}$$

If we now identify $x = x_0 + h$, $y(x) = f(x)$, $y_0 = f(x_0)$, (1.4) can be related to the equation of the tangent line at x_0 to the curve $y = f(x)$, i.e.

$$y = y_0 + m(x - x_0),$$

where the slope $m = df/dx$ at the point $x = x_0$.