

## 1.1: Background Material

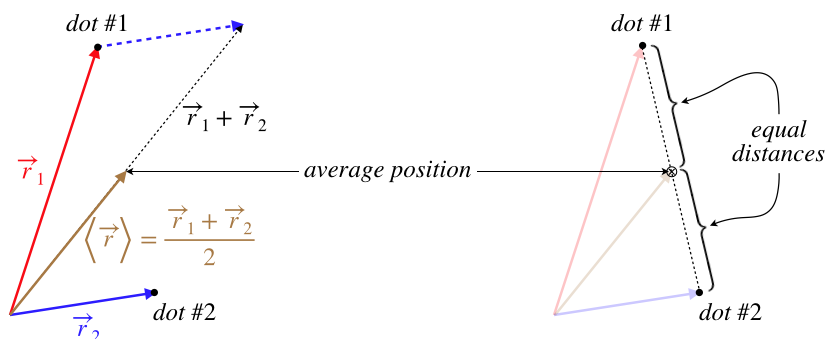
### Average Position

In this lab we will be doing several trials that all produce dots on a piece of paper that measure the position of a marble as it strikes the floor. These dots can all be thought of as existing at the heads of position vectors,  $\vec{r}_1$ ,  $\vec{r}_2$ ,  $\vec{r}_3$ , and so on. As is typically the case for experiments, we will be interested in an average quantity over many trials – in this case the average position at which the marble lands. Finding an average vector is no different from finding an average number, namely:

$$\text{average position} = \langle \vec{r} \rangle = \frac{\vec{r}_1 + \vec{r}_2 + \cdots + \vec{r}_n}{n} \quad (1.1.1)$$

We will not want to actually choose an origin and draw all these vectors, so it is helpful to come up with some way to find an average position directly from the positions of the dots. It isn't hard to show that the average position for two trials is just the point that is located halfway between the positions of the two trials.

**Figure 1.1.1 – Average of Two Position Vectors**



Finding the average position of more than two dots does not have quite as simple of a method, but we will use a trick to keep things from getting too complicated. Notice that if we have four dots, we can write the average position vector this way:

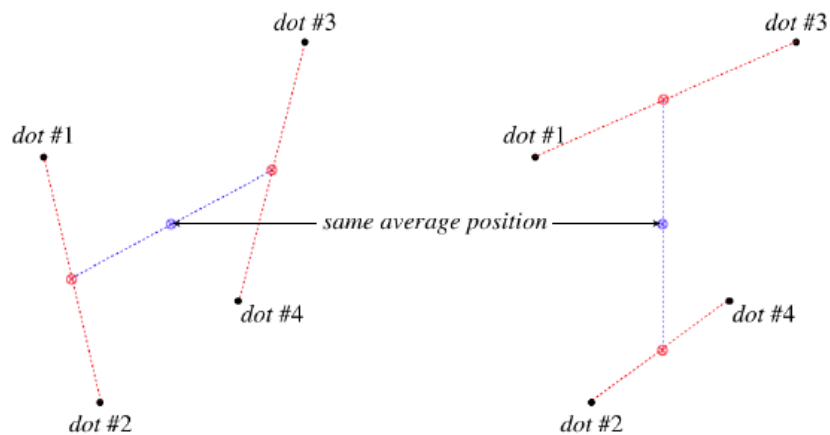
$$\langle \vec{r} \rangle = \frac{\vec{r}_1 + \vec{r}_2 + \vec{r}_3 + \vec{r}_4}{4} = \frac{\frac{\vec{r}_1 + \vec{r}_2}{2} + \frac{\vec{r}_3 + \vec{r}_4}{2}}{2} \quad (1.1.2)$$

This shows that we can get the average position of four dots by first finding the average positions of two pairs of dots, and then finding the average of those averages. This allows us to just use a ruler to locate the halfway points between pairs of dots to find the average position of all the dots. Notice that this procedure requires that we have some power-of-2 number of dots (2, 4, 8, 16, etc.). We could do it for a different number, but then we lose the "halfway between points" simplicity, and since we have control over the number of trials, we will stick with this method.

It should also be noted that it doesn't matter how we pair-off the points – in the end we end up with the same average position:

$$\langle \vec{r} \rangle = \frac{\frac{\vec{r}_1 + \vec{r}_2}{2} + \frac{\vec{r}_3 + \vec{r}_4}{2}}{2} = \frac{\frac{\vec{r}_1 + \vec{r}_3}{2} + \frac{\vec{r}_2 + \vec{r}_4}{2}}{2} \quad (1.1.3)$$

**Figure 1.1.2 – Average Position of Four Dots Found Two Ways**



## Statistical Uncertainty

When we perform an experiment, we are interested in more than just the average value we obtain from many trials, we want to know to what extent this average can be trusted. That is, we want to know how *uncertain* we are that what have measured can be applied to any conclusions we might wish to draw. In the experiment we will perform, we will be "aiming" the marbles at a particular point on the paper, and the scatter of the dots is a result of uncertainty in our aim.

Whenever experimental runs have results that are scattered either because of human involvement or because the apparatus is not good at repeating a run very precisely, we determine the uncertainty *statistically*. This consists of computing what is called the *standard deviation*, which goes as follows:

- compute the average of all the data points

$$\langle x \rangle = \frac{x_1 + x_2 + \cdots + x_n}{n} \quad (1.1.4)$$

- compute how far each data point deviates from the average

$$\Delta x_1 = x_1 - \langle x \rangle, \Delta x_2 = x_2 - \langle x \rangle, \dots \quad (1.1.5)$$

- square all the deviations from the average

$$\Delta x_1^2, \Delta x_2^2, \dots \quad (1.1.6)$$

- average the square deviations

$$\langle \Delta x^2 \rangle = \frac{\Delta x_1^2 + \Delta x_2^2 + \cdots + \Delta x_n^2}{n} \quad (1.1.7)$$

- compute the square root of the average

$$\sigma_x = \sqrt{\frac{\Delta x_1^2 + \Delta x_2^2 + \cdots + \Delta x_n^2}{n}} \quad (1.1.8)$$

This description of the computation of standard deviation makes it easy to remember, as we are just computing averages (first of the data points, then of the squares of the deviation of the data point values from the mean), but technically in these situations where we compute a mean from the actual data, there is a actually a slightly more accurate formula for standard deviation. It involves dividing the sum of the square deviations by  $n - 1$ , rather than by  $n$ . We won't go into the technical details of why this is so, but it is important to note that the difference between these can become significant when  $n$  is quite small, as it often will be in our experiments. We will therefore henceforth use the so-called "unbiased" version of the standard deviation:

$$\sigma_x = \sqrt{\frac{\Delta x_1^2 + \Delta x_2^2 + \cdots + \Delta x_n^2}{n - 1}} \quad (1.1.9)$$

The way this method of measuring uncertainty works for our present experiment should be clear: First use the method described above to determine the place on the paper that is the average landing point. Second, measure the distance from each dot to the

average landing point. This is the "deviation from the mean" ( $\Delta x$ , measured in centimeters) of each data point. Then do the math from there.

## Estimated Uncertainty

Another place where we introduce uncertainty in our results is in measurements. For example, if we are measuring a distance with a ruler, we would not expect our measurements to be accurate down to the micron ( $\frac{1}{1000}$  millimeter), and we would estimate the uncertainty of these measurements to be more like in the range of perhaps a few millimeters. So in the experiment described above, our measurements of distances between dots and between the average landing points and dots introduces uncertainty into our results, because our measuring device does not measure these distances exactly. However, in this case we find that the tiny "few millimeter" estimated uncertainty of these distance measurements is insignificant compared to the statistical uncertainty associated with human aim (which is in the "few centimeter" range). We can therefore ignore the estimated uncertainty associated with ruler measurements for this experiment, as it contributes a negligible amount. Though both types of uncertainty typically occur in any experiment, it is usually true that only one type is the dominant version, allowing us to ignore the other. The simplest way to get a sense of this is to do a few repetitions of (what should be identical) runs, to get a sense of how much the results "scatter." While this experiment has this scatter greatly exceed the measurement uncertainty, more often than not, the reverse will be true. This is because we will use apparatuses that do a decent job of repeating runs.

So how do we make a decent estimate of a measurement uncertainty? Without going into the details that you may have encountered in a statistics class (like the nuances of the central limit theorem and the assumption of a normal distribution for our measurements), we will say that the range of uncertainty is such that we will expect that *roughly two-thirds of the data points will land within one standard deviation of the average*. While this works out automatically when we do the uncertainty statistically, we will use this as the standard for making our uncertainty estimates for measurements as well. That is, estimate the uncertainty of a measurement such that you would expect the true value of the measurement to lie within the uncertainty range of your recorded measurement roughly two-thirds of the time.

## A Good Example to Keep in Mind

If you find the notions of statistical and estimated uncertainties confusing, here is a good model to keep in mind for them. Suppose you perform an experiment that involves measuring both a distance between two well-defined points, and a time interval between two events. So for example, a bouncing marble has pretty well-defined landing points, and the times between the landings is a time interval. The separation of the two points involves using a ruler or tape measure, and you can look at the markings on the measuring device to get an idea of the uncertainty in that measurement. This is an estimated uncertainty. The time interval is different – the device that launched the marble may not be consistent, but more importantly, if a human being is pressing a stopwatch when they see the marble land, then the uncertainty in this measurement is more amenable to statistical calculation – measure the time interval for several "identical" cases many times and compute the standard deviation.

---

This page titled [1.1: Background Material](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Tom Weideman](#) directly on the LibreTexts platform.