Photoelectric Effect

(Dated: May 20, 2025)

I. THEORY

Photoelectric effect is just the particle-particle interaction of an electron with a photon. It is a quantum mechanical phenomenon. It emphasizes the particle nature of light.

Imagine a long glass cylindrical tube with vacuum inside. Put two metal plates at the ends from inside. Connect these plates to the positive and negative ends of an ammeter outside. What do we measure? No current.

Then we shine light on one plate, and we measure a nonzero current, which we call *photocurrent*. What is going on?

In a metal or essentially anything that behaves as a nice enough conductor or even a semi-conductor, nuclei form a lattice and sit tight in place. Electrons are shared among neighbors and are free to travel around, which is what we call conductivity in a nutshell. We refer to this as the *electron sea*. This is slightly different than an electron in a bound state in hydrogen. For an electron in the electron sea, there is a range of allowed energies, which are different for different materials. When we shine light on this material, each light quantum, or *photon*, might kick an electron off the sea if it has sufficient energy. This interaction leaves a net positive charge behind, and what we see if this process continues is a net accumulation of positive charge, or equivalently a nonzero electric current.

Let us focus on an electron that gets kicked off, which we refer to as a *photoelectron*. It is a free particle now. Even though the process is quantum mechanical, it is not necessarily relativistic. The electron flies off with a nonrelativistic kinetic energy:

$$K = \frac{1}{2} m_e v^2. \tag{1}$$

We have an electron in the sea, it literally absorbs a photon of energy $E_\gamma = hf$, and assuming it is sufficient, the electron flies off with some kinetic energy. The photon energy is spent to remove the electron from the sea, which we refer to as the *work function*, denoted W, and give the electron its kinetic energy, K. Logic dictates that

$$E_\gamma \geq K + W. \tag{2}$$

Depending on the initial energy of the electron, there is actually a range of kinetic energies and hence speeds. For the most energetic photoelectron, we have

$$E_\gamma = K_{\max} + W. \tag{3}$$

We can use this relation to find the ratio of two universal constants, namely h/e. Suppose we apply potential to the other metal plate–the one on which we do not shine light–to repel away the incoming stream of photoelectrons. If we apply a sufficiently high voltage, this stops the most energetic photoelectrons, and we make sure that we stop such photoelectrons once the ammeter reading drops to zero. We refer this value of potential the *stopping potential*, denoted V_s.

If we have an incoming electron with energy $K_{\max}$ and if it takes V_s amount of voltage to bring it to rest, then the conservation of energy gives us

$$K_{\max} + (-e)V_s = 0, \tag{4}$$

or

$$K_{\max} = eV_s. \tag{5}$$

Then we can write

$$hf = eV_s + W, \tag{6}$$

or, after slight rearrangement,

$$V_s = \frac{h}{e} f - \frac{W}{e}. \tag{7}$$

Given that h, e, and W are all positive, this relation represents a straight line with a positive slope, negative y-intercept, and positive x-intercept. Since the negative values of the potential doesn't really make sense in this case (because we *apply* potential to stop the electrons), this relation says there must be a *cut-off* frequency, which is found to be $f_c = W/h$. This is the activation frequency for the photon to liberate a photoelectron.

This is the entire theory. We clearly see the effect of frequency on the stopping potential. We still need to think what happens if we increase the light intensity? What do you expect to find out? How can you justify?

II. EXPERIMENT

A. Initial setup

1. Turn on the mercury lamp and wait for it to warm up for about 10 minutes.
2. Put the diffraction grating 2 m from the wall. Mark the spot and remove the grating.
3. Play with the lens to see the clearest possible image on the wall. Then put back the grating.
4. On one side of the line of sight or of the central maximum on the desk, put the sensor. The other side should remain free of obstacles because we want the diffraction grating pattern on the wall. Focus on the wall.
5. See three lines: yellow, green, and violet. If you put a printing paper on the wall, due to its fluorescent properties you will see two UV lines, so you have 5 lines in total. (You might see multiple dim yellow and green lines, and they are due to impurities in the lamp.) Using the formula for diffraction grating maxima,

$$d\sin(\theta) = m\lambda \tag{8}$$

   with $m = 1$ and $\sin(\theta) \approx \tan(\theta) = y/L$, where y is the distance from the central maximum and $L = 2$ m, obtain the wavelengths. Keep track of uncertainties!
6. Using the relation $f\lambda = c$, compute the frequencies along with the propagated uncertainties.

B. Effect of light intensity on stopping potential

1. Focus on one of the UV lines. Remove all the gadgets off the sensor. Make sure that light enters the aperture. Also check the cap behind the white piece. Don't forget to close it again.
2. Use the intensity filter. (There are three filters on your desks: a green filter, a yellow filter, and an intensity filter.) Align "100%" with the aperture, hit the red button behind the sensor, and the record the voltage reading after it stabilizes after a couple of seconds. This is the stopping voltage.
3. Repeat for 80%, 60%, 40%, and 20%.

C. Effect of frequency on stoppping potential

1. Focus on the yellow line. Align it properly on the aperture, then put the yellow filter on. Then hit the red button behind the sensor and record the voltage once it stabilizes. Decide if you need to introduce uncertainties for V_s measurements.
2. Repeat for green with the green filter.
3. Repeat for violet and UV lines without any filter. The sensor can differentiate these lines from the ambient light.

III. STATISTICAL ANALYSIS

Suppose you have a dataset in the form:

Frequency [Hz]	Stopping potential [V]
$f_1 \pm \delta f_1$	$V_{s1} \pm \delta V_{s1}$
$f_2 \pm \delta f_2$	$V_{s2} \pm \delta V_{s2}$
$f_3 \pm \delta f_3$	$V_{s3} \pm \delta V_{s3}$
$f_4 \pm \delta f_4$	$V_{s4} \pm \delta V_{s4}$
$f_5 \pm \delta f_5$	$V_{s5} \pm \delta V_{s5}$

The model, also known as the fit model, also known as the fit function, is linear:

$$\hat{V}_s = b_0 + b_1 f, \tag{9}$$

where $\hat{V}_s$ is the predicted value for the stopping potential, and b_0 and b_1 are the fit parameters. The former is called the y-intercept and the latter is referred to as the slope. The goal is to obtain the best-fit values for the fit parameters, $b_0 = \beta_0 \pm \delta\beta_0$ and $b_1 = \beta_1 \pm \delta\beta_1$, as well as the correlation between the fit parameters.

We note that if we perform a naive linear regression for example using MS Excel's `LINEST()` function, we incorrectly estimate the uncertainties for the fit parameters. Below, we discuss all possible cases for the given data set and present a minimal working example (MWE) to illustrate the differences.

A. Case of $\delta f_i = 0$ and $\delta V_{si} = 0$ for all i

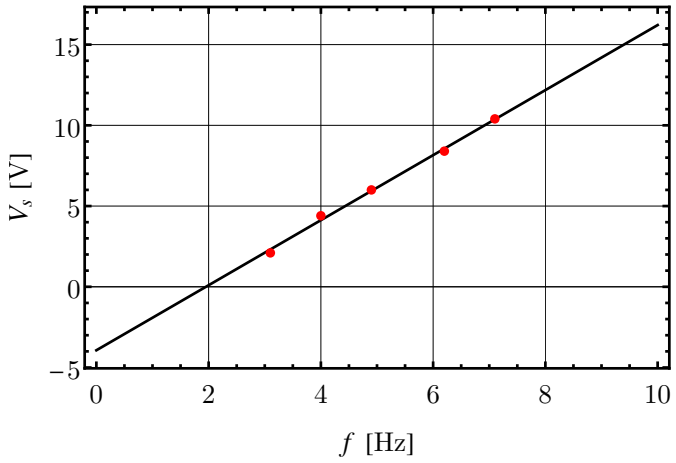
Use the simple linear regression by all means and trust in $\delta\beta_0$ and $\delta\beta_1$.

Consider the example dataset:

```
f = {3.1, 4.0, 4.9, 6.2, 7.1};
Vs = {2.1, 4.4, 6.0, 8.4, 10.4};
```

We obtain

$$b_0 = -3.93165 \pm 1.62777, \tag{10}$$
$$b_1 = 2.01416 \pm 0.309315, \tag{11}$$
$$\rho = -0.961519. \tag{12}$$



B. Case of $\delta f_i = 0$ and $\delta V_{si} \neq 0$ for all i

When the dependent variable has uncertainties, we perform a *weight fit*. We define a χ^2 function as

$$\chi^2 = \sum_{i=1}^{5} \frac{[V_{si} - \hat{V}_s(f_i)]^2}{\delta V_{si}^2} = \sum_{i=1}^{5} \frac{[V_{si} - (b_0 + b_1 f_i)]^2}{\delta V_{si}^2}. \tag{13}$$

This is just a quadratic function of b_0 and b_1. We compute the first partial derivatives with respect to b_0 and b_1, set them equal to zero, and solve them for $b_0 = \beta_0$ and $b_1 = \beta_1$:

$$\left[\frac{\partial \chi^2}{\partial b_0}\right]_{b_0=\beta_0} = 0, \quad \left[\frac{\partial \chi^2}{\partial b_1}\right]_{b_1=\beta_1} = 0. \tag{14}$$

Then we compute the hessian of the χ^2 function and evaluate it at $b_0 = \beta_0$ and $b_1 = \beta_1$:

$$\mathcal{F} = \frac{1}{2}\begin{pmatrix} \frac{\partial^2 \chi^2}{\partial b_0^2} & \frac{\partial^2 \chi^2}{\partial b_0 \partial b_1} \\ \frac{\partial^2 \chi^2}{\partial b_0 \partial b_1} & \frac{\partial^2 \chi^2}{\partial b_2^2} \end{pmatrix}_{b_0=\beta_0, b_1=\beta_1}. \tag{15}$$

This is called the Fisher information matrix. The inverse of the Fisher matrix gives the covariance matrix, $\mathcal{V}$, which looks like

$$\mathcal{V} = \mathcal{F}^{-1} = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix}, \tag{16}$$

where σ_0 and σ_1 are the uncertainties in β_0 and β_1, respectively, and ρ is their correlation.

Consider the example dataset:

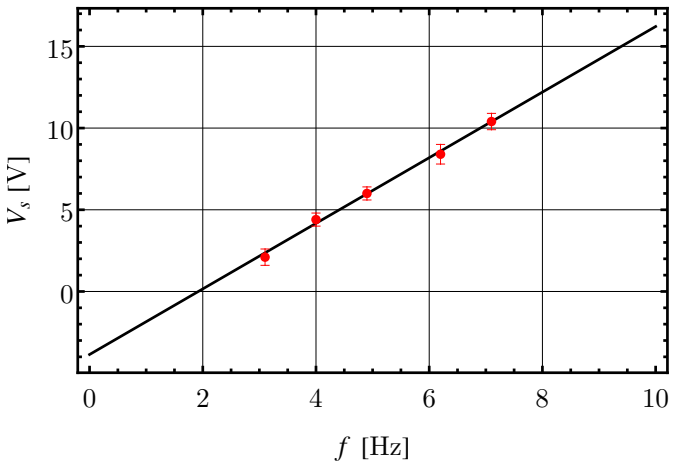
```
f = {3.1, 4.0, 4.9, 6.2, 7.1};
Vs = {2.1, 4.4, 6.0, 8.4, 10.4};
dVs = {0.5, 0.4, 0.4, 0.6, 0.5};
```

We obtain

$$b_0 = -3.85697 \pm 0.780734, \tag{17}$$

$$b_1 = 2.00722 \pm 0.154176, \tag{18}$$

$$\rho = -0.964117. \tag{19}$$



C. Case of $\delta f_i \neq 0$ and $\delta V_{si} = 0$ for all i

The weighted fit of the previous section will not work here because of vanishing uncertainties in the dependent variable. The trick is to define a new model by treating V_s as the independent variable and f as the dependent one:

$$\hat{f} = c_0 + c_1 V_s. \tag{20}$$

We repeat the analysis of the previous section by simply swapping fs by V_ss and by replacing bs by cs. We obtain the best-fit values for the fit parameters as $c_0 = \gamma_0 \pm \delta\gamma_0$ and $c_1 = \gamma_1 \pm \delta\gamma_1$. Then noting that

$$V_s = \frac{f - c_0}{c_1} = -\frac{c_0}{c_1} + \frac{1}{c_1} f, \tag{21}$$

which gives us

$$b_0 = -\frac{c_0}{c_1}, \quad b_1 = \frac{1}{c_1}, \tag{22}$$

we obtain the best-fit values for b_0 and b_1 by letting the errors propagate through

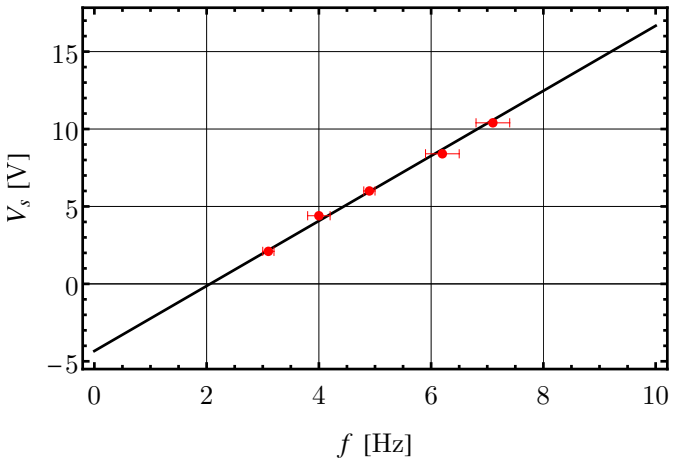
$$\beta_0 \pm \delta\beta_0 = -\frac{\gamma_0 \pm \delta\gamma_0}{\gamma_1 \pm \delta\gamma_1}, \quad \beta_1 = \frac{1}{\gamma_1 \pm \delta\gamma_1}. \tag{23}$$

Consider the example dataset:

```
f = {3.1, 4.0, 4.9, 6.2, 7.1};
df = {0.1, 0.2, 0.1, 0.3, 0.3};
Vs = {2.1, 4.4, 6.0, 8.4, 10.4};
```

We obtain

$$b_0 = -4.33441 \pm 0.524907, \tag{24}$$
$$b_1 = 2.10011 \pm 0.119775, \tag{25}$$
$$\rho = -0.967084. \tag{26}$$



D. Case of $\delta f_i \neq 0$ and $\delta V_{si} \neq 0$ for all i

In this most general case, we apply the method of *orthogonal distance regression*, where our χ^2 function is of the form

$$\chi^2 = \sum_{i=1}^{5} \left(\frac{\Delta f_i^2}{\delta f_i^2} + \frac{\Delta V_{si}^2}{\delta V_{si}^2} \right), \tag{27}$$

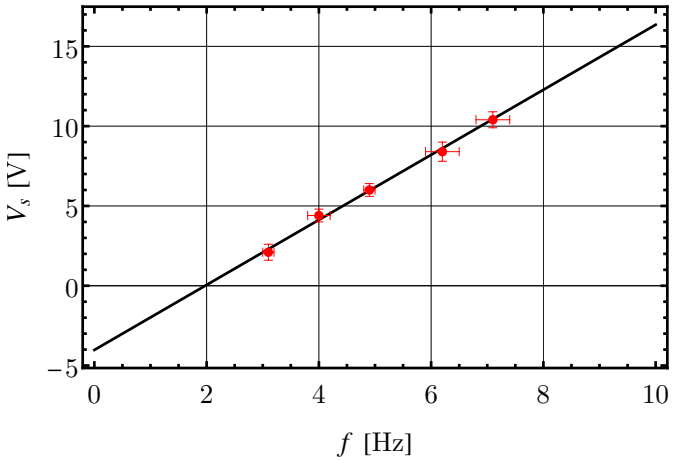
where each Δf_i is now an auxiliary fit parameter, and $\Delta V_{si} = b_0 + b_1(f_i + \Delta f_i) - V_i$. This is now a quadratic function of seven variables (two original fit parameters b_0 and b_1, and five Δf_i); nevertheless, the idea is the same: set the first partial derivatives equal to zero, obtain the best-fit values to obtain the Fisher information matrix. Once we have the Fisher matrix, the rest is the same as in earlier sections, namely the parts where we invert the Fisher matrix to obtain the uncertainties and the correlation.

Consider the example dataset:

```
f = {3.1, 4.0, 4.9, 6.2, 7.1};
df = {0.1, 0.2, 0.1, 0.3, 0.3};
Vs = {2.1, 4.4, 6.0, 8.4, 10.4};
dVs = {0.5, 0.4, 0.4, 0.6, 0.5};
```

We obtain

$$b_0 = -4.02244 \pm 1.03016, \tag{28}$$
$$b_1 = 2.03759 \pm 0.213709, \tag{29}$$
$$\rho = -0.966788. \tag{30}$$



***

This document can be obtained from

https://kagsimsek.github.io/files/teaching/lab7_notes.pdf

The Mathematica notebook with the code for all these cases can be obtained from

https://kagsimsek.github.io/files/teaching/mwe_stat_analysis.nb