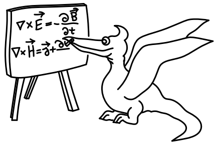

Serial: Spectroscopy in Astronomy

After visiting small scales, in the last part of this year's serial, we will move to large scales—into the universe. With a few exceptions[1], we gain knowledge about distant bodies in astronomy by observing electromagnetic radiation. In the following text, we will focus on its narrow part between the near-ultraviolet and infrared regions, where techniques you might have encountered in optics during physics classes are used. Specifically, we will focus on spectroscopy – a technique that describes the dependence of light intensity on wavelength.

What Do We Measure With?

An instrument that disperses light into a spectrum and allows its recording is called a spectrograph. However, observed objects are faint in astronomy, so it is first necessary to concentrate light into the spectrograph using a telescope. The basic parameters of a telescope are

- the objective lens diameter D, which determines both the amount of light we are able to collect, and also the angular resolution Ψ of the instrument;
- the focal length f, indirectly determining the focal ratio, magnification, and field of view of the instrument.

For the angular resolution when observing at a wavelength λ, the Rayleigh criterion is known as

$$\Psi = 1.22\frac{\lambda}{D}\,.$$

The focal ratio of a telescope, similar to a camera lens, is given in the form of $f/\#$, where the number is $\# = f/D$. Usual values are $f/4$ to $f/20$. In principle, the focal ratio describes the divergence of the beam of rays falling into the focal plane – the smaller it is, the wider the beam is, and the construction of the instrument is technically more demanding (we get further away from the paraxial approximation).

Light from the observed object in the focal plane (where we would place a camera chip if we wanted to photograph the sky) falls onto a slit, through which we select the specific light source whose spectrum we want to measure. We then collimate the passing diverging beam (usually by reflection on a parabolic mirror, but it is also possible to use a converging lens[2]) into a parallel beam. This is dispersed into a spectrum by a dispersive element – for each wavelength, we get a set of parallel rays, the direction of which depends precisely on the wavelength. Finally, it is necessary to image the obtained beam onto a camera chip, either by another mirror (in which case we speak of a second collimator) or by a lens (here, for a change, we speak of an objective lens). It is important to realize that for each wavelength, the spectroscope is an imaging optical instrument. The slit of the spectrograph is therefore imaged on the chip as a narrow rectangle. A narrower slit increases the resolution of the instrument at the cost of less light. Many modern instruments use an optical fiber instead of a slit, one end of which is in the focal plane of the telescope and the other in the optical laboratory at the entrance to the spectrograph. The

[1] For example, gravitational waves and particle radiation.

[2] In practice, it is hardly used; mirrors are easier to manufacture than lenses with the correct shape, and moreover, they do not have chromatic aberration.

advantage is that the spectrograph is not exposed to external conditions and can be thermally and mechanically better stabilized, or even placed in a vacuum.

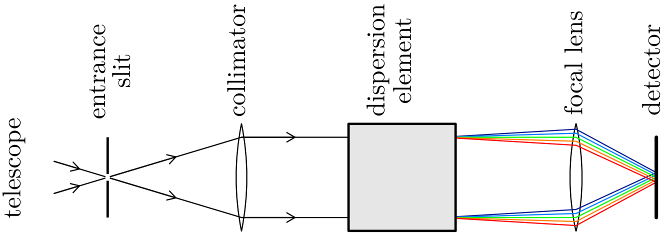

Figure 1: Diagram of a spectrograph.

Optical Prism

Historically, the first dispersive elements were glass prisms, with which, for example, Isaac Newton demonstrated that sunlight consists of a spectrum of color components that are further indivisible and can be recombined into white light. A light ray falling on a prism at an angle α_1 is refracted toward the normal, passes through the glass prism with a refractive index n and an apex angle β, and exits the prism after refracting away from the normal at an angle α_4 to the normal. We are interested in the deviation of the ray from its original direction δ. For the refraction of light at a planar interface, Snell's law holds as

$$\sin\alpha_1 = n\sin\alpha_2\,, \qquad\qquad n\sin\alpha_3 = \sin\alpha_4\,.$$

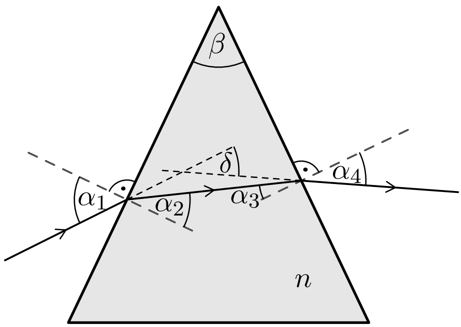

Figure 2: Diagram of a prism.

For the angles α_2, α_3, and β, from the triangle formed by the apex of the prism and both points of incidence, we have $\beta + (90^\circ - \alpha_2) + (90^\circ - \alpha_3) = 180^\circ$, thus $\beta = \alpha_2 + \alpha_3$. Similarly, for the deviation of the ray from the original direction in which it fell on the prism (i.e., the deviation of the ray), we have

$$\delta = \alpha_1 - \alpha_2 - (\alpha_3 - \alpha_4) = \alpha_1 + \alpha_4 - \beta\,.$$

By gradually combining these relations, we get

$$\delta = \alpha_1 - \beta + \arcsin(n\sin\alpha_3) = \alpha_1 - \beta + \arcsin\left(n\sin\left(\beta - \arcsin\left(\frac{\sin\alpha_1}{n}\right)\right)\right).$$

For a given value of β, the deviation is large at perpendicular incidence, then decreases and reaches a minimum for the situation with $\alpha_1 = \alpha_4 = \alpha_\mathrm{c}$, $\alpha_2 = \alpha_3$, and then increases again. For practical reasons, the configuration with minimum deviation is used – in other cases, the circular cross-section of the incident collimated cylindrical beam is changed to an elliptical one. In the case of lower angles of incidence, the resulting beam is compressed in the direction of dispersion, whereby we get a lower resolution due to the finite size of the slit, whose image on the detector widens in the same ratio. It might seem that angles larger than α_c would increase the resolution – this is indeed the case, but at the cost of a larger radius of the objective lens and higher light losses by unwanted partial reflection of light at the first interface[3]. For the angle of incidence in the case of minimum deviation, we have

$$\alpha_2 = \frac{\beta}{2} \quad \Rightarrow \quad \alpha_c = \arcsin\left(n\sin\left(\frac{\beta}{2}\right)\right).$$

For the deviation, we have

$$\delta = 2\left(\alpha_1 - \alpha_2\right) = 2\arcsin\left(n\sin\left(\frac{\beta}{2}\right)\right) - \beta\,.$$

For a prism as a dispersive element, the dependence of the deviation on the refractive index is exactly what is important. This is because it depends on the wavelength, while for glass it approximately holds that

$$n \approx A + \frac{B}{\lambda^2}\,,$$

where A and B are material constants. Two parameters are important for the description of dispersion: angular dispersion

$$D_\lambda = \frac{\mathrm{d}\delta}{\mathrm{d}\lambda} = \frac{\mathrm{d}\delta}{\mathrm{d}n}\frac{\mathrm{d}n}{\mathrm{d}\lambda} = \frac{2\sin\left(\frac{\beta}{2}\right)}{\sqrt{1 - n^2\sin^2\left(\frac{\beta}{2}\right)}}\frac{\mathrm{d}n}{\mathrm{d}\lambda} \approx -\frac{4\sin\left(\frac{\beta}{2}\right)}{\sqrt{1 - n^2\sin^2\left(\frac{\beta}{2}\right)}}\frac{B}{\lambda^3}$$

and the resolving power R defined as

$$R = \frac{\lambda}{\Delta\lambda}\,,$$

where $\Delta\lambda$ is the distance between two structures that we can still resolve in the region of the spectrum around λ. Let us imagine a beam of rays passing through a prism in the minimum deviation configuration with a width h inside the prism. If one extreme ray passes through the apex of the prism, then h is the height of the prism. Between the bottom extreme rays for two wavelengths λ and $\lambda + \Delta\lambda$ (which we want to resolve), there is an optical path difference of

$$\Delta s = 2n(\lambda + \Delta\lambda)\,h\tan\left(\frac{\beta}{2}\right) - 2n(\lambda)\,h\tan\left(\frac{\beta}{2}\right) = \frac{\mathrm{d}n}{\mathrm{d}\lambda}b\Delta\lambda\,,$$

[3] The ratio of reflected and transmitted light intensity at the interface of two media is given by the Fresnel equations.

where b is the length of the prism's base. We can resolve the rays if this difference is at least one wavelength.[4]

By rearranging, we get

$$R = \frac{\lambda}{\Delta\lambda} = b\frac{\mathrm{d}n}{\mathrm{d}\lambda} \approx b\frac{2B}{\lambda^3}\,.$$

Typical values of the resolving power of a prism range from the high tens to the high single-digit thousands. Let us note that the resolving power of a prism strongly depends on the wavelength – it is higher for short wavelengths in the blue region.

Optical Grating

An optical grating is usually a planar surface with periodically repeating parallel lines, grooves, or other structures. Gratings can be either transmissive, where the "grooves" are slits; or reflective, where the reflective surface is usually of a sawtooth profile with a distance d between the teeth. The distance between the periodically repeating structures is called the grating constant. Let us consider a pair of parallel monochromatic rays falling on the grating at an angle α at points separated by d, which are subsequently reflected at an angle β. The reflected rays will interfere constructively if their path difference is an integer multiple of the wavelength

$$m\lambda = a - b = d(\sin\alpha - \sin\beta)\,, \quad m \in \mathbb{Z}\,.$$

This relation is called the grating equation. Gratings are often used in the so-called Littrow configuration, where the direction of the incident and reflected rays is identical ($\alpha = -\beta$), thanks to which we get a simpler relation

$$m\lambda = 2d\sin\alpha\,.$$

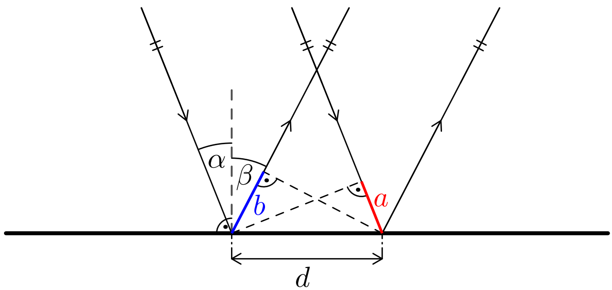

Figure 3: Diagram of diffraction on a grating.

One of the signs of diffraction on a grating is precisely the fact that rays with a given $m\lambda$ interfere constructively in a given direction. If the source used has a spectrum with a large range of wavelengths, overlapping of spectra occurs for the successive orders m and $m+1$ when

$$m(\lambda + \Delta\lambda) = (m+1)\lambda \quad\Rightarrow\quad \Delta\lambda = \frac{\lambda}{m}\,.$$

[4] There are several different criteria for resolution; in general, it would be necessary to calculate the intensity profile of the monochromatic beam emerging from the prism. Thus, our calculation represents only a simplified approach, which, however, yields the correct result (up to a possible numerical coefficient close to 1).

For example, in the third order, the wavelength of 400 nm overlaps with the wavelength of 600 nm in the second order, 1 200 nm in the first order, and 300 nm in the fourth order. If we want to measure the spectrum in the vicinity of 600 nm in a given direction, we must use a filter for the other wavelengths. We derive the dispersion of the grating from the grating equation by simple differentiation, on the left with respect to λ and on the right with respect to β, as

$$m\,\mathrm{d}\lambda = -d\cos\beta\,\mathrm{d}\beta \quad \Rightarrow \quad \frac{\mathrm{d}\beta}{\mathrm{d}\lambda} = -\frac{m}{d\cos\beta}\,,$$

The differential quotient is thus directly proportional to the order of the spectrum. For the resolution of a grating, let us again consider the extreme rays incident on the grating, illuminating N slits, thus a length of Nd of the grating. Again, we can resolve wavelengths for which the path difference is equal to the wavelength

$$\lambda = \Delta s = Nd\frac{\mathrm{d}\sin\beta}{\mathrm{d}\beta}\frac{\mathrm{d}\beta}{\mathrm{d}\lambda}\Delta\lambda = Nd\cos\beta\frac{m}{d\cos\beta}\Delta\lambda\,,$$

where we substituted from the relation for dispersion. By rearranging, we obtain the relation for the resolution

$$R = \frac{\lambda}{\Delta\lambda} = Nm\,.$$

Gratings have resolutions from the low thousands up to 100 000, thus significantly higher than a prism. A disadvantage, however, is that due to the overlapping of orders, we can observe only a smaller part of the spectrum at once.

Other Configurations

In some cases, a different system configuration is advantageous. An echelle spectrograph (from the French word for stairs) utilizes the fact that the dispersion of a grating depends on the angle of incidence. By using a specially prepared grating with an angle of incidence and reflection of up to 80°, it is possible to achieve the same resolution even with a smaller number of grooves per millimeter. We therefore look into the region with high orders of the spectrum m. While for a standard grating spectrograph we use filters to remove the unwanted overlapping of spectral orders, in the case of an echelle spectrograph, we fully utilize this phenomenon. By incorporating a second dispersive element with dispersion in the perpendicular direction and with low resolution, we separate the individual orders from each other.

In the so-called "white pupil" configuration (as in the serial problem), a high-resolution spectrum is first created, and in the place where its image is located (and usually the detector is situated there), a mirror is placed reflecting light into a "second" spectrograph with a low resolution. For each point of the high-resolution spectrum, a low-resolution spectrum is thus formed, containing a series of points with wavelengths corresponding to different orders. Instead of a single track, we will therefore see on the detector a series of several parallel tracks corresponding to different values of m, and thus containing different intervals of wavelengths. Thanks to this, such an instrument can capture a large range of wavelengths, often the entire visible spectrum, in a single image at high resolution (even around 100 000).

Another special arrangement for observation at low resolution is the principle of "slitless/field spectrography", when our instrument does not contain a slit, and light from the

[5] Available from `[https://commons.wikimedia.org/wiki/File:Starlight](https://commons.wikimedia.org/wiki/File:Starlight)_(6815972618).jpg`.

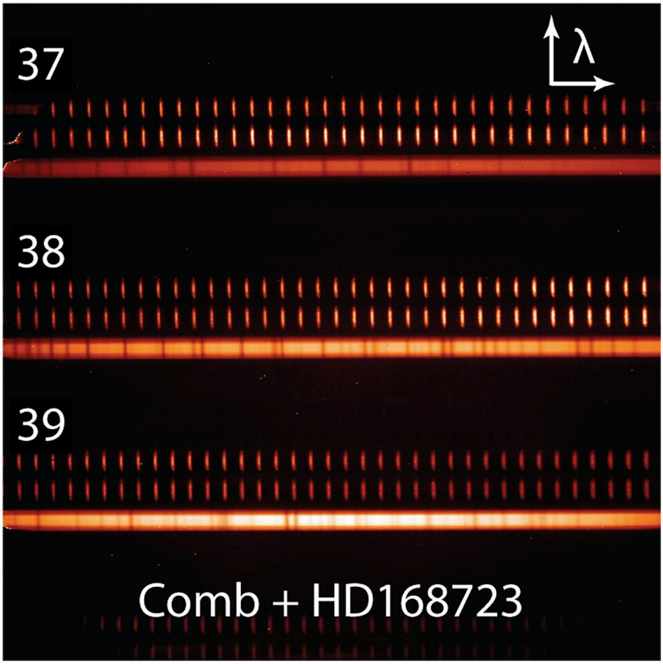

Figure 4: An excerpt of an echelle spectrum. For each order of dispersion, we see three tracks: one for the star (the bottom of the three) with dark absorption lines, and two from the calibration source (a laser frequency comb) with narrow emissions of known wavelengths. The wavelength within an order increases from left to right, and across orders from bottom to top. The original black-and-white image was colored according to the measured intensity; from white (the most signal) through orange and red to black (no signal).[5]

entire field of view falls on the dispersive element. Historically, an optical prism was used in front of the telescope's objective; today, the use of a grating in front of the detector, in the place where photometric filters are normally placed, is more common. With this method, it is possible to obtain the spectrum of a large number of objects simultaneously in a single exposure, albeit at the cost of potential overlapping of spectra belonging to different objects.

Processing and Calibration

Most spectrographs use a CCD or CMOS chip cooled to a low temperature by liquid nitrogen as a detector. We need to convert the obtained image of such a camera into the dependence of intensity on wavelength. For this purpose, besides the image of the observed object, we also need other so-called calibration images:

[6]Available from `[https://esahubble.org/images/heic1412c/](https://esahubble.org/images/heic1412c/)`. The source is NASA&ESA/Hubble.

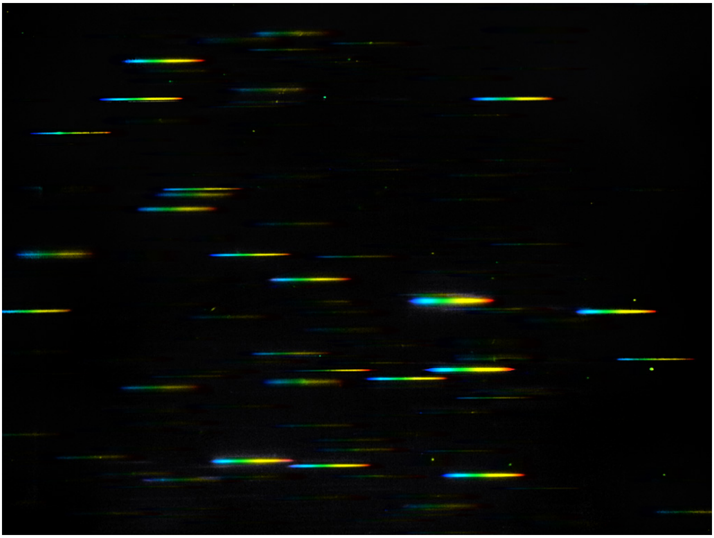

Figure 5: Using a so-called grism, Hubble can simultaneously obtain spectra across the entire field of view. The original black-and-white image was artificially colored according to the corresponding wavelengths of light.[6]

- *Zero*—an image obtained at a zero exposure time of the camera describing the zero level of the camera and the noise of the readout electronics.
- *Dark*—an image at an exposure time corresponding to the science image with the camera shutter closed, now without any incident light. This image serves to remove thermal noise (signal caused by the thermal motion of atoms in the camera); however, for modern cameras cooled by liquid nitrogen, it is negligible and is therefore often not used.
- *Flat*—an image obtained when illuminating the spectrograph with a light source having the smoothest possible spectrum without spectral lines, such as an incandescent light bulb.
- *Comp*—an image of a reference source for which the wavelengths of the spectral lines are known. Typically, cathode lamps filled with a noble gas and a hollow cathode made of a chemically pure metal with a large number of spectral lines are used, e.g., a thorium-argon and uranium-neon lamp.[7]

The first step is to correct the zero point of the individual pixels of the science image, the *flat* image, and the *comp* image by using the *zero* and *dark* images. Subsequently, the profile of the spectrum is identified on them (*order tracing*), and the pixel values are summed in columns perpendicular to the spectrum. The result is a profile of a sort of intensity versus position on the chip in the direction along the spectrum (in pixels). By dividing the profile of the science spectrum by the profile of the *flat* spectrum, we remove small differences in linearity between pixels and the intensity profile caused by the optics (the so-called blaze function). The shape

[7] For more precise calibration, a Fabry-Pérot etalon or a laser frequency comb can be used. These are technically more demanding devices, but they provide uniform coverage with lines of the same intensity.

of the resulting spectrum therefore depends on the *flat* lamp used – it is thus customary to further normalize the spectrum, i.e., to identify the position of the continuum and divide the spectrum by its profile. The resulting spectrum thereby has a value of 1 outside the lines.[8]

However, we still have the spectrum as intensity as a function of position. To convert position into wavelengths, we use the spectrum obtained from the *comp* image. In it, we identify the position of the spectral lines with known wavelengths most commonly by a Gaussian fit. Through the obtained pairs (x, λ), we subsequently fit the dispersion relation, which we use to convert positions into wavelengths for each pixel of the acquired science spectrum. The resulting spectrum is thus a set of intensities and wavelengths corresponding to the pixels on the detector.

And What Do We Need These Spectra For?

Spectroscopy allows us to determine what the observed bodies are composed of and how they move.

Motions of Bodies

For a moving source, we have a change in wavelength from the Doppler effect

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c}\,,$$

where v is the radial component of the velocity.[9] The position of the spectral lines therefore shifts, and we can measure this shift; whether on individual lines or globally across the entire spectrum. In addition to the global motion of an object, it is also possible to measure its internal velocities (such as the rotational speed of stars or galaxies) from the broadening and shape of the spectral lines. In the case of stellar observations, the first step is usually to create a global profile representing an "average" spectral line by utilizing a list of present/expected lines and their depths/weights. A common method is the use of cross-correlation of the observed spectrum and a mask – the resulting function/profile reaches a maximum if the match between the mask and the observation is the best. However, the method of deconvolution (LSD, least-squares deconvolution) is more illustrative. The observed spectrum can be represented as a convolution of a spectrum with thin lines (given by the physical state of the observed substance with zero radial velocity) and the velocity profile of the observed material. Deconvolution allows us to obtain precisely this profile. From its width, we can determine the rotational speed of the star, and from the position of its center, the velocity of the star relative to the observer.

It is important to realize that the instrument moves in space and is therefore not inertial. Indeed, we are observing from the surface of a rotating planet orbiting the Sun. The measured velocities must be corrected for the motion of the observatory relative to the center of mass of the Solar System – the barycentric correction, which reaches up to $30\,\mathrm{km\cdot s^{-1}}$. Meanwhile, the orbital velocities of binary stars have values of a similar order of magnitude. In the case of exoplanets, where we observe the motion of the star caused by the planet's gravity, it even involves values of $100\,\mathrm{m\cdot s^{-1}}$ for hot Jupiters[10] down to units of $\mathrm{cm\cdot s^{-1}}$ for Earth-mass planets

[8] In some cases, it is possible to perform absolute calibration into physical units by utilizing a radiation source with a known profile, or if we know the spectral sensitivity of our instrument. The latter case applies to most satellite instruments, such as the Hubble or Webb telescopes. Observations of suitable stars obtained by them can subsequently be used to calibrate certain instruments on the Earth's surface.

[9] We do not consider relativistic velocities; for them, the relation is more complicated.

[10] I.e., planets with Jupiter's mass with orbital periods of a few days

in a one-year orbit. In this case, it is also necessary to include the Moon and the planets for the calculation of the barycentric correction. It is interesting that a resolution of 100 000 corresponds to a velocity of $3\,\mathrm{km\cdot s^{-1}}$. Measuring radial velocity with an accuracy of 3 to 4 orders of magnitude is possible precisely by utilizing a large number (10^3 to 10^4) of spectral lines.

State of Matter and Its Composition

The spectrum of stars is determined by the interaction of radiation coming from inside the star and the matter that forms the star. In the spectrum, we thus see the spectral lines of atoms and molecules that make up stars. An exact description is complex and requires knowledge of the electronic structure of substances given by quantum physics (energies and probabilities of transitions between energy levels), statistical physics (describing the occupation of individual levels), and the theory of radiative transfer. Let us therefore describe the observed spectrum of stars qualitatively.

In practice, a comparison of the observed spectrum with a theoretically calculated one is used to determine the parameters. The presence of spectral lines depends both on their inherent strength and on the number of particles radiating at the given wavelength. With increasing temperature, the energy of the levels occupied in the substance increases. For the coolest stars with temperatures below 3 000 K, bands of titanium and vanadium oxides dominate in the spectra. At higher temperatures, the dissociation of these molecules occurs, and lines of metals such as magnesium, sodium, or iron prevail in the spectra. Stars similar to the Sun with temperatures around 5 000 K thus look like balls of metal at first glance. At higher temperatures, lines of ionized metals (they have maximum intensity around 6 000 K) and hydrogen lines (maximum intensity around 9 000 K) begin to appear in the spectra.

For temperatures in the range approximately from 20 000 K to 30 000 K, the lines of neutral helium are dominant, which has the highest ionization energy among neutral atoms; and hydrogen lines gradually disappear, since hydrogen is ionized into a proton and electron pair, which do not have line spectra. Other chemical elements can be found in states with a high degree of ionization. For the hottest stars with temperatures above 40 000 K, even helium is ionized. By comparing the depths of the lines of differently ionized/excited elements with theoretical values, it is then possible to determine the temperature of the star and the abundance of individual chemical elements. Besides rotation, spectral lines are also broadened by thermal motion and mutual collisions. From the shape of the hydrogen spectral lines[11], it is therefore possible to determine the gravitational acceleration on the surface, which, through pressure, determines the frequency of collisions. Broadening can also be caused by a magnetic field; for strong fields, even the splitting of lines by the Zeeman effect occurs. To determine magnetic fields, spectra obtained in polarized light are used, where the observed signal is directly proportional to the magnetic field.

Conclusion

This part was largely dedicated to the optical principles of instruments used in astronomy. In the serial problem, you will have the opportunity to try out a (rough) design of an echelle spec-

[11] As the lightest element, it is most affected by collisions.

[12] Available from `[https://commons.wikimedia.org/wiki/File:Spectral](https://commons.wikimedia.org/wiki/File:Spectral)_lines_in_spectral_classes.svg`.

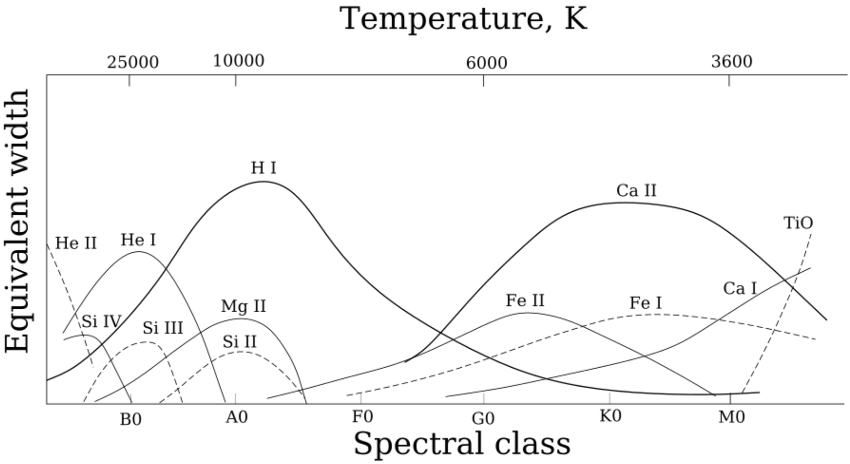

Figure 6: The evolution of the strength (equivalent width) of spectral lines in the spectrum of stars as a function of their effective temperature.[12]

trograph. Modern astronomy is based precisely on opto-mechanically high-quality instruments; I therefore hope that this part provided a more advanced demonstration of the use of optics subject matter.

A Few Words on This Year's Serial

Over the course of the entire year, you had the opportunity to peek into the amazing world of experimental physics. We have revealed how to measure extremely small and extremely distant things, and how to even find out what the world around us is composed of and how objects interact within it. In the problems, you tried what it is like to interpret measured data from real experiments, and found out what information is hidden within them.

However, the beauty of experimental physics is not limited to the ability to measure something no one before you has ever noticed. You can gain access to the most modern instruments and technologies that utilize laboriously discovered physical principles and operate on the very edge of what humanity can build. Of course, this can be accompanied by various malfunctions and subsequent repairs of instruments or some DIY improvements; from a technical and engineering perspective, however, it is undoubtedly amazing work, to which one must take off their imaginary hat.

Of course, it was not possible to explain all aspects of the given experiments in a few pages, let alone to go deep into the discussed techniques. You as the solvers might have even been slightly lost in the multitude of new concepts. However, the goal of these texts was not to teach you specific techniques or how to properly process data, but rather we wanted to show you the boundaries of what can be found out about the world around us, and to broaden your horizons on how new knowledge is acquired. It would be a great satisfaction for us if our serial ever helped you find yourselves in connection with experimental physics, or even if our demonstration of some physics branches helped you in your future choices. In any case, we hope that wherever your life's path takes you, you will already have an idea of how physicists get to know the world.

Jozef Lipták
`liptak.j@fykos.org`

FYKOS is organized by students of Faculty of Mathematics and Physics of Charles University. It's part of Media Communications and PR Office and is supported by Institute of Theoretical Physics of CUNI MFF, his employees and The Union of Czech Mathematicians and Physicists. The realization of this project was supported by Ministry of Education, Youth and Sports of the Czech Republic.