Chapter 4

The Derivative

The earlier chapters are the analytical preludes to calculus. This chapter begins the study calculus proper, starting with the study of *differential calculus*, also known as the calculus of derivatives. We will develop all of the fundamental derivative computations in this chapter.

Once we complete our development of derivatives and see many of their most immediate applications in this chapter and in Chapter 5, we will then look towards *integral calculus,* where roughly speaking we see how to reverse what we do here.[1] Integral calculus begins with Chapter 6, with the advanced computational techniques being introduced in Chapter 7. Subsequent chapters develop several topics which are either offshoots of differential and integral calculus, or are greatly extended by these.

As we will see, differential calculus addresses rates at which quantities change with resepct to each other, while integral calculus addresses how quantities (or the changes in quantities) accumulate. Once we lay the foundations of both differential and integral calculus, we will further develop and apply both in very diverse circumstances for the remainder of the text.

4.1 The Derivative: Rates of Change, Velocity and Slope

Suppose that we are passengers in a car driving west to east along a highway. Further suppose that we cannot see the speedometer (measuring speed) but the highway is marked at regular intervals so we can measure our position accurately.[2] Using a stopwatch, we see that we traveled a total of 130 miles in 2 hours for the whole trip. Then we would say our average velocity (with positive measured in the eastward direction) was

$$\frac{130 \text{ mi}}{2 \text{ hr}} = 65 \text{ mi/hr}.$$

Now suppose that during the trip we would like to know our actual velocity at a particular time t_1. The average for the whole trip does not usually reflect the velocity at any particular time t_1 with acceptable accuracy, since we could have been stopped for a break at that particular time, or speeding up to pass a truck, or even driving in reverse (for a negative velocity). One way to attempt to approximate the velocity at time t_1 is to begin our stopwatch at t_1, measure how far we traveled in the next minute, and calculate the average velocity for the time interval $t \in [t_1, t_1 + 1 \text{ minute}]$.

[1] Here we take a function and find its derivative. Later we take the derivative and determine what was the underlying function. This reverse process is often a more formidable task, as we will see in later chapters. However a thorough understanding of the material in this chapter greatly simplifies the learning of integral calculus.

[2] Alternatively, we have a very accurate odometer or global positioning device in plain view.

At this point some notation will be useful. We will take the position at time t to be $s(t)$. It is a function, which we will call the *position function*; its input is time t and its output is our position at time t. We will denote a change in t by Δt, read "delta t."[3] With t_1 as the initial time in our experiment to approximate velocity, and $t_2 = t_1 + \Delta t$ as the final time, we see the change is indeed $t_2 - t_1 = \Delta t$. The *average velocity* over any time interval $[t_1, t_2]$ is thus

$$\frac{s(t_2) - s(t_1)}{t_2 - t_1} = \frac{s(t_1 + \Delta t) - s(t_1)}{\Delta t}.$$

If we take $\Delta s = s(t_1 + \Delta t) - s(t_1)$ to be the change in $s(t)$ which results from the change in t from t_1 to $t_1 + \Delta t = t_2$, then we have the average velocity also equal to $(\Delta s)/(\Delta t)$, i.e.,

$$\frac{s(t_2) - s(t_1)}{t_2 - t_1} = \frac{s(t_1 + \Delta t) - s(t_1)}{\Delta t} = \frac{\Delta s}{\Delta t}.$$

This is akin to the old-fashioned "rate equals distance divided by time" that is taught in grade school.[4] With this we can get back to the problem of attempting to find the velocity at time t_1. If we let Δt be equal to one minute, then we look to see how far we traveled in that one minute, and find the average velocity for that minute. *If the velocity did not change very much in that time interval, then the average velocity will closely approximate the actual velocity*, which we will denote $v(t_1)$, where $v(t)$ is the actual velocity at time t (what we would have read on the speedometer—except for a possible sign difference—were it available):

$$\frac{\Delta s}{\Delta t} = \frac{s(t_1 + 1\text{ minute}) - s(t_1)}{1\text{ minute}} \approx v(t_1).$$

On the other hand, many things can happen in a minute which can cause the velocity to change significantly. Perhaps we have a true velocity of 65 miles/hour at t_1, but then slow to a stop at a toll booth during that minute, and thus unacceptably underestimate $v(t_1)$ as approximated by the average velocity for the time interval $[t_1, t_2]$. If possible, it would likely be much better to measure how far we traveled in the first *second* after t_1, since most cars cannot change velocity as significantly in such a time interval except in catastrophic circumstances (e.g., collisions). Thus[5]

$$v(t_1) \approx \frac{s(t_1 + 1\text{ second}) - s(t_1)}{1\text{ second}}.$$

Following the same line of thinking, it seems reasonable that we can better approximate the actual value of $v(t_1)$ by taking the average velocity over an interval $[t_1, t_1 + \Delta t]$ with smaller and smaller values of Δt (such as one minute, one second, 0.001 seconds, etc.). For this reason we actually *define* the velocity at time t_1 by the following, 0/0-form limit:

$$v(t_1) = \lim_{\Delta t \to 0} \frac{s(t_1 + \Delta t) - s(t_1)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}.$$

Recall that we have to consider $\Delta t \to 0^-$ as well as $\Delta t \to 0^+$ in this calculation. This is not unreasonable, as we could also approximate $v(t_1)$ by considering how far we went in the minute, second, 0.001 second, etc., ending with t_1. Now we state the formal definition of velocity.

[3] Note that we take Δt as one quantity. It is *not* "Δ times t." One can read Δt to be synonymous with the *change* in t. Occasionally we will write $\Delta t = (\Delta t)$ to remove ambiguity and reinforce that it is one quantity. (Here Δ is the capital Greek letter delta.)

[4] The grade school formula is lacking in that it always assumes velocity is constant, and does not distinguish between "distance" and "displacement," or "distance" and "position." (Distance only carries a nonnegative sign.) It is only mentioned here because of its familiarity.

[5] Of course we need to convert units to be consistent, e.g., 1 second = (1/3600) hour, and so on.

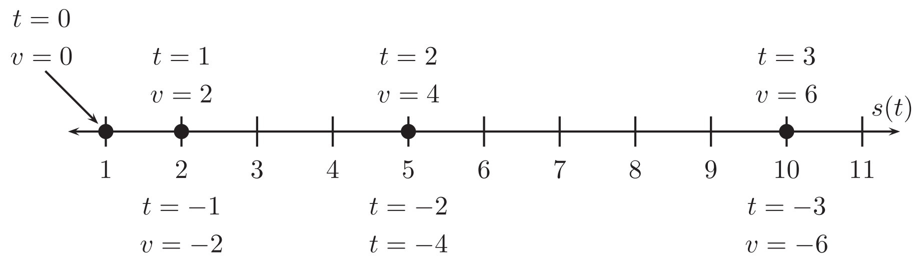

Figure 4.1: Here we trace the one-dimensional motion $s(t) = t^2 + 1$ as a position on a number line for times $t = -3, -2, -1, 0, 1, 2, 3$ (note the progression of those positions in the figure). The velocities $v = 2t$ are also given. The graph reflects how the particle comes in from the right for negative t, stops at $t = 0$ ($s = 1$, $v = 0$), and moves back out towards the right for positive t (faster and faster as $t > 0$ increases). Note that no "t-axis" appears explicitly.

Definition 4.1.1 *Given a position function $s(t)$, define the* **velocity** *(or* **instantaneous velocity**, *to distinguish it from average velocity) at a time t to be the function given by the limit*

$$v(t) = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \lim_{\Delta t \to 0} \frac{s(t + \Delta t) - s(t)}{\Delta t}, \tag{4.1}$$

for each t for which the limit (4.1) *exists, and where we also define*

$$\Delta s = s(t + \Delta t) - s(t). \tag{4.2}$$

For now it is the second part of (4.1) that will be most useful. If we are lucky enough to know an algebraic formula for $s(t)$ as a function, then we can use the limit to calculate $v(t)$. This describes a situation known to physicists as *one-dimensional motion.* Note how (4.1) allows for $\Delta t \to 0^-$ as well as $\Delta t \to 0^+$. Note also that for continuous $s(t)$—a reasonable assumption in classical physics—limits of form (4.1) will be of 0/0 form.

Example 4.1.1 *Suppose that position is given by $s(t) = t^2 + 1$. We can use* (4.1) *to calculate the velocity function for any fixed t as follows. As this limit will be a 0/0 form, we perform algebra to attempt to can cancel the Δt factor in the denominator.*[6]

$$\begin{aligned}
v(t) &= \lim_{\Delta t \to 0} \frac{s(t + \Delta t) - s(t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\big((t + \Delta t)^2 + 1\big) - \big(t^2 + 1\big)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{(t^2 + 2t\Delta t + (\Delta t)^2 + 1) - (t^2 + 1)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{t^2 + 2t\Delta t + (\Delta t)^2 + 1 - t^2 - 1}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{2t\Delta t + (\Delta t)^2}{\Delta t} \\
&= \lim_{\Delta t \to 0} (2t + \Delta t) \\
&= 2t.
\end{aligned}$$

We showed that $s(t) = t^2 + 1 \implies v(t) = 2t$. Some position and velocity data are given for various times t in Figure 4.1. *Note that when $s > 0$ the position is to the right of $s = 0$ (as is always the case here), and when $s < 0$ position is to the left. Also, when $v > 0$ the motion is to the right, and when $v < 0$ the motion is to the left. For example, at time $t = 2$ we have the position $s(2) = 2^2 + 1 = 5$, and velocity $v(2) = 2(2) = 4$. If s is measured in meters, and t in seconds, then the*

[6] Here we treat t as a constant in the calculation of $v(t)$; t is fixed while $\Delta t \to 0$.

units in the limit (4.1) *are meters/second, as we would hope.*

The ability to find a nonconstant velocity function is a tremendous leap from the grade school notion of "rate = distance/time." Having limits at our disposal made it possible. Note that $v(t)$ is really a *limit* of an expression of the form "position change/time change," i.e., $(\Delta s)/(\Delta t)$ so it has some of the spirit of the grade school notion.

Such limits are useful in more than just "position $\Longrightarrow$ velocity" problems; we will have use for them throughout the text in numerous contexts. Because they are ubiquitous we generalize the notation and call the functions which arise from these limits *derivatives*. (The following definition should be commited to memory.)

Definition 4.1.2 *Given any quantity Q which is a function of the variable x, i.e., $Q = Q(x)$.*

- *The* **derivative** *of Q* **with respect to** *x is the function $Q'(x)$, read* **"Q-prime of x,"** *defined by*
$$Q'(x) = \lim_{\Delta x \to 0} \frac{Q(x + \Delta x) - Q(x)}{\Delta x} \tag{4.3}$$
wherever that limit exists and is finite.

- *If this limit does not exist or is infinite at a given x_0, we say $Q'(x_0)$* **does not exist**. *If the limit does exist as a finite number at $x = x_0$, we say $Q(x)$ is* **differentiable** *at x_0.*

- *$Q'(x)$ is also called the* **instantaneous rate of change of $Q(x)$ with respect to** *x.*[7]

To define the derivative $Q'(x)$ at a given value x, we require not only that the limit (4.3) exists, but also that it is finite (i.e., exists as a real number). We will make more use of the term *differentiable* in later sections where its justification is clearer.

Note that in most cases we expect the limit (4.3) which defines the derivative to be of 0/0 form, requiring the usual techniques of algebraic simplification to compute.

Definition 4.1.3 *We also define the* average rate of change *over an interval as before: if the initial value of x is x_0 (pronounced "x-naught" or "x sub(script) zero"), and the final value is x_f, then the* **average rate of change of $Q(x)$ with respect to** *x for $x \in [x_0, x_f]$ or $x \in [x_f, x_0]$ (depending upon whether $x_0 < x_f$ or $x_0 > x_f$) is given by the* **difference quotient**
$$\frac{Q(x_f) - Q(x_0)}{x_f - x_0} = \frac{Q(x_0 + \Delta x) - Q(x_0)}{\Delta x} = \frac{\Delta Q}{\Delta x} \tag{4.4}$$
where
$$\Delta x = x_f - x_0, \tag{4.5}$$
$$\Delta Q = Q(x_f) - Q(x_0) = Q(x_0 + \Delta x) - Q(x_0). \tag{4.6}$$

So we see that the derivative (4.3) is just the limit of the average rate of change in Q given in (4.4) on an interval with endpoints x and $x + \Delta x$, assuming that limit as $\Delta x \to 0$ is finite.

With this notation, we can rewrite the (instantaneous) velocity function for a given $s(t)$ as:
$$v(t) = s'(t). \tag{4.7}$$

[7] "Instantaneous" rate of change means the rate of change "at that instant," as opposed to an average rate of change of the output variable which occurs over an entire interval's length of values of the input variable.