Simple derivation of the Doppler effect

Klaus Kassner 25 July 2024

I was surprised to discover that in a question on Research Gate posed back in 2015 and titled "Is the doppler effect of light an actual energy shift of photons or it is only a relativistic connection of different reference frames?" Stefano Quattrini supported the view that the Doppler effect is based on energy/momentum conservation and results from an energy shift transmitted by an absorber to the photon on absorption (obviously he considered the case where the absorber was moving towards the light source, not away from it), rather than being just due to the fact that the frequency of a wave is normally different in two reference frames moving with respect to each other (which is even true outside of relativity). Now, I don't know of any derivation of the velocity Doppler effect based on energy conservation. Rather, it is the Compton effect that is usually derived via energy-momentum conservation, but that is an entirely different matter. The Compton effect is a quantum mechanical phenomenon, usually discussed in terms of the particle picture for photons. On the other hand, the Doppler effect is a genuine wave phenomenon, the notion of *photon* is not even needed in its discussion.

I was even more flabbergasted in finding out that Quattrini had not bothered to learn enough about the physics of the Doppler effect in the last nine years to have corrected his opinion. Quite to the contrary, he considered the discussion "closed" in favor of his original point of view. That is pretty sad...

It is easy to prove that his point of view is wrong. An elementary derivation of the effect without reference to energy or momentum conservation is available that he should have been able to develop himself, just by looking at what goes on physically. I will give such a derivation here, for the benefit of those followers of the question that may be confused by the contradictory and sometimes weird lines of argument that can be found in the long history of this question. The issue is simple enough that being strong at logic is not necessary to achieve complete understanding. There are more elegant and more general derivations of the Doppler effect, based on the invariance properties of the phase of the wave, but for readers not familiar with that kind of reasoning, they may be less convincing than a physics first approach.

For reasons of transparency, I will first derive the longitudinal Doppler effect for acoustic waves, where what happens physically is quite evident and then show the – minor – modifications needed to arrive at a correct description for electromagnetic waves. To keep things simple, I will consider only the situation where the relative velocity between the emitter and receiver is aligned with the straight line connecting the two (whence longitudinal Doppler effect).

In the acoustic case, there is a physically preferred frame of reference, to which all velocities can be referred – that is the rest frame of the medium carrying the sound; this could be air, for example.[1] Then two situations are of particular interest: the case of a moving emitter and resting receiver ("listener") and the case of a resting emitter and moving receiver.

It is helpful to visualize these situations by drawings, because that gives an immediate idea why there is a Doppler effect and provides an approach of how to express it in a mathematically exact way. In Fig. 1, I have assumed that the emitter is a point source, so it emits a spherical wave, constant-phase positions of which reduce to circles in the picture. However, as I only wish to discuss the frequency and wavelength shifts along the straight line connecting emitter and receiver, I might as well have considered an emitter that is extended in the directions

[1] The velocity of this medium can be measured by physical devices, in any frame, this way determining the velocity of the measuring frame with respect to the preferred one.

vertical to that line, so the wave would become planar (but intersection points of constant-phase surfaces with the connecting line remain the same). The results would be unaffected, but the picture would only look half as nice...

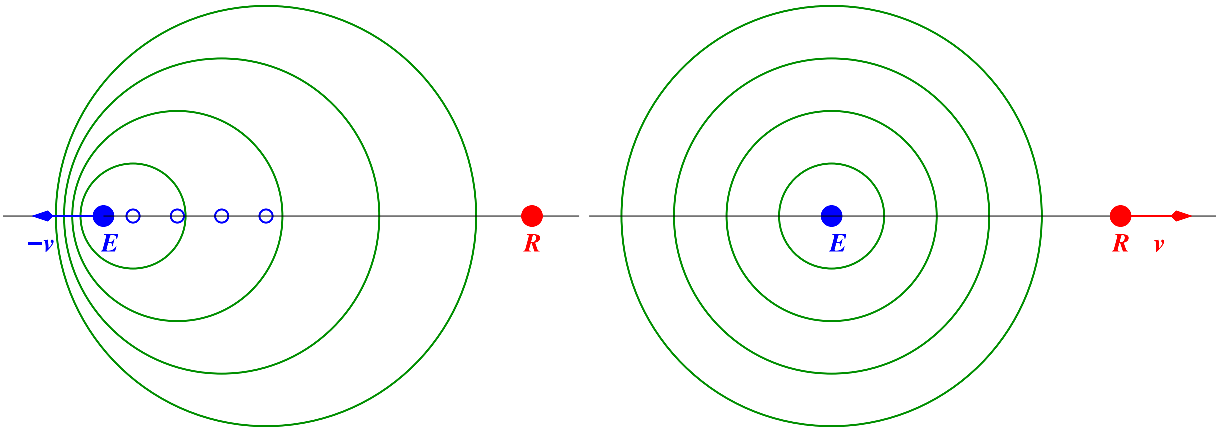

Fig. 1:
Left: The emitter E moves at velocity $-v$ (i.e., to the left). The receiver R is at rest.
Right: The emitter E is at rest. The receiver R moves at velocity v (i.e., to the right).
Green circles indicate the instantaneous positions of some successive wave crests. Open blue circles in the left image indicate the (earlier) positions of the emitter when it emitted the wave crest at the center of which the blue circle is located.

From the left panel of the figure we can essentially read off, *why* there is a Doppler effect, when the emitter is moving. Due to the motion of the emitter, successive wave crests are not emitted concentrically, so the wavelength is increased for the wave part emitted opposite to the motion of the emitter and decreased for the wave part emitted into the half space to which the emitter velocity points. The speed of sound c_s does not depend on the speed of its source, so the wavelength change means a decrease in frequency in the half space to the right of the emitter and an increase in frequency in the half space to its left. It is very easy to make the relation quantitative for a wave traveling along the line ER. Let $\Delta\tilde{t}$ be the time between the emission of two successive wave crests, then for an emitter at rest the wavelength will simply be $\tilde{\lambda} = c_s\Delta\tilde{t}$, because this is the distance one crest will move before the next is emitted, afterwards they will continue to move at the same speed, and the distance between two successive wave crests is, by definition, the wavelength. If the emitter moves at velocity $-v$, then the second crest is created at the relative position $-v\Delta\tilde{t}$, when the first reaches relative position $c_s\Delta\tilde{t}$, so the wavelength becomes $(c_s + v)\Delta\tilde{t}$, and this is $\lambda = (1 + v/c_s)\tilde{\lambda}$.

The frequency received by R then is $\nu = c_s/\lambda = c_s/((c_s + v)\Delta\tilde{t})$. Obviously, the frequency ν' emitted by E is the inverse of the time interval between the emission of two crests, i.e. $\nu' = 1/\Delta\tilde{t}$ (again by definition: the frequency is the number of oscillations per time and $\Delta\tilde{t}$ is the time for one oscillation). Combining the formulas for ν and ν', we obtain

$$\boxed{\nu = \frac{\nu'}{1 + v/c_s}}, \tag{1}$$

which is the correct (and well-known) result for the acoustic Doppler effect in the case where the source moves away from the receiver.[2]

Let us now consider the case depicted in the right panel of Fig. 1, of an emitter that is at rest in the sound-propagating medium, while the receiver moves (to the right) at velocity v. With

[2] We get the result for a source moving towards the receiver either by noting that the wavelength of the left-moving wave in the left panel is $(1 - v/c_s)\Delta\tilde{t}$ or, formally, by replacing v with $-v$ in Eq. (1). The result is $\nu = \nu'/(1 - v/c_s)$.

the source at rest, the wavelength of the emitted concentrical spherical wave is $\lambda' = c_s/\nu'$. But the time interval between two successive crests arriving at the position of R is not λ'/c_s as it would be, if the receiver was at rest. Because the receiver is moving, the wave crests have to "chase" after it. So the time interval $\Delta\hat{t}$ that the second of two successive crests takes to also catch up with the receiver (after the first has reached it), is given by

$$c_s \Delta\hat{t} = \lambda' + v\Delta\hat{t} \quad \Rightarrow \quad \Delta\hat{t} = \frac{\lambda'}{c_s - v} = \frac{1}{\nu'(1 - v/c_s)} \,. \tag{2}$$

Now, the inverse of this time interval obviously is the frequency "observed" by the receiver, hence[3]

$$\boxed{\nu = \left(1 - \frac{v}{c_s}\right) \nu'} \,. \tag{3}$$

Note that the wavelength remains λ' in this second case (the wave crests emitted are all concentric), but the frequency changes, because the wave velocity is different from c_s for R. We have $c_{sR} = \nu\lambda' = (1 - v/c_s)\,\nu'\lambda' = c_s - v$.

So a pretty clear picture emerges for the Doppler effect for sound. The Doppler effect in the case of a moving emitter is due to a wavelength shift caused by the displacement of the emitter during the time of an oscillation, at constant speed of sound. Modification of the wavelength at constant wave speed leads to a modification of the frequency. In the case of a moving receiver, the wavelength is unaffected, but the frequency of reception of successive wave crests changes, because the wave must either run after the receiver or is moving towards it. This leads to a change of the effective wave speed (which is either $c_s - v$ or $c_s + v$), so a frequency change is compatible with a fixed wavelength.

A few points are noteworthy. Of course, the two situations are physically different. We have different formulas for the Doppler effect with a moving emitter and that with a moving receiver. (The case where both are moving, can be easily treated as a combination of both effects.) The derivation does not require any reference to energy or momentum conservation, it is entirely based on wave geometry and kinematics. In the case of the moving emitter, the wavelength change is evidently present everywhere between the emitter and the receiver, it does not happen only on absorption. If the receiver is moving, the frequency of the wave is not changed on absorption from what it has been reduced to due to the "catch up" effect. By the way, Quattrini's idea of a "small" frequency change on absorption is completely off here. The formulas containing $1 - v/c_s$ are valid for all velocities in the interval $[0, c_s)$ and may result in enormous frequency shifts. The formulas containing $1 + v/c_s$ are even valid for supersonic velocities v.[4]

Now that we completely have understood the (longitudinal) Doppler effect for sound, let us try how far we get with the same way of reasoning in the case of electromagnetic waves in vacuum. As it will turn out, the derivation is achievable with just two technical modifications: the speed of sound has to be replaced with the vacuum speed of light and the effect of time dilation between the emitter and the receiver has to be figured in.

At first sight, a conceptual problem may seem to arise due to the fact that in the case of light we do not have a physically preferred frame of reference. Light waves travel through vacuum[5] which does not have a well-defined state of motion that might render possible the

[3] And for a receiver moving towards the emitter, we have $\nu = (1 + v/c_s)\,\nu'$, $(v > 0)$.

[4] It is easy to verify that Eq. (1) holds for a supersonic emitter. The wave moving along ER to the right is not qualitatively different, whether there is a Mach cone to the left or not.

[5] A property they share with a number of other wave phenomena, such as gravitational waves or matter waves.

unique definition of a velocity with respect to that state. Nor does the ether that some people feel to be necessary as a medium supporting electromagnetic waves[6] provide such a definition; since it is not detectable experimentally, it does not give us a reference frame that would be picked out by physical laws.

However, the absence of a preferred frame of reference need not deter us, because in each of the two scenarios for sound, we had a particular object at rest with respect to the preferred frame. In the left panel that was the receiver, in the right one the emitter. Therefore, rather than describing the two cases in terms of a "third-party" frame of reference, we ascribe the left panel to an *emitter E moving in the frame of reference of the receiver R* and the right one to a *receiver R moving in the frame of reference of the emitter E*. With the velocities chosen (and the standard notion of simultaneity), the relativity principle then tells us that the two situations are physically identical. But we are not obliged to take this into consideration. We may treat the two cases as physically distinct, and should we find their Doppler effects to be different, then we would have proven them to be different and the relativity principle not to apply. Let us see what happens...

Starting with the case of a moving emitter E, we note that the wave crests still lie on non-concentric spheres, and the wavelength of the right-moving wave is $(c+v)\Delta\tilde{t}$, where c is the vacuum speed of light and $\Delta\tilde{t}$ is the time between the emission of two crests *in the frame of the receiver R*. But the *emission frequency* will not be $1/\Delta\tilde{t}$, because the time interval between the emission of two successive crests is different from $\Delta\tilde{t}$ *in the frame of the emitter* (and that is of course the frame relevant for the emission frequency). In fact, we know from the standard result for time dilation between two inertial frames of reference in special relativity that if $\Delta t'$ is the time interval between the emission of two wave crests in the emitter frame (meaning that $\nu' = 1/\Delta t'$), then we have $\Delta t' = \Delta\tilde{t}/\gamma(-v) = \sqrt{1-v^2/c^2}\Delta\tilde{t}$.[7] We obtain $\lambda' = c\Delta t'$ and

$$\nu = \frac{c}{(c+v)\Delta\tilde{t}} = \frac{c}{(c+v)\Delta t'/\sqrt{1-v^2/c^2}} = \frac{\sqrt{1-v^2/c^2}}{1+v/c}\frac{1}{\Delta t'} = \left(\frac{1-v/c}{1+v/c}\right)^{1/2}\nu'$$

$$\boxed{\nu = \left(\frac{1-v/c}{1+v/c}\right)^{1/2}\nu'}\,. \tag{4}$$

This is a minor modification with respect to the sound result.[8] We can attribute the Doppler effect in the case of a moving emitter to a shift in the wavelength plus a frequency shift due to time dilation.

Let us now consider the case of a moving receiver. Again, we construct the formula along the

[6] Maxwellian electrodynamics shows that electromagnetic fields can carry energy density and that the electromagnetic field is capable of supporting stresses – this is described by the Maxwell stress tensor. Since the purpose of the medium in general wave phenomena is to support density and pressure variations giving rise to oscillatory behavior, it is not clear what an ether would be needed for in the electromagnetic case, where energy density and pressure (= negative stress) are already supported by the field itself. Moreover, it seems that proponents of a Lorentzian ether which would be the carrier medium for electromagnetic waves, hardly ever care for the question of a gravitational ether that would be necessary in addition in order to support gravitational waves. Nor does the question of an ether necessary for each matter field seem to ever cross their minds, probably because they always consider matter consisting of particles, in spite of the interfereence experiments successfully performed with electrons, neutrons and even buckyballs.

[7] It may be useful to recapitulate, how much of special relativity goes into this argument and what elements of the theory are not used. The formula for time dilation is derivable from the universality of the one-way vacuum speed of light alone. (I will show that in a different essay.) That is, the relativity principle is not needed, but the time coordinate used is Einstein synchronized. Indeed, time dilation of clocks kept at a distance is synchronization dependent (whereas the compounded time dilation of the twin paradox is not dependent on synchronization, because in it the two clocks are together at both events where time is measured).

[8] The result for an emitter moving *towards* the receiver is again obtained by replacing v with $-v$ in the formula.

argument of the sound case. First, the analog of (2) becomes

$$c\Delta\hat{t} = \lambda' + v\Delta\hat{t} \quad \Rightarrow \quad \Delta\hat{t} = \frac{\lambda'}{c-v} = \frac{1}{\nu'(1-v/c)} \ , \tag{5}$$

where $\Delta\hat{t}$ is the time interval between two successive crests being detected by R, but it is a time interval *measured in the emitter frame.* Since the receiver is moving in that frame, it observes a shorter time interval $\Delta t = \Delta\hat{t}/\gamma(v) = \sqrt{1-v^2/c^2}\Delta\hat{t}$, and we obviously have $\nu = 1/\Delta t$. Hence,

$$\nu = \frac{\nu'(1-v/c)}{\sqrt{1-v^2/c^2}} = \nu'\sqrt{\frac{1-v/c}{1+v/c}}$$

$$\boxed{\nu = \left(\frac{1-v/c}{1+v/c}\right)^{1/2} \nu'} \ . \tag{6}$$

Here, the Doppler effect is due to two frequency shifts, one coming from time dilation (increasing the frequency at the receiver end), the other from the fact that the observer moving away from the emitter will receive wave crests at a decreased rate.[9]

The result does not give us any indication for a violation of the relativity principle, rather suggests its confirmation, especially in view of the fact that, contrary to the case of sound, the propagation speed of the wave is the same in both frames and, hence, there is an analogous formula to Eqs. (4) and (6) for the Doppler wavelength change:[10]

$$\lambda = \left(\frac{1+v/c}{1-v/c}\right)^{1/2} \lambda' \ . \tag{7}$$

It may be difficult to visualize that for light traveling through vacuum the two wave patterns depicted in Fig. 1 actually correspond to the same wave. Yet, this is a straightforward consequence of the universality of the vacuum speed of light, implying that light waves from a point source must, in every inertial frame of reference, be spherical. So the only possible difference in different frames lies in whether successive surfaces of equal phase are concentric or not (and by how much the centers of these spherical surfaces are shifted between two neighboring ones).

Some of the confusion in trying to picture this may arise due to the fact that we expect moving spherical bodies to be distorted by Lorentz contraction. However, there is a difference between a solid moving shape and the geometry of a wave front. This can be most easily seen, if we imagine our emitter to be at the center of a thin spherical shell made of fluorescent material and assume it to send off a short light pulse. The spherical wave will expand until it hits the sphere and there it will trigger the emission of fluorescence radiation (having a lower frequency than the light of the original pulse). The whole shell will start fluorescing at one moment of time.

Now consider the same scenario as a moving observer. Then the shell will be Lorentz contracted along the direction of movement, so it has the shape of an oblate ellipsoid. The light pulse emitted at its interior expands as a spherical wave, part of which will first hit one of the "short ends" of the ellipsoid and trigger fluorescence there. Rather than the whole material surface emitting fluorescent light, the fluorescence pattern will spread as a ring from the point where it first appeared; a little later it will appear on the opposite side of the ellipsoid and

[9] Again the case of an observer moving towards the emitter is deduced by replacing v with $-v$ in the formula.

[10] Remember that in the case of sound, the wavelength is unchanged for a moving receiver. In the case of electromagnetic waves in vacuum, it changes the same way for a moving receiver as for a moving emitter.

expand as a ring from there, too. The two rings will move towards the large circumference of the ellipsoid, join there, and the emission of fluorescence will then end. That is, whereas fluorescence happens at one point in time in the frame of the emitter, it spreads over a time interval in the frame of the observer.[11]

Now, this is what we should expect from the relativity of simultaneity anyway. In the emitter frame, there is an instant, where the shape of the shell and the wave front coincide, whereas in the observer frame, the shape of the shell is constant and an ellipsoid, whereas the shape of the wave front is an expanding sphere; during a certain time interval, the two shapes intersect (in circular, i.e. ring-shaped regions), but they never coincide fully. This is clearly consistent with the predictions of special relativity.

Returning to the Doppler effect, our derivation was purely geometric and kinematical, there was no necessity of invoking any conservation law. Nevertheless, I would like to give some physical arguments in addition, demonstrating the idea to be wrong that the frequency shift somehow has to do with an energy exchange during the absorption process.

For concreteness, let us assume that in the scenario of Fig. 1, the wavelength emitted is 450 nm (corresponding to blue light and a frequency of $6.67\times10^{14}\mathrm{s}^{-1}$) and the speed of the observer moving away from the emitter is $0.35\,c$. This leads to a Doppler shifted wavelength of 650 nm (corresponding to red color and a frequency of $4.62\times10^{14}\mathrm{s}^{-1}$) at the receiving end. Now let us put a color filter between the emitter and the receiver and let it be at rest with respect to the receiver. Suppose the filter lets almost 100% of the light through in the wavelength range 640 nm to 660 nm, but essentially nothing at frequencies outside this band.

Will the observer see light from the emitter? Our derivation clearly says yes. Since the filter is moving precisely as the observer, it will "see" the Doppler shifted wavelength, which is 650 nm and therefore goes through the filter. According to Quattrini, the frequency shift comes about by absorption, so the light should arrive with a wavelength of 450 nm at the filter, which means, it will not go through and be completely absorbed, even before reaching the receiver.[12] Our derivation tells us that in the frame of the receiver, the wavelength is Doppler shifted before any interaction with the receiver.

If we use a filter instead that is at rest with respect to the emitter, it will have to be transparent to wavelengths around 450 nm, in order for light to get to the receiver, showing that in the frame of the emitter the wavelength is 450 nm on the whole segment between the emitter and the receiver.

Clearly, the frequency and the wavelength of an electromagnetic wave depend on the frame of reference, in which they are to be determined. This should not be too surprising. The frequency of a wave is defined as the number of oscillations it performs per time unit at a fixed position. (For a frequency-modulated signal, the frequency varies between different positions.) But the meaning of "fixed position" is frame dependent. A fixed position in the emitter frame is a variable position in the receiver frame (it moves at velocity $-v$). A fixed position in the receiver frame is variable in the emitter frame (it moves at velocity v). If relativistic speeds are considered, then in addition the time unit is frame dependent. It is then evident that photon

[11] What seems to happen *visually*, is yet another thing. Any observer will see, at one point in time, the light that *arrives* at his eye at that time, *not* light that was *emitted* at one point in time. A known consequence is that the Lorentz contraction is not visible as a contraction. Rather, for distant objects, it will make them appear rotated. Hence, a sphere will not appear as the ellipsoid it is when moving, but as a sphere (with changed radius). Images on its surface will appear distorted. There has been a strong research activity, originally in Tübingen (in the group of Hanns Ruder), leading to the computation of visual aspects of many fast-moving objects according to special relativity: `https://www.spacetimetravel.org/tompkins`.

[12] The idea of using a filter is of course motivated by the fact that we know the frequency of light *not* being absorbed by it. And it is this frequency that is Doppler shifted, so the shift cannot be due to absorption.

frequencies are frame dependent (similar to quantities such as the kinetic energy, which is also frame dependent).

Finally, I would like to debunk Quattrini's absorption argument in a different frequency range. Our derivation is universal, i.e., it works for arbitrary frequencies. Quattrini's absorption argument does not. Consider electromagnetic waves with radar frequencies, i.e., in the range between 1 and 100 GHz. In this frequency range, there do not exist single-photon detectors. Detectors are essentially antennas and they work by the electromagnetic field making the electrons in the metal oscillate along with the frequency of the infalling wave. That produces an oscillatory electric current that can be used for detection. The absorption process is described by classical electrodynamics and results in taking energy from the electromagnetic field, thus reducing its amplitude – without changing its frequency. Also the momentum of the electromagnetic field is reduced (as can be seen from its Poynting vector), but again without any change in the energy of the single photons. The energy and momentum transfer is always that of a huge number of photons and proportional to that number. The photons are converted into electronic excitations of the detector metal, with the same frequency. Hence, the absorption process is not accompanied by a frequency change (as it might be at higher frequencies, where quantum mechanical processes such as the Compton effect may play a role). Absorption works very differently in different frequency domains, the Doppler effect always works the same (the factor between the emitted and received frequencies is frequency independent). The absorption argument as an explanation of the Doppler shift therefore fails.

Lastly, I would like to give a result that you will hardly find in the literature and probably have not seen before. It is also new to myself and I do not have independent sources for its verification. But I trust in the correctness of the approach that I have given here (which provides a straightforward derivation of that result). What I am talking about is *the relativistic correction to the Doppler effect for sound*. Since my derivation of the sonic Doppler effect was based on Newtonian mechanics, it did not contain relativistic effects. In general, these are negligible, because the sound velocity is very small in comparison with the speed of light. So it is not normally worthwhile to derive them. But it should be clear how to do it, hence I will give the result. All that has to be done, is taking into account time dilation in addition to the wavelength and frequency effects in the cases of a moving emitter and receiver, respectively. This provides us with

$$\nu = \frac{\sqrt{1 - v^2/c^2}}{1 + v/c_s}\, \nu' \,, \qquad \text{(moving emitter)} \tag{8}$$

$$\nu = \frac{1 - v/c_s}{\sqrt{1 - v^2/c^2}}\, \nu' \,. \qquad \text{(moving receiver)} \tag{9}$$

Of course, in all everyday applications, we have $v/c \ll v/c_s$, so the correction in comparison with Eqs. (1,3) is tiny. However, if we were to consider the sonic Doppler effect inside a neutron star, we could easily have $c_s = c/2$, so the frequency change of a seismic wave during a starquake, while it moves from the pole to the equator (where the surface is spinning very fast, up to hundreds of revolutions per second) or vice versa, would have to be calculated taking relativistic effects into account.