# LAB 7: PHOTOELECTRIC EFFECT

## 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 glass 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 are you going to measure? No current.

Now shine light on one plate. We will measure a nonzero current, called *photocurrent*. What's going on?

In a metal or essentially anything that is a nice conductor or even a semi-conductor, the nuclei form a lattice and sit tight. The electrons are shared between neighbors and are free to travel around, which is how conductivity works in a nutshell. We call this an *electron sea*. This is slightly different that an electron in a bound state in hydrogen. Now there is a range of energies, though the details are immaterial. These energies are different for different materials. Now, when we shine light on this metal, each light quantum, or photon, might kick an electron off if the photon has sufficient energy. This interactions leaves a net positive charge behind, so what we see if this process continues is a net accumulation of positive charge, or equivalently a nonzero electric current.

Now, let's focus on this electron, 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 will fly off with a nonrelativistic kinetic energy:

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

We had an electron in the sea, it absorbed a photon of energy $E_\gamma = hf$, and assuming it was sufficient, it flew off with the maximum possible kinetic energy. So, the photon energy is spent to remove the electron from the sea, which we call the *work function*, denoted $W$, and give it its kinetic energy, $K$. Logic dictates that

$$E_\gamma \geq K + W.$$

Now, 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

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

We can use this relation to find the ratio of two universal constants, namely $h/e$. This is how. Suppose we apply potential to the other metal plate (the one on which we don't shine light) to repel away the incoming electrons. If we apply a sufficiently high voltage, this will stop the most energetic photoelectrons, and we will make sure that we have stopped such photoelectrons once the ammeter gives us zero. We call this potential the *stopping potential*, denoted $V_s$. Now, if we have an incoming electron with energy $K_{\text{max}}$ and if it takes $V_s$ amount of voltage, then the conservation of energy gives us

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

or

$$K_{\text{max}} = eV_s.$$

Then we can write

$$hf = eV_s + W,$$

or

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

Given that $h$, $e$, and $W$ are all positive, this is a straight line with 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 says there must be a *cut-off frequency*, which is $f = W/h$. 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. You need to think about this: What happens if just increase the light intensity? What do you expect to find out? How can you justify?

**Experiment**

*I. Initial setup*

1. Turn on the mercury lamp and wait for it to warm up for about 10 minutes.
2. Put the grating 2 meters 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 bring 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. We will get to the sensor shortly. Focus on the wall.
5. You will 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 green and yellow lights, and they are due to impurities in the lamp.) Using the formula for diffraction grating maxima, $d\,\sin(\theta) = m\lambda$ 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 uncertainty.

*II. Effect of light intensity on stopping potential*

1. Focus on one of the UV wavelengths. 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 record the voltage value after waiting for a couple of seconds for the value to stabilize. That's your stopping voltage.
3. Repeat this for 80%, 60%, 40%, and 20%.

*III. Effect of frequency on stopping 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 after waiting for a couple of seconds. Decide if you need to introduce uncertainties for $V_s$ measurements.
2. Repeat for green with green filter.
3. Repeat for violet and UV lines without filter. The sensor can differentiate these lines from the ambient light.

**Statistical analysis**

Suppose you have the following data:

| 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, aka the fit model, aka the fit function, is linear:

$$\hat{V}_s = b_0 + b_1 f.$$

$b_0$ and $b_1$ are the fit parameters. The former is called the $y$-intercept and the latter is referred to as the slope. If you perform a naive linear regression, say using MS Excel's `LINEST()` function, you will probably underestimate the uncertainties for the values you obtain for the fit parameters, $b_0 = \beta_0 + \delta\beta_0$ and $b_1 = \beta_1 + \delta\beta_1$. Below, we discuss widely used methods to improve upon the simple linear regression.

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

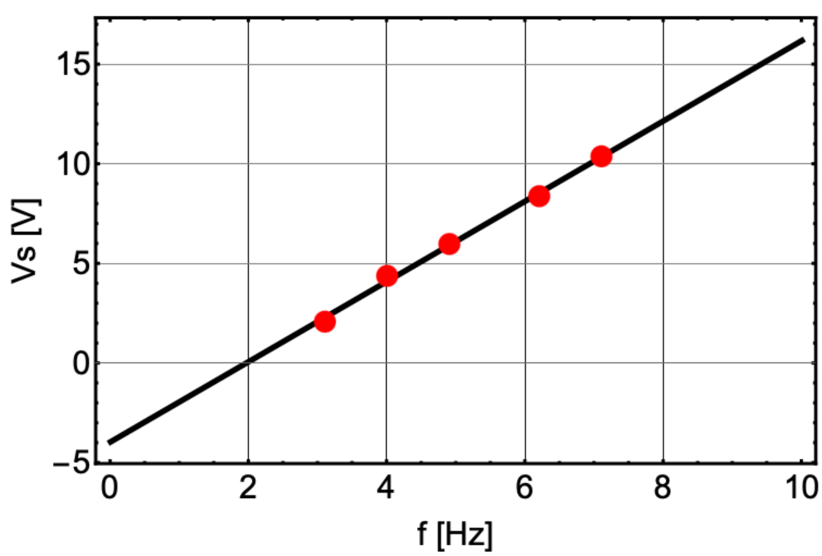
Use the simple linear regression by all means and trust in $\delta\beta_0$ and $\delta\beta_1$.

Consider the example data: (See the Mathematica code for details.)

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

We obtain

$$b_0 = -3.93165 \pm 1.62777$$
$$b_1 = 2.01416 \pm 0.309315$$
$$\rho = -0.961519$$



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

Define a $\chi^2$ function as

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

This is just a quadratic function of $b_0$ and $b_1$. Compute its 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 $b_1 = \beta_1$:

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

Now compute the hessian of the $\chi^2$ function and evaluate it at $b_0 = \beta_0$ and $b_1 = \beta_1$:

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

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

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

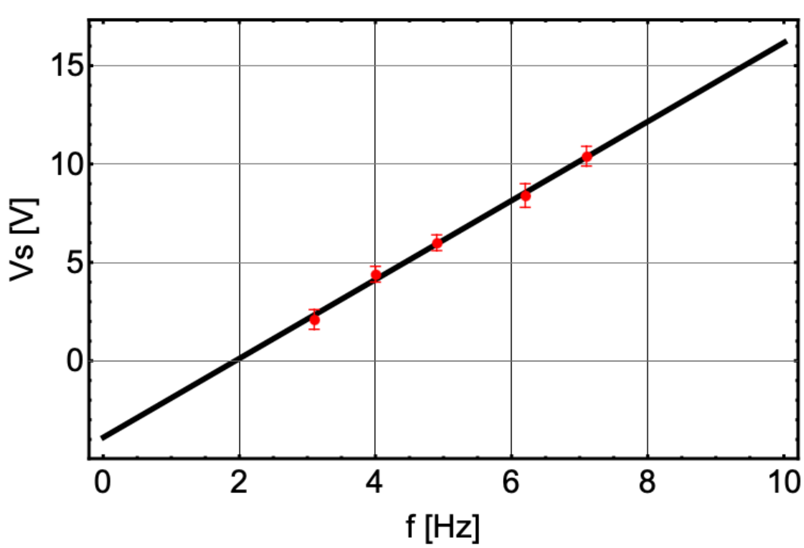
where $\sigma_0$ and $\sigma_1$ are the uncertainties in $\beta_0$ and $\beta_1$, respectively, and $\rho$ is their correlation.

Consider the example data:

```
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

$$b_0 = -3.85697 \pm 0.780734$$
$$b_1 = 2.00722 \pm 0.154176$$
$$\rho = -0.964117$$



*III. 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:

$$f = c_0 + c_1 V_s.$$

Repeat the analysis in the previous section by simply swapping $f$s by $V_s$s. 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

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

which gives us

$$b_0 = -\frac{c_0}{c_1}, \qquad b_1 = \frac{1}{c_1},$$

we obtain the best-fit values for $b_0$ and $b_1$ by

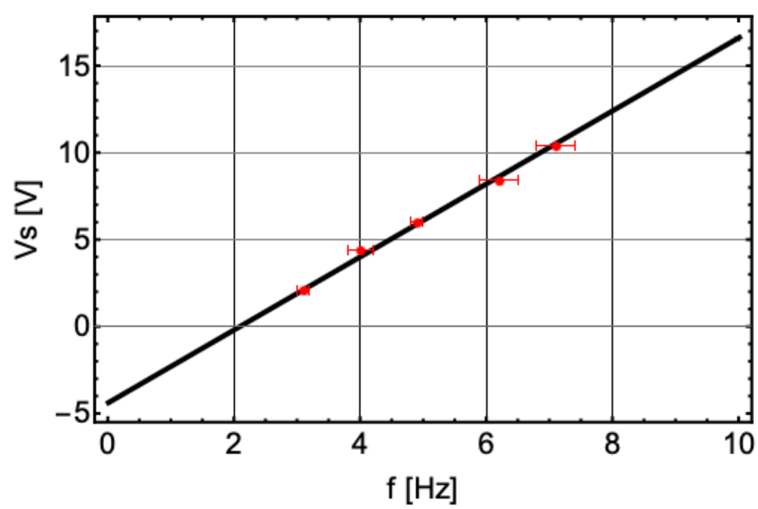
$$\beta_0 = -\frac{\gamma_0 \pm \delta\gamma_0}{\gamma_1 \pm \delta\gamma_1}, \qquad \beta_1 = \frac{1}{\gamma_1 \pm \delta\gamma_1}.$$

Consider the example data:

```
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

$$b_0 = -4.33441 \pm 0.524907$$
$$b_1 = 2.10011 \pm 0.119775$$
$$\rho = -0.967084$$




*IV. 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 $\chi^2$ function is of the form

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

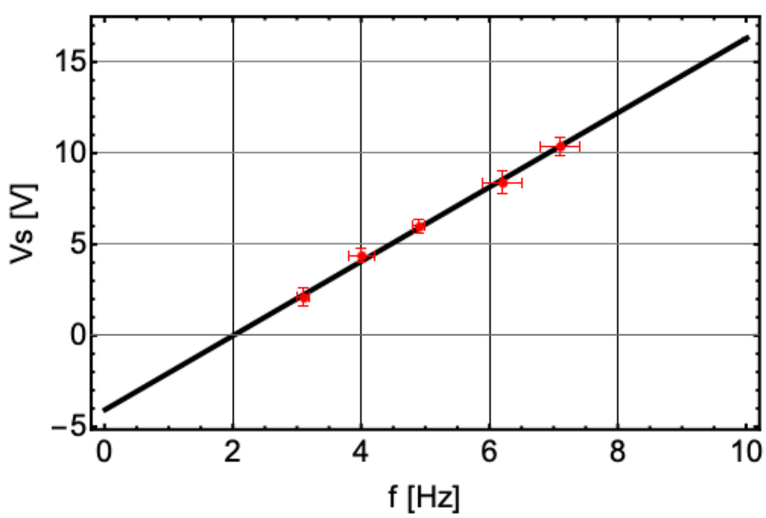
where each $\Delta f_i$ is now an auxiliary fit parameter, and $\Delta V_i = b_0 + b_1(f_i + \Delta f_i) - V_i$. This is a highly nonlinear function of seven variables (two original fit parameters $b_0$ and $b_1$, and five $\Delta f_i$); nevertheless, the idea is the same: set the first partial derivatives equal to zero, obtain the best-fit values of all the seven fit parameters, and evaluate the hessian at these best-fit values to obtain the Fisher information matrix. Once you have the Fisher information matrix, the rest is the same as in previous sections (namely the parts where you invert the Fisher matrix to obtain the uncertainties).

Consider the example data:

```
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

$$b_0 = -4.02244 \pm 1.03016$$
$$b_1 = 2.03759 \pm 0.213709$$
$$\rho = -0.966788$$




***

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