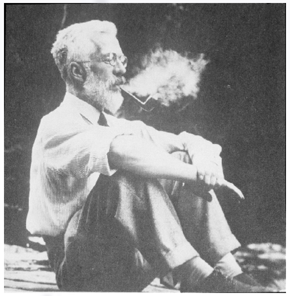
Applied Statistics

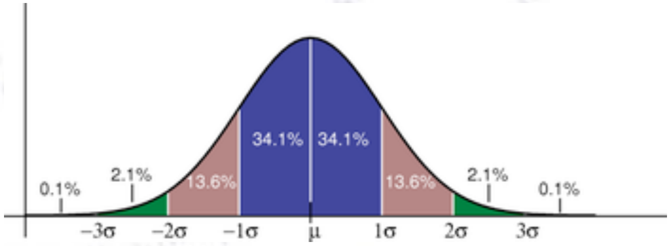
Project objectives and evaluation points








Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense"

Project objective

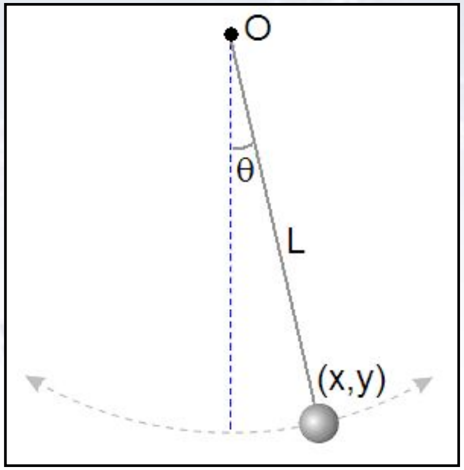
The project in Applied Statistics is to
measure the gravitational acceleration,

g

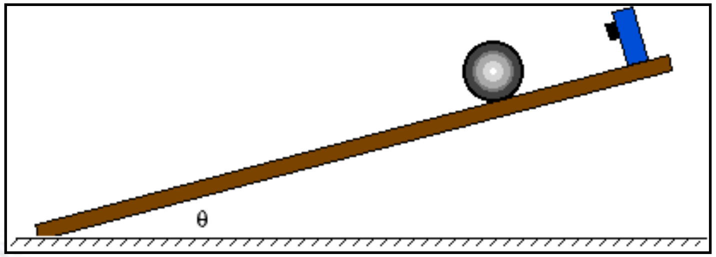
with the greatest possible <u>correct</u> precision
and the most possible <u>cross checks</u>,
using two different experiments

Applied Statistics - Project

The project in Applied Statistics uses two different experiments:



Simple pendulum:
Measure **length** and **period** of the pendulum. Length is measured with a measuring band and a laser, and time by your hand.



Ball rolling down incline:
Measure **incline angle**, **distance between gates**, **ball radius**, **rail distance** and **gate passage times**. First four are measured by hand, while timing is extracted from data files.

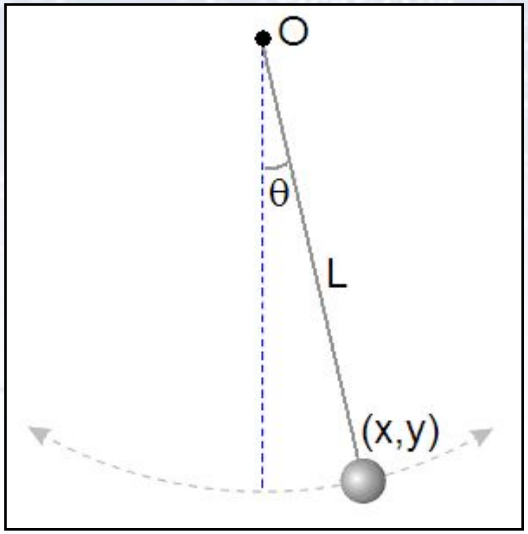
The project purpose is to learn how to extract, minimise and propagate errors. Before doing the experiments, please consider through error propagation, which of the measurements are going to be most challenging/limiting.

For more information, please look at the project webpage.

Experiment formulae

The pendulum formula is well known:

$$g = L\left(\frac{2\pi}{T}\right)^2$$



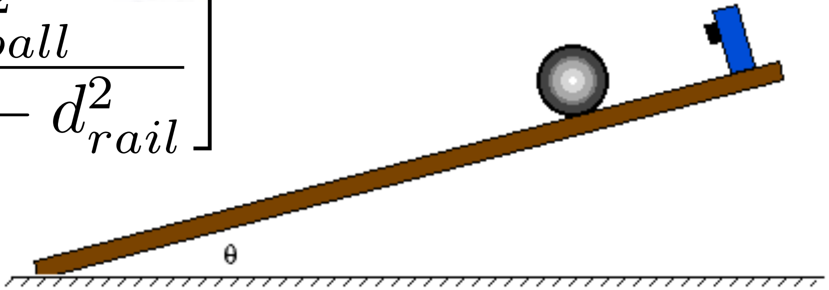
The resulting error formula is easy:

$$\sigma_g^2 = \left(\frac{2\pi}{T}\right)^4 \sigma_L^2 + \left(-2L\frac{(2\pi)^2}{T^3}\right)^2 \sigma_T^2$$

For the ball on incline, the formula is a bit more involved:

$$g = \frac{a}{sin(\theta \pm \Delta\theta)}\left[1 + \frac{2}{5}\frac{D_{ball}^2}{D_{ball}^2 - d_{rail}^2}\right]$$

The resulting error formula is in this case not that nice, but certainly doable.



This is a case, where the numerical solution is a good cross check!

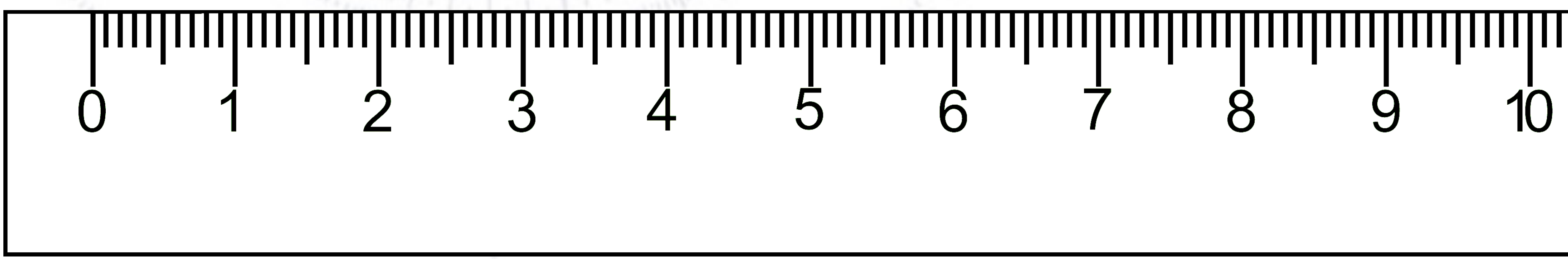
Estimating uncertainties

Estimating uncertainties is never easy, and always yields an inaccurate and possibly **doubtful or flawed result**.

A **rule of thumb** is, that one can read off at a precision of 1/2 the smallest instrument division (i.e. 0.5mm on a folding rule). **But…**

- For some instruments, it can be done more precisely (e.g. large goniometer).
- For some setups, it is not the instrument that limits the precision, but rather experimental conditions (e.g. long pendulum).

Much better is to **estimate the uncertainty from the data itself**. That is why one should think about the design of an experiment, and also ensure to make multiple independent measurements.



Measurement situation

There are four possible situations in experimental measurements of a quantity:

One measurement, no error:

X = 3.14

Situation: You are f***ed!

You have no clue about uncertainty, and you can not obtain it!

Several measurements, no errors:

X1 = 3.14
X2 = 3.21
X3 = …

Situation: You are OK

You can combine the measurements, and from RMS get error on mean.

One measurement, with error:

X = 3.14 ± 0.13

Situation: You are OK

You have a number with error, which you can continue with.

Several measurements, with errors:

X1 = 3.14 ± 0.13
X2 = 3.21 ± 0.09
X3 = …

Situation: You are on top of things!

You can both combine to a weighted average and check with a chi-square.

Measurement situation

There are four possible situations in experimental measurements of a quantity:

One measurement, no error:

X = 3.14

Situation: You are f***ed!

You have no clue about uncert
and you can not obtain it

Several measurements, n rrors:

X1 = 3.14
X2 =

: You are OK

u can combine the measurements,
and from RMS get error on mean.

Commit this line of thinking to memory!

One measure ror:

.13

Situation: You are OK

You have a number with error,
which you can continue with.

Several measurements, with errors:

X1 = 3.14 ± 0.13
X2 = 3.21 ± 0.09
X3 = …

Situation: You are on top of things!

You can both combine to a weighted,
average and check with a chi-square.

Measurement situation

Commit this line of thinking to memory!

For project: Repeat measurements in an independent way to get uncertainties.

There are four possible situations in experimental measurement … antity:

One measurement, no error:

X = 3.14

Situation: You are f…

You have n…
and y…

Several … rrors:

… measurements,
… get error on mean.

O…

Situatio… are OK

You have a number with error,
which you can continue with.

…eral measurements, with errors:

X1 = 3.14 ± 0.13
X2 = 3.21 ± 0.09
X3 = …

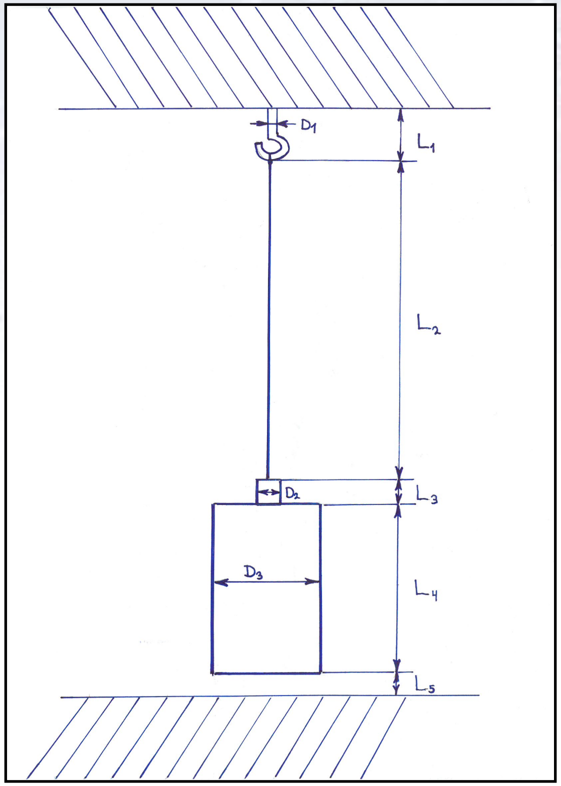
Situation: You are on top of things!

You can both combine to a weighted,
average and check with a chi-square.

Pendulum objectives

What should you have measured in order to have everything needed for measuring g?

$$g = L\left(\frac{2\pi}{T}\right)^2$$



The answer is clear from the formula, but each measurement consists of several measurements!

It is generally worthwhile to make a good drawing ahead of doing the measurements.

Avoid bouncing pendulum, as it changes its length!

<u>Make sure that you answer the following:</u>

- What is the timing precision of **each person** in the group?
- What is the gravitational acceleration g and the errors from:
    - ✦ Length of pendulum.
    - ✦ Period of pendulum.

Ball on incline objectives

What should you have measured in order to have everything needed for measuring g?

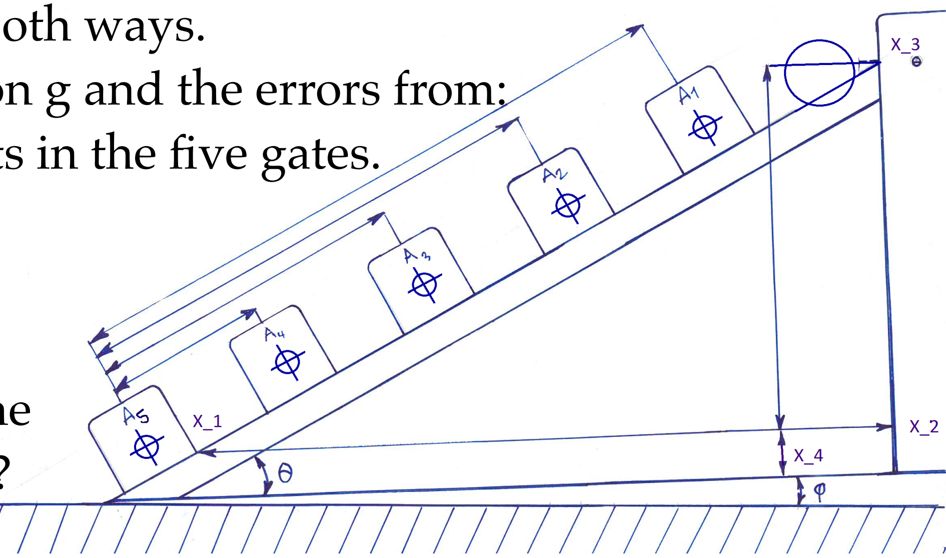
$$g = \frac{a}{sin(\theta \pm \Delta\theta)} \left[1 + \frac{2}{5} \frac{D_{ball}^2}{D_{ball}^2 - d_{rail}^2} \right]$$

Ask yourself, what the critical measurements are. Where do you expect the largest impact on the result and uncertainty to come from?

<u>Make sure that you answer the following:</u>

- What is the angle of the rail θ, and what is the angle of the table, Δθ.
  - You should measure the angle in both ways.
- What is the gravitational acceleration g and the errors from:
  - Timing and distance measurements in the five gates.
  - Ball radius and rail distance.
  - Angle(s) of the rail.

Finally, perhaps you can eliminate some of your uncertainty by making θ = 90º?



Combining measurements

Given repeated measurements (by individual group members) of several quantities, that can be combined, what is the best way forward?

Combine at the end of analysis:

<u>Measurements:</u>
L1 = 3.543 ± 0.002 m
T1 = 3.942 ± 0.002 s
⇒ g1 = 9.821 ± 0.005 m/s²

L2 = 3.545 ± 0.003 m
T2 = 3.940 ± 0.003 s
⇒ g2 = 9.827 ± 0.007 m/s²

L3 = 3.523 ± 0.002 m
T3 = 3.944 ± 0.003 s
⇒ g3 = 9.771 ± 0.006 m/s²

<u>Combination:</u>
g = 9.806 ± 0.004 m/s²
Chi2 = 28.3, Ndof = 2
Prob(Chi2,Ndof) = 7.5 x 10^{-7}

Combine each quantity first:

<u>Measurements:</u>
L1 = 3.543 ± 0.002 m
L2 = 3.545 ± 0.003 m
L3 = 3.523 ± 0.002 m
⇒ L = 3.537 ± 0.002 m

Chi2 = 30.8, Ndof = 2
Prob(Chi2,Ndof) = 2.1 x 10^{-7}

T1 = 3.942 ± 0.002 s
T2 = 3.940 ± 0.003 s
T3 = 3.944 ± 0.003 s
⇒ T = 3.942 ± 0.002 s

Chi2 = 1.3, Ndof = 2
Prob(Chi2,Ndof) = 0.52

<u>Combination:</u>
g = 9.806 ± 0.004 m/s²

Combining measurements

Given repeated measurements (by individual group members) of several quantities, that can be combined, what is the best way forward?

Combine at the end of analysis:

Measurements:
$L1 = 3.543 \pm 0.002$ m
$T1 = 3.942 \pm 0.002$ s
$\Rightarrow$ **$g1 = 9.821 \pm 0.005$ m/s²**

$L2 = 3.545 \pm 0.003$ m
$T2 = 3.940 \pm 0.003$ s
$\Rightarrow$ **$g2 = 9.827 \pm 0.007$ m/s²**

$L3 = 3.523 \pm 0.002$ m
$T3 = 3.944 \pm 0.003$ s
$\Rightarrow$ **$g3 = 9.771 \pm 0.006$ m/s²**

Combination:
$g = 9.806 \pm 0.004$ m/s²
Chi2 = 28.3, Ndof = 2
Prob(Chi2,Ndof) = 7.5×10^{-7}

I would argue for combination within each quantity, to check for consistency.

Combine each quantity first:

Measurements:
$L1 = 3.543 \pm 0.002$ m
$L2 = 3.545 \pm 0.003$ m
$L3 = 3.523 \pm 0.002$ m
$\Rightarrow$ **$L = 3.537 \pm 0.002$ m**

Chi2 = 30.8, Ndof = 2
Prob(Chi2,Ndof) = 2.1×10^{-7}

$T1 = 3.942 \pm 0.002$ s
$T2 = 3.940 \pm 0.003$ s
$T3 = 3.944 \pm 0.003$ s
$\Rightarrow$ **$T = 3.942 \pm 0.002$ s**

Chi2 = 1.3, Ndof = 2
Prob(Chi2,Ndof) = 0.52

Combination:
$g = 9.806 \pm 0.004$ m/s²

Cross checks are VITAL

The two experiments are relatively simple, but you should **imagine that they are more complicated** (and potentially ground breaking), and that you need to **convince others**, that what you're doing is **correct and accurate**.

Imagine the following question from a reviewer:

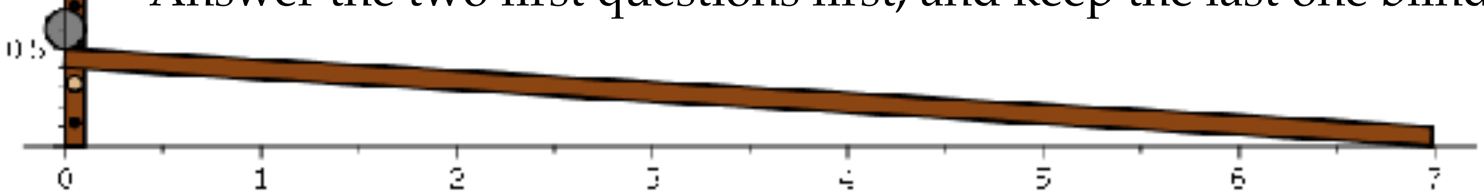
"Do you know that you measure the angle correctly with the goniometer?"

I know that it is unlikely, but in your next experiment, you'll be standing with a complicated Xmeter (which you build yourself?), and not being sure.

Measuring things in **two independent ways** yielding consistent results is VERY convincing. For the angle measured both with the goniometer and through trigonometry (and turning 180^o) allow such a cross checks:

- Do they give consistent values for $\Delta\theta$?
- Does the goniometer show a consistent change, when you turn it 180^o?
- Do the two methods give consistent measures of the angle, θ?

Answer the two first questions first, and keep the last one blinded!



Notes on the angles

Discussion of the angle θ

The angle θ, between the rail and the direction of gravity, can (and should) be measured in **two independent ways**, which allows for a vital cross check:

With the goniometer: $$\theta = \theta_{\mathrm{gonio}}$$

Using trigonometry and turning experiment: $$\theta = \theta_{\mathrm{trig}} + \Delta\theta_{\mathrm{turn}}$$

You might think, that doing things in two independent ways is needless. But this is very important in experiments (which might be extremely complicated and rely on many assumptions!), as this ensures the correctness of the central value, and also tests if the uncertainties are realistic.

For this reason, the formula for g for the ball-on-incline experiment has two versions, depending on angular measurement, and with the above one has:

$$g = \frac{a}{\sin(\theta)} \left[1 + \frac{2}{5} \frac{D_{ball}^2}{D_{ball}^2 - d_{rail}^2} \right]$$

Note on $\Delta\theta$

The angle of the table of the Ball-on-Incline (BoI) experiments - denoted $\Delta\theta$ - can be determined in two ways (thus again allowing for cross check).

1. Using a goniometer before and after turning the experiment 180 degrees.

2. Measuring the acceleration before (normal direction, "norm") and after (reverse direction, "rev") turning the experiment 180 degrees, and equating the value for g between the two measurements:

$$\frac{a_{\mathrm{norm}}}{\sin(\theta + \Delta\theta)} = g = \frac{a_{\mathrm{rev}}}{\sin(\theta - \Delta\theta)}$$

As we can measure the acceleration in both configurations and also the angle θ, we have one equation with one unknown, which happen to have an analytical solution:

$$\Delta\theta = \frac{(a_{\mathrm{norm}} - a_{\mathrm{rev}})\sin(\theta)}{(a_{\mathrm{norm}} + a_{\mathrm{rev}})\cos(\theta)}$$

Notes on your report

Report content

Your report is intended for your fellow students, and you therefore do not need to make a long description of the experimental setup.

However, from your report, your fellow students (and we) should be able to **repeat/reproduce your experiment and subsequent data analysis**. Thus you have write what measurements you make (can be put in appendix, see next page), and exactly what you do with them.

Particularly important is, that you apply cross checks and Chi2 evaluations, whenever you can, and use these to evaluate uncertainties and possibly exclude measurements. This description is very important.

In the end, we simply want to see that you can get from raw data to final results, and that you can convince others (your peers and us), that what you have done is correct.
Therefore, make sure that you go through your numbers and errors and check that they are "reasonable". If they are not, find and correct the error or at least comment.

Example of appendix

APPENDICES

A Pendulum Experiment

A1 Experimental Setup

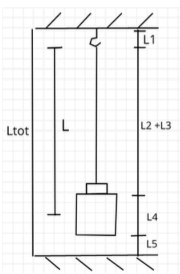


Fig. A1: Schematic representation of the pendulum experimental setup.

Oscillation number	Time [s]
1	2.7956
2	5.6492
3	8.4672
4	11.2882
5	14.0722
6	16.8781
7	19.7594
8	22.4898
9	25.3526
10	28.1713
11	30.9811
12	33.8216
13	36.6129
14	39.4555
15	42.2498
16	44.9828
17	47.9462
18	50.6521
19	53.5126
20	56.3769
21	59.1553
22	61.9397
23	64.7498
24	67.6178
25	70.4369

Table AII: Obtained values for one measurement process, used in Fig. 1. The number of decimals is not significant, but corresponds to the digital precision of the script used to obtain the timing values.

A2 Experimental Results

L_1^i[cm]	$L_{(2+3)}^i$[cm]	L_4^i[cm]	L_5^i[cm]	L_{tot}^i[mm]
3.55	195.45	2.95	7.65	2099
3.65	195.25	3.05	7.71	2108
3.60	195.30	3.00	7.60	2108
3.50	195.30	3.10	7.70	2108

Table AI: Tape and laser measurements of different parts of the pendulum experimental setup. The laser measurement, L_{tot}, is an estimation of the distance from top (where the pendulum is hung), to the bottom of the floor (the laser's initial point).

T[s]	σ_T[s]	t_0[s]	σ_{t0}[s]
2.82	0.03	0.0	0.4
2.82	0.03	-0.1	0.4
2.82	0.03	-0.1	0.4
2.81	0.03	0.0	0.4

Table AIV: Results of the first linear fit, where t_0 represents the offset parameter in the fit $y = mx + b$. σ_T represents the error on the slope estimation, which is not the actual error on T (that one is estimated and shown in Table *AIII*).

T_i[s]	σ_T^i[s]	χ^2	$P(\chi^2)$
2.8165	0.0009	24.7	0.37
2.8177	0.0015	32.0	0.10
2.818	0.003	14.3	0.92
2.8121	0.0012	19.0	0.70

Table AIII: Results of the second linear fit. A χ^2 test is performed. By construction a $P(\chi^2) \sim 0.5$ was expected. The second test provides a really small $P(\chi^2)$, meaning that the fitting function does describe the histogrammed data very well (*i.e.*, it is not gaussian), whereas the 3rd and 4th tests provide a rather large probability, that we could have got out of luck.

μ_{ts}^i[s]	σ_{ts}^i[s]
0.003	0.03
-0.047	0.06
-0.088	0.11
0.006	0.04

Table AV: Obtained means, μ, and RMS, σ, from the gaussian fits to the binned time residuals.

$\bar{T}$[s]	$\sigma_{\bar{T}}$[s]	χ^2	$P(\chi^2)$
2.8156	0.0006	11.89	0.008

Table AVI: Values of resulting T and chi-squared test (see the Discussion section for an interpretation).

B Ball on Incline Experiment

B1 Experimental setup

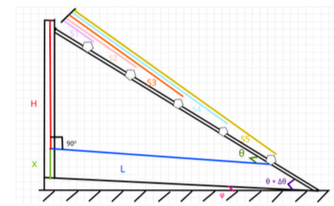


Fig. B1: Schematic representation of the experimental setup used to obtain g by throwing a ball down an inclined plane.

B2 Experimental Results

Gate 1[cm]	*Gate* 2[cm]	*Gate* 3[cm]	*Gate* 4[cm]	*Gate* 5[cm]
21.90	36.95	52.55	69.35	87.15
21.96	36.84	52.51	69.33	87.05
21.95	36.90	52.55	69.35	87.10
21.85	36.80	52.45	69.35	87.15

Table BI: Measurements of the position of each gate in the experimental setup.

L[cm]	x[cm]	H[cm]
83.85	3.90	20.25
83.99	3.98	20.46
83.58	4.00	20.30
83.75	4.05	21.15

Table BII: Measurements of the experimental setup used to obtain a trigonometric value of θ (see Fig. B1).

D_{small}[mm]	D_{big}[mm]	d_{rail} [mm]
10.98	12.70	5.95
10.88	12.70	6.18
10.84	12.67	5.98
10.82	12.70	6.10

Table BIII: Measurements obtained with a slide gauge in order to estimate the value of the gravitational constant.

Side A		Side B	
θ_{norm}[°]	$\theta_{180°}$[°]	θ_{norm}[°]	$\theta_{180°}$[°]
14.90	14.20	13.10	13.20
14.60	14.20	13.20	13.10
14.90	14.20	13.10	13.00
14.50	14.20	13.20	13.00

Table BIV: Goniometer measurements of the angles.

$Time_{s1}$[s]	$Time_{s2}$[s]	$Time_{s3}$[s]	$Time_{s4}$[s]	$Time_{s5}$[s]
0.5197	0.7273	0.8831	1.0197	1.1435
0.1193	0.3305	0.4877	0.6248	0.749
0.0671	0.2848	0.4442	0.5829	0.7081
0.2966	0.521	0.6829	0.8227	0.9489

Table BV: Estimated passage times when the big ball is thrown and the experiment is facing side A.

a[m/s²]	σ_a[m/s²]	v_0[m/s]	σ_{v_0}[m/s]	s_0[m]	σ_{s_0}[m]
B.B norm. side					
1.560	0.011	-0.252	0.009	0.140	0.003
1.565	0.010	0.356	0.004	0.1655	0.0009
1.559	0.010	0.413	0.004	0.1879	0.0007
1.56	0.03	0.006	0.009	0.1416	0.0017
Resulting a	**(1.56 ±**	**0.03)**	m/s²		
B.B rev. side					
1.412	0.009	0.203	0.005	0.1438	0.0012
1.414	0.009	-0.259	0.008	0.156	0.003
1.412	0.09	-0.099	0.007	0.1325	0.0025
1.414	0.009	-1.225	0.015	0.656	0.011
Resulting a	**(1.41 ±**	**0.03)**	m/s²		
S.B norm. side					
1.503	0.000	-2.155	0.001	1.651	0.001
1.500	0.010	0.114	0.006	0.1186	0.0017
1.504	0.010	0.786	0.001	0.339	0.000
1.502	0.010	-0.354	0.009	0.157	0.004
Resulting a	**(1.50 ±**	**0.03)**	m/s²		
S.B rev. side					
1.360	0.009	-0.150	0.007	0.139	0.003
1.360	0.008	-1.599	0.015	1.097	0.013
1.363	0.009	-0.387	0.009	0.189	0.004
1.357	0.010	-0.283	0.009	0.126	0.004
Resulting a	**(1.36 ±**	**0.04)**	m/s²		

Table BVI: Results obtained for the fit done to a parabola of the form $y(x) = \frac{1}{2}ax^2 + v_0x + s_0$ for the values of the gate distances in function of the time elapse. Where a, v_0, and s_0 corresponds to the value of acceleration, the initial velocity and the initial position respectively. The resulting values of the acceleration is the average of the values obtained from the fit and the error is given by the RMS.

$\bar{a}$[m/s²]	$\sigma_{\bar{a}}$[m/s²]	χ^2	$P(\chi^2)$
1.56	0.03	0.01	0.99
1.41	0.03	0.001	0.99
1.50	0.03	0.005	0.99
1.36	0.04	0.002	0.99

Table BVII: Values of the resulting accelerations, χ^2-test and Prob(χ^2).

B3 Error propagation on g

$$\frac{\delta g}{\delta a} = \frac{1}{\sin(\theta \pm \Delta\theta)} \left(1 + \frac{2}{5} \frac{D_{ball}^2}{D_{ball}^2 - d_{rail}^2}\right) \tag{1.12}$$

Example of writing up "raw" measurements, making the analysis reproducible!

Project evaluation

Project evaluation

Pendulum:

- Did you measure $T \pm \sigma(T)$ correctly? Combine with Chi2 and comments?
- Did you measure $L \pm \sigma(L)$ correctly? Combine and check correctly?
- Did you provide the individual T and L precisions/uncertainties on g?
- Did you measure each team members timing precision and submit these?

Ball on incline:

$$\left.\begin{array}{l}\bullet\ T \pm \sigma(T)\\ \bullet\ L \pm \sigma(L)\end{array}\right\} \Rightarrow a \pm \sigma(a), \text{ with Chi2 and comments.}$$

- θ, $\Delta\theta$ obtained correctly and
- d, R and errors propagated correctly?

Generally:

- **Correctly propagated uncertainties, showing individual contributions.**
- **Using Chi2 and its probability, whenever possible.**
- All necessary figures and tables there? 2-3 essential figures needed.
- Text enough to understand results? Clear and fitting captions?
- Comment on result (especially inconsistencies) and correct significant digits.

Collect results: Pendulum (T, L, g) and Ball on Incline (T, L, a, θ, $\Delta\theta$, d, R)

Project challenge

The project consist of experiments and data analysis, which well resembles those in real life.

There is TONS of experience to gather from these!!!

For this reason, we give as challenge to persons/groups, if you can manage the following:

- Pendulum measurement better than 1/1000 with full and correct data analysis and error propagation consistent with g.
- Ball on incline measurement better than 1/100 with full and correct data analysis and error propagation consistent with g.

It is perfectly alright NOT to do this, and one is of course allowed to continue in person, and just submit a personal addition.

Bonus Slides

Different equation versions

$$g = \frac{a}{\sin(\theta \pm \Delta\theta)} \left[1 + \frac{2}{5} \frac{R^2}{R^2 - \left(\frac{d}{2}\right)^2} \right]$$

$$g = \frac{a}{sin(\theta \pm \Delta\theta)} \left[1 + \frac{2}{5} \frac{r_{ball}^2}{r_{ball}^2 - (d_{rail}/2)^2} \right]$$

$$g = \frac{a}{sin(\theta \pm \Delta\theta)} \left[1 + \frac{2}{5} \frac{D_{ball}^2}{D_{ball}^2 - d_{rail}^2} \right]$$