Calculus: Exploration Sheet

Suvrat Raju*

Maths Circle, 8 November 2025

1 Introduction

In this session, we will explore Calculus. This is a vast and important subject. It is not an overstatement to say that, without Calculus, there would be no modern physics! We cannot do justice to the entire subject in two sessions. So the objective of this exploration sheet is to introduce you to some simple ideas. We will take a perspective that is different from what the senior students might have seen in their textbooks, although this perspective has started to become common in some modern computational fields.

This week, we will explore a topic called "differentiation". In the next session, we will explore "integration".

You can either start by doing the problems below and then read the discussion in the next section. But, if you prefer, you can start with the discussion and return to the problems below.

2 Differentiation

2.1 Dual Numbers

We will start by introducing a new symbol, ϵ. We will establish the following rules for dealing with this symbol.

1. It is possible to multiply ϵ with ordinary numbers and also add it to ordinary numbers. So we can study expressions like
$$2 + 3.5\epsilon,$$
or
$$\frac{7}{6} + \sqrt{2}\epsilon.$$
2. We can also multiply expressions of the form above using the usual rules of multiplication. But whenever we encounter ϵ^2, we set it to zero.

You can think of ϵ as a special number with the property that $\epsilon^2 = 0$ but $\epsilon \neq 0$.

If this seems strange to you, we can formalize the notion of ϵ by extending the ordinary numbers to ordered pairs
$$(a, b) \equiv a + b\epsilon. \tag{1}$$

*Send comments or corrections to suvrat@icts.res.in

Addition on these ordered pairs, and multiplication by ordinary numbers is defined by

$$(a, b) + (c, d) = (a + c, b + d), \tag{2}$$

and

$$z \times (a, b) = (za, zb), \tag{3}$$

where z is an ordinary number. Multiplication of two of these ordered pairs is defined by

$$(a, b) \cdot (c, d) = (ac, ad + bc). \tag{4}$$

You should check that (2), (3), (4) are equivalent to the two rules we specified for dealing with ϵ above.

P1 Your first problem is to work out the rules for division. If e, f are defined through

$$\frac{a + b\epsilon}{c + d\epsilon} = e + f\epsilon, \tag{5}$$

find e, f in terms of a, b, c, d. When is the ratio undefined? This relation can also be written as

$$\frac{(a, b)}{(c, d)} = (e, f), \tag{6}$$

in terms of the e, f that you found, which supplements (4), (3) and (2).

Addition, subtraction, multiplication and division are the basic arithmetic operations. Since we have extended them, any function that is defined for ordinary numbers in terms of these elementary operations is defined for this extended class of numbers.

2.2 Derivatives

Let $f(x)$ be some function that is obtained by composing the elementary operations of addition, subtraction, multiplication and division.

P2 Argue that for any function f, we must have

$$f(x + \epsilon) = f(x) + g(x)\epsilon, \tag{7}$$

for some function $g(x)$ that depends on $f(x)$.

The function $g(x)$, defined via the relation (7), is called the *derivative* of $f(x)$. It is denoted by

$$g(x) = \frac{df(x)}{dx}. \tag{8}$$

You might also see

$$g(x) = f'(x). \tag{9}$$

For those of you who have seen the derivative before, you can consult the next section for an intuitive explanation of why this corresponds to the usual definition of the derivative. Others can simply take this to be definition of the derivative.

Let us work out this derivative in a few cases.

P3 Let

$$f(x) = x^n, \tag{10}$$

where n is some natural number. Evaluate $f(x+\epsilon)$ using the binomial theorem or using induction and, therefore, find the expression for $f'(x)$. Next, let

$$f(x) = \frac{1}{x^n}, \tag{11}$$

where n is again a natural number. By evaluating $f(x+\epsilon)$, find the expression for $f'(x)$.

If you can think of other functions that you can define by combining the elementary arithmetic operations, you can go ahead and evaluate derivatives for those functions using the method above.

If we know the derivatives of two functions, and a new function is defined by adding, multiplying or composing those functions, there are standard rules that tell us how to compute the derivative of the new function. The next problem asks you to work out these rules.

P3 Let $f(x)$ and $g(x)$ be two functions. Find the expressions for the derivative of $h(x)$ in each of the following cases.

1. Sum of functions.
$$h(x) = f(x) + g(x). \tag{12}$$
2. Product of functions.
$$h(x) = f(x)g(x). \tag{13}$$
The rule for this case is called the "product rule" for differentiation.
3. Ratio of functions.
$$h(x) = \frac{f(x)}{g(x)}. \tag{14}$$
Also specialize the answer to the case where $f(x) = 1$.
4. Composition of functions.
$$h(x) = f(g(x)). \tag{15}$$
The rule for this case is called the "chain rule" for differentiation.
5. Finally, we consider inverse functions. This means that $h(x)$ is defined by the relation
$$h(f(x)) = x. \tag{16}$$
As a special case, compute the derivative of $\sqrt{x}$ using this method.

Especially in computation, a function is often specified via an algorithm that gives more and more accurate approximations to the function. For example, here is an algorithm to find the square root of a number $x > 0$. Start with a guess $y_0 > 0$ and iteratively define

$$y_{n+1} = \frac{1}{2}(y_n + \frac{x}{y_n}). \tag{17}$$

This algorithm is an example of a numerical technique called "bisection". For any y_n, we know that square root must be between y_n and $\frac{x}{y_n}$. In each successive approximation, we cut the size of this interval by half and eventually "squeeze" the square root in a small-enough range.

The definition of the derivative given above is especially useful in such cases since we can calculate the function and its derivative at the same time.

P4 Run the square-root algorithm with $x = 2 + \epsilon$ and $y_0 = 1$ for 3 steps, i.e. compute y_3. (You can do more if you want!) Feel free to use a calculator. You should get an excellent approximation to $\sqrt{2}$, and also to the derivative of $\sqrt{x}$ at $x = 2$.

3 Discussion

Some of you might have already studied calculus in school or elsewhere. Physically, the derivative of a function measures its "rate of change". Consider the graph of a function shown in Figure 1. At point x_0, the function is changing rapidly whereas at point x_1 it more-or-less flat. The derivative makes this precise: if one draws a tangent to the curve at a point, the derivative is the slope of the tangent.

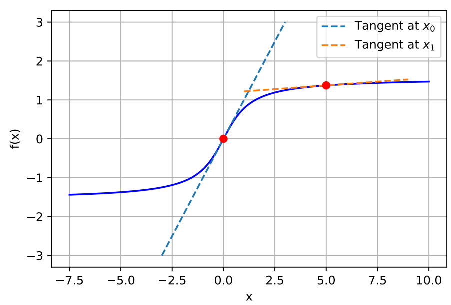
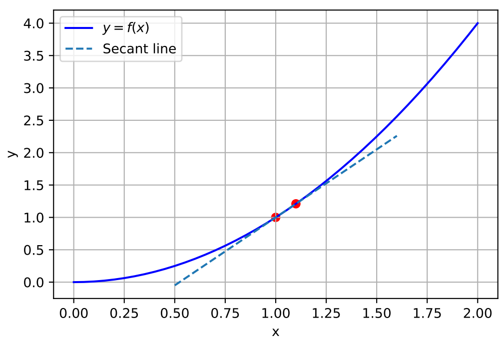


Figure 1: *A function $f(x)$ with tangents marked at two points, x_0 and x_1. The derivative is large in magnitude at x_0 and small in magnitude at x_1.*

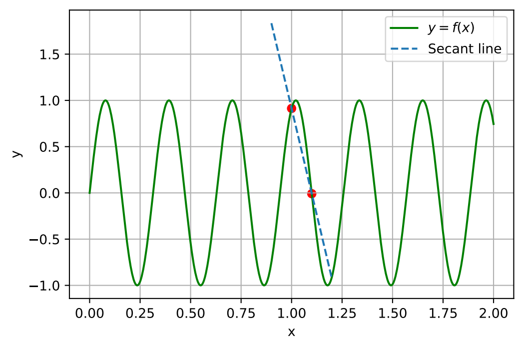
If we are allowed to numerically evaluate the function, then the derivative can be approximated via

$$f'(x) \approx \frac{f(x+\delta) - f(x)}{\delta}, \tag{18}$$

which gives the slope of a secant that joins points on the graph at x and $x + \delta$. This is a good approximation to the slope of the tangent if δ is "small enough". But what does "small enough" mean? Figure 2 shows a case where $\delta = 0.1$ is a good approximation, and another case where it is not.



(a) *For this function, the secant joining the points at $x = 1$ and $x = 1.1$ gives a reasonable approximation to the tangent at $x = 1$.*



(b) *For this function, the secant joining the points at $x = 1$ and $x = 1.1$ is a very poor approximation to the tangent at $x = 1$: the secant has negative slope although the tangent has a positive slope.*

Figure 2: *In both cases above, we attempt to evaluate the derivative at $x = 1$ by drawing the secant that connects points at $x = 1$ and $x = 1.1$. We obtain an accurate result in one case, and a poor result in the other.*

We would like to take δ to be infinitesimally small to avoid the problems that appear in Figure 2b, and to get as good an approximation to the slope of the tangent as possible. One way to define this formally is via a limiting procedure where we take δ to be smaller-and-smaller to get successively better-and-better approximations to the derivative.

On the other hand, this limiting procedure is hard for computers to implement. When δ is very small, (18) involves the ratio of the difference of two numbers that are very close, with a third very small number. Computers operate with finite precision, and can make errors in computing such ratios.

The method of dual numbers introduced above is an alternate way of formalizing the notion of a derivative. We can think of ϵ, defined via the rules of section 2.1, as a precise version of our notion of an infinitesimal quantity. It makes precise the idea that we need to keep track of first-order terms in the infinitesimal, but second order terms can be neglected. Such dual numbers are used in modern computational work — especially in the fields of artificial intelligence and machine learning — to perform “automatic differentiation”. Often, a function and its derivative are evaluated at the same time, just as you did in the example above.