Some sins in physics didactics

Hendrik van Hees

May 9, 2015

1 Introduction

There are many sins in physics didactics. Usually they occur, because teachers, professors, textbook or popular-science-book writers, etc. try to simplify things more than possible without introducing errors in reasoning, or they copy old-fashioned methods of explaining an issue, leading to the necessity to "erase" from the students' heads what was hammered in in a careless way before. Some examples are the introduction of a velocity-dependent mass in special relativity, which is a relic from the very early years after Einstein's ground-breaking paper of 1905, the use of Bohr's atomic model as an introduction to quantum theory, which provides not only quantitatively but even qualitatively wrong pictures about how an atom is understood nowadays in terms of "modern quantum theory". In this blog, I like to address some of these questionable cases of physics didactics. Of course, this is a quite subjective list of "sins".

For each case, I'll first give a rather non-technical review, which should be understandable by a high-school student. Then I'll give a more technical description of the point of view of contemporary (theoretical) physics.

2 The photoelectric effect and the abuse of the notion of photons

Particularly seductive is quantum theory to the well-intentional teacher. This has several reasons. First of all it deals with phenomena at atomic or even subatomic scales that are not within our daily experience, and this realm of the natural world can described only on quite abstract levels of mathematical sophistication. So it is difficult to teach quantum theory in the correct way, particularly on an introductory level, let alone on a level understandable to lay people.

In this blog I address readers who are already familiar with modern nonrelativistic quantum theory in terms of the Dirac notation.

2.1 Historical development

Often introductory texts on quantum theory start with a heuristic description of the photoelectric effect, inspired by Einstein's famous paper on the subject (1905). There he describes the interaction of light with the electrons in a metallic plate as the scattering of "light particles", which have an energy of $E = \hbar\omega$ and momentum $\vec{p} = \hbar\vec{k}$, where $\hbar$ is the modified Planck constant, ω the frequency of monochromatic light and $\vec{k}$ the wave number.

To kick an electron out of the metal one needs to overcome its binding energy W, and the conservation of energy thus implies that the kicked-out electrons have a maximal energy of

$$E = \hbar\omega - W, \tag{1}$$

and this formula is often demonstrated by letting the photo-electrons run against an electric field, which just stops them, and measuring the corresponding stopping voltage as a function of the light's frequency ω nicely confirms Einstein's Law.

After Planck's discovery and statistical explanation of the black-body-radiation law in 1900, this work of Einstein's started the true quantum revolution. Planck's derivation was already mind-puzzling enough, because he realized that he had to assume that electromagnetic radiation of frequency ω can only be absorbed in energy portions of the size $\hbar\omega$. In addition he had to apply a pretty strange method to count the number of microstates for the given macroscopic situation of radiation at a fixed temperature in a cavity in order to use Boltzmann's famous relation between the entropy and this number of microstates, which in fact was written down first by Planck himself in explicit terms: $S = k_B \ln \Omega$, where Ω is the number of microstates.

Although already this was breaking with the classical picture, and Planck tried to "repair" this radical consequences of his own discovery till the very end of his long life, Einstein's paper was much clearer about how deep this departure from the principles of classical physics indeed was. First of all Einstein (re)introduced the idea of a particle nature of light, which was abandoned pretty much earlier due to the findings of wavelike phenomena like interference effects as in Young's famous double-slit experiment, demonstrating the refraction of light. Finally, Maxwell's theory about electromagnetism revealed that light might be nothing else than waves of the electromagnetic field, and H. Hertz's experimental demonstration of electromagnetic waves with the predicted properties, lead to the conviction that light indeed is an electromagnetic wave (in a certain range of wavelengths, the human eye is sensitive to).

Second, Einstein's model (which he carefully dubbed a "heuristic point of view" in the title of the paper) introduced wave properties into the particle picture. Einstein was well aware that this "wave-particle duality" is not a very consistent description of what's going on on the microscopic level of matter and its interaction with the electromagnetic field.

Nevertheless, the wave-particle duality of electromagnetic radiation was an important step towards the modern quantum theory. In his doctoral dissertation L. de Broglie introduced the idea that wave-particle duality may be more general and may also apply to "particles" like the electron. For a while it was not clear what the stuff in vacuum tubes might be, particles or some new kind of wave field, until in 1897 J. J. Thomson could measure that the corresponding entity indeed behaves like a gas of charged particles with a fixed charge-mass ratio by studying how it was moving in electro- and magnetostatic fields.

All these early attempts to find a consistent theory of the microcosm of atoms and their constituents were very important steps towards the modern quantum theory. Following the historical path, summarized above, the break through came in 1926 with Schrödinger's series of papers about "wave mechanics". Particularly he wrote down a field-equation of motion for (nonrelativistic) electrons, and in one of his papers he could solve it, using the famous textbook by Courant and Hilbert, for the stationary states (energy eigenstates) of an electron moving in the Coulomb field of the much heavier proton, leading to an eigenvalue problem for the energy levels of the hydrogen atom, which where pretty accurate, i.e., only lacking the fine structure, which then was thought to be a purely relativistic effect according to Sommerfeld's generalization of Bohr's quantum theory of the hydrogen atom.

Now the natural question was, what the physical meaning of Schrödinger's wave function might be. Schrödinger himself had the idea that particles have in fact a wavy field-like nature and might be "smeared out" over finite regions of space rather than behaving like point-like bullets. On the other hand, this smearing was never observed. Free single electrons, hitting a photo plate, never gave a smeared-out pattern but always a point-like spot (within the resolution of the photo-plate, given by the size of the grains of silver salt, e.g., silver nitrate). This brought Born, applying Schrödinger's wave equation to describe the scattering of particles in a potential, to the conclusion that the square of the wave function's modulus, $|\psi(\vec{x})|^2$, gives the **probability density** to find an electron around the position $\vec{x}$.

A bit earlier, Heisenberg, Born, and Jordan had found another "new quantum theory", the "matrix mechanics", where the matrices described transition probabilities for a particle changing from one state of definite energy to another. Heisenberg had found this scheme during a more or less involuntary holiday on the Island of Helgoland, where he moved from Göttingen to escape his hay-fever attacks, by analyzing the most simple case of the harmonic oscillator with the goal to use only observable quantities and not theoretical constructs like "trajectories" of electrons within an atom or within his harmonic-oscillator potential. Back home in Göttingen, Born quickly found out that Heisenberg had reinvented matrix algebra, and pretty rapidly he, Jordan, and Heisenberg wrote a systematic account of their new theory. Quickly Pauli could solve the hydrogen problem (also even before Schrödinger with his wave mechanics!) within the matrix mechanics.

After quarter of a century of struggle of the best theoretical physicists of their time to find a consistent model for the quantum behavior of microscopic particles, all of a sudden one had not only one but even two of such models. Schrödinger himself could show that both schemes were mathematically equivalent, and this was the more clear, because around the same time another young genius, Dirac, found another even more abstract mathematical scheme, the so-called "transformation theory", by introducing non-commuting "quantum numbers" in addition to the usual complex "classical numbers", which commute when multiplied together. The final step for the complete mathematical resolution of this fascinating theory came with a work by von Neumann, who showed that states and observables can be described as vectors in an abstract infinite-dimensional vector space with a scalar product, a so-called Hilbert space (named after the famous mathematician) and so-called self-adjoint operators acting on these state vectors.

In the next section we shall use this modern theory to show, what's wrong with Einstein's original picture and why it is a didactical sin to claim the photoelectric effect proves the quantization of the electromagnetic field and the existence of "light particles", now dubbed **photons**.

2.2 Modern understanding of the photoelectric effect

Let us discuss the photoelectric effect in the most simple approximation, but in terms of modern quantum theory. From this modern point of view the photoelectric effect is the induced transition of an electron from a bound state in the metal (or any other bound system, e.g., a single atom or molecule) to a scattering state in the continuous part of the energy spectrum. To describe induced transitions, in this case the absorption of a photon by an atom, molecule, or solid, we do not need to quantize the electromagnetic field at all but a classical electromagnetic wave will do, which we shall prove now in some detail.

The bound electron has of course to be quantized, and we use the abstract Dirac formalism to describe it. We shall work in the interaction picture of time evolution throughout, with the full bound-state

Hamiltonian,

$$\hat{H}_0 = \frac{\hat{\vec{p}}^2}{2\mu} + V(\hat{\vec{x}}), \tag{2}$$

which we have written in terms of an effective single-particle potential, leading to bound states $|E_n, t\rangle$, where n runs over a finite or countable infinite number (including possible degeneracies of the energy spectrum, which don't play much of a role in our treatment) and a continuous part $|E, t\rangle$ with $E \geq 0$. It is important to note that in the interaction picture the eigenvectors of operators that represent observables are time dependent, evolving with the unperturbed Hamiltonian, which is time-independent in our case, according to

$$|o, t\rangle = \exp\left[\frac{\mathrm{i}}{\hbar}(t - t_0)\hat{H}_0\right]|o, t_0\rangle. \tag{3}$$

For the eigenvectors of the unperturbed Hamiltonian this implies

$$|E, t\rangle = \exp\left[\frac{\mathrm{i}}{\hbar}(t - t_0)E\right]|E, t_0\rangle. \tag{4}$$

The operators which represent observables themselves move accordingly as

$$\hat{O}(t) = \exp\left[\frac{\mathrm{i}}{\hbar}(t - t_0)\hat{H}_0\right]\hat{O}(t_0)\exp\left[-\frac{\mathrm{i}}{\hbar}(t - t_0)\hat{H}_0\right]. \tag{5}$$

The classical radiation field is for our purposes best described by an electromagnetic four-vector potential in the non-covariant radiation gauge, i.e., with

$$A^0 = 0, \quad \vec{\nabla} \cdot \vec{A} = 0. \tag{6}$$

Then the electromagnetic field is given by

$$\vec{E} = -\frac{1}{c}\partial_t \vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A}. \tag{7}$$

This field is coupled to the particle in the minimal way, i.e., by substitution of

$$\hat{\vec{p}} \rightarrow \hat{\vec{p}} + \frac{e}{mc}\hat{\vec{A}} \quad \text{with} \quad \hat{\vec{A}} = \vec{A}(t, \hat{\vec{x}}) \tag{8}$$

in (2). For a usual light wave we can assume that the corresponding field is very small compared to the typical field the electron "feels" from the binding potential. Thus we can restrict ourselves to the leading linear order in the perturbation $\vec{A}$. We can also assume that a typical electromagnetic wave has much larger wavelengths than the dimensions of the typical average volume the electron is bound to within the atom, i.e., we can take

$$\hat{\vec{A}} \simeq \vec{A}(t) = \vec{A}_0 \cos(\omega t) = \frac{\vec{A}_0}{2}[\exp(\mathrm{i}\omega t) + \exp(-\mathrm{i}\omega t)]. \tag{9}$$

Then $\vec{A}$ is a pure external c-number field and commutes with $\hat{\vec{p}}$. To linear order the perturbation ("interaction") Hamiltonian thus reads

$$\hat{H}_{\mathrm{I}} = \frac{e}{mc}\vec{A} \cdot \hat{\vec{p}}. \tag{10}$$

Now in the interaction picture the equation of motion for the state vector of the electron reads

$$\mathrm{i}\hbar\partial_t|\psi(t)\rangle = \hat{H}_{\mathrm{I}}|\psi(t)\rangle. \tag{11}$$

The formal solution is the time-ordered exponential [see any good textbook on quantum theory, e.g., J. J. Sakurai, Modern Quantum Mechanics, 2nd Edition, Addison Wesley (1994)],

$$|\psi(t)\rangle = \hat{C}(t,t_0)|\psi(t_0)\rangle, \quad \hat{C}(t,t_0) = \mathscr{T}\exp\left[-\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}\mathrm{d}t'\hat{H}_{\mathrm{I}}(t')\right]. \tag{12}$$

In leading order the exponential reads

$$\hat{C}(t,t_0) = 1 - \frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}\mathrm{d}t'\hat{H}_{\mathrm{I}}(t'). \tag{13}$$

Now we want to evaluate the transition probability that the electron which is assumed to have been at time t_0 in a bound state $|\psi(t_0)\rangle = |E_n\rangle$ to be found in a scattering state $|E\rangle$. The corresponding transition-probability amplitude is given by

$$a_{fi} = \langle E, t_0|\hat{C}(t,t_0)|E_n,t_0\rangle = -\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}\mathrm{d}t'\langle E,t_0|\hat{V}_{\mathrm{I}}(t')|E_n,t_0\rangle. \tag{14}$$

For the matrix element, because of (10), we only need

$$\langle E,t_0|\hat{\vec{p}}(t')|E,t_0\rangle = \exp\left(\mathrm{i}\omega_{fi}t'\right)\langle E,t_0|\hat{\vec{p}}(t_0)|E_n,t_0\rangle, \tag{15}$$

where we have used the time evolution (5) for the momentum operator and the abbreviation $\omega_{fi} = [E-E_n]/\hbar$.
Plugging this into (14) we find

$$\begin{aligned} a_{fi} &= -\frac{\alpha}{2\hbar}\left[\frac{\exp[\mathrm{i}(\omega_{fi}-\omega)(t-t_0)]-1}{\omega_{fi}-\omega} + \frac{\exp[\mathrm{i}(\omega_{fi}+\omega)(t-t_0)]-1}{\omega_{fi}+\omega}\right] \\ &= -\frac{\mathrm{i}\alpha}{\hbar}\left[\exp[\mathrm{i}(\omega_{fi}-\omega)(t-t_0)/2]\frac{\sin[(\omega_{fi}+\omega)(t-t_0)/2]}{\omega_{fi}-\omega} + (\omega\to-\omega)\right], \end{aligned} \tag{16}$$

where

$$\alpha = \vec{A}_0\cdot\langle E,t_0|\hat{\vec{p}}(t_0)|E_n,t_0\rangle \tag{17}$$

Now we are interested in the probability that the electron is excited from a bound state with energy $E_i < 0$ to a scattering state with energy $E_f = E > 0$,

$$\begin{aligned} P_{fi} = |a_{fi}|^2 =& \frac{\alpha^2}{\hbar^2}\frac{\sin^2[(\omega_{fi}-\omega)(t-t_0)]}{(\omega_{fi}-\omega)^2} \\ &+ \frac{\alpha^2}{\hbar^2}\frac{\sin^2[(\omega_{fi}+\omega)(t-t_0)]}{(\omega_{fi}+\omega)^2} \\ &+ \frac{2\alpha^2}{\hbar^2}\cos(\omega t)\frac{\sin[(\omega_{fi}-\omega)(t-t_0)]}{\omega_{fi}-\omega}\frac{\sin[(\omega_{fi}+\omega)(t-t_0)]}{\omega_{fi}+\omega}. \end{aligned} \tag{18}$$

For $t-t_0 \to \infty$ we can use

$$\frac{\sin[(t-t_0)x]}{x} \simeq \pi\delta(x), \quad \frac{\sin^2[(t-t_0)x]}{x^2} \simeq \pi(t-t_0)\delta(x). \tag{19}$$

Thus, after a sufficiently long time the transition rate, becomes

$$w_{fi} = \dot{P}_{fi} \simeq \frac{\alpha^2}{\hbar^2}\delta(\omega_{fi}-\omega). \tag{20}$$

This shows that the transition is only possible, if

$$\omega_{fi} = \omega \Rightarrow E = E_i + \hbar\omega. \tag{21}$$

Now $E_i = -W < 0$ is the binding energy of the electron in the initial state, i.e., before the light has been switched on. This explains, from a modern point of view, Einstein's result (1) of 1905, however without invoking any assumption about "light particles" or photons.

We note that the same arguments, starting from Eq. (18) hold for $\omega_{fi} < 0$ and $\omega = -\omega_{fi}$. Then one has

$$E_f = E_i - \hbar\omega, \tag{22}$$

which describes the transfer of an energy $\hbar\omega$ from the electron to the radiation field due to the presence of this radiation field. This is called **stimulated emission**. Again, we do not need to invoke any assumption about a particle nature of light.

Where this feature truly comes into the argument can be inferred from a later work by Einstein (1917): One can derive Planck's black-body-radiation formula (1900) only under the assumption that despite the absorption and stimulated emission of energy quanta $\hbar\omega$ of the electromagnetic field, there is also a **spontaneous emission**, and from a modern point of view, this can indeed only be explained from the quantization of the electromagnetic field (in addition to the quantization of the electron). Then indeed, for the free quantized electromagnetic field, there is a particle-like interpretation, leading to a consistent picture of the electromagnetic field, interacting with charged particles, **Quantum Electrodynamics**.