Functional differential equations. 2: The classical hydrogen atom

C. K. Raju

AlBukhary International University
05200 Alor Setar, Malaysia
ckr@ckraju.net

(Submitted 22-07-2013)

Abstract

The previous part closed with a doubt: textbook electrodynamics mentions only ordinary and partial differential equations (ODEs and PDEs), never any functional differential equations (FDEs). So are FDEs or past data really needed? In fact, to solve Maxwell's PDEs we need Cauchy data on fields, which is equivalent to past data on particle motions. Electrodynamics actually involves a *coupled* system of ODEs and PDEs, equivalent to FDEs which are currently the best way to solve such a coupled system. The full electrodynamic force is commonly approximated by the Coulomb force, but this approximates FDEs by ODEs, and is error prone. Specifically, circular or central orbits are not valid solutions of the electrodynamic 2-body problem, in the *absence* of radiation damping. We can see this heuristically, for the classical hydrogen atom, where the full electrodynamic force involves a delay torque. Hence, the century-old argument that classical electrodynamics is inadequate is based on erroneous reasoning (irrespective of whether the conclusion is valid).

1 Recap

To recapitulate the first part, the motion of two charges is determined by the Heaviside-Lorentz force and Maxwell's equations. This leads to retarded FDEs. Such FDEs present a paradigm shift in physics, because past data is needed to solve them. However, the need for such a paradigm shift (or the need to prescribe past data) was contested in the Groningen debate. Zeh argued that the Heaviside-Lorentz force (and Newton's second law) lead to ODEs, while Maxwell's equations are PDEs. For solving either ODEs or PDEs, only initial data is needed. Hence, he maintained, there is no need for past data, and no paradigm shift. The puzzle was that the same underlying physics (Heaviside-Lorentz force, and Maxwell's equations) seems to lead to two opposite conclusions about the need for past data.

2 The resolution of the Groningen debate

Subsequently, I gave a simple resolution of this puzzle, which was eventually published in the *Foundations of Physics* 2004[1] which had Zeh on its editorial board. The simple resolution which relates FDEs to ODEs and PDEs is as follows.

2.1 Hyperbolic PDEs

For ODEs initial data determines a unique solution, as we know, and Peano's existence theorem provides a formal proof. For PDEs the situation *seems* the same, but is actually a little more complicated.

First, initial data are appropriate only for certain *types* of PDEs, called hyperbolic PDEs. The wave equation is a canonical example of a hyperbolic PDE: the effect (the wave) depends upon a specific cause (the source) in the past. The wave travels out and gets disconnected from the source. In contrast, the potential equation (Laplace equation, Poisson equation) is an example of an elliptic PDE. An elliptic PDE involves action at a distance: the Newtonian gravitational potential or the Coulomb field is decided everywhere and for all time by specifying the position of its source. The field remains attached to its source, so that boundary data are more appropriate than initial data.

2.2 Cauchy data

Initial data—formally called Cauchy data—involve values of the function (and its partial derivatives for a higher order equation). These must be specified on an appropriate hypersurface, called a Cauchy hypersurface. [A hypersurface is an $n-1$ dimensional subspace of an n dimensional space. Thus, a plane is a hypersurface in 3-dimensions, while 3-d space is a hypersurface in 4-dimensional spacetime.] Prescribing Cauchy data for a PDE is *analogous* to prescribing initial data for an ODE, since the "Cauchy hypersurface" may be just the "present instant", or the hypersurface in spacetime given by the equation $t = 0$. However, the analogy may be misleading.

2.3 What Cauchy data is needed for Maxwell's equations?

Many PDEs of physics (like the Hilbert-Einstein equations or Navier-Stokes equations) are not strictly hyperbolic, but let us just assume that Maxwell's equations are hyperbolic for (in the Lorenz gauge) they are just inhomogeneous wave equations.

Exactly *what* Cauchy data do we need for

Maxwell's equations? Since those equations involve the derivatives of the fields E and B, we must prescribe E and B on the Cauchy hypersurface, say $t = 0$. **That is, to solve Maxwell's equations, we must prescribe the electric and magnetic fields over *all* space at one instant of time.** But how do we *do* that?

2.4 Cauchy data for fields = past data for particles

In our case of two charged particles, consider a point P on the Cauchy hypersurface $t = 0$. The field at P must come from the motion of the two particles, for they are the only sources of fields by assumption. Since we have assumed retarded potentials, the field must come from their *past* motions. Which past motions? To determine the field at P, we must again construct the backward null cone with vertex at P and determine the two points (say P_A, and P_B) where it intersects the worldlines of the two particles, A, and B.

As the point P moves further away in space, the corresponding points P_A and P_B (at which the backward null cone from P intersects the two world lines) will move further back in time. As P runs over the entire hypersurface, the corresponding points P_A and P_B will cover the entire past world lines of the two particles. That is, (assuming retarded propagators) prescribing Cauchy data for fields on a hypersurface (i.e., at one instant of time) is equivalent to prescribing *past* data on the particle world-lines.

That is to solve Maxwell's equations, we need to prescribe the electric and magnetic fields over all space at only *one* instant of time. *But to do that we need to know the past motions of the two particles for all time.*

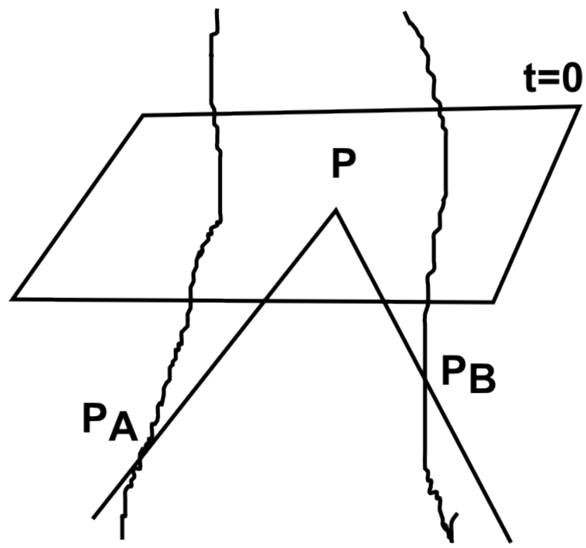

Figure 1: **I**nitial data for fields is past data for particles. To solve Maxwell's equations we must prescribe the fields $\vec{E}$ and $\vec{B}$ over all space at one instant of time $t = 0$. For the 2-body problem the field at a point P on $t = 0$ depends on the past motions of particles A and B at the retarded positions P_A and P_B.

So, past data *is* needed to solve for the motion of two charges, whichever way we look at it: whether in terms of fields or particles. Because the mathematical theory of PDEs is a little more complicated than that of ODEs, and many physicists are not very familiar with it, they just unthinkingly extended the Newtonian paradigm of ODEs to PDEs.

Further, physicists have unfortunately got carried away by the intuitive picture of a field.

The particle pictures pins us down and reveals the reality of retarded action at a distance (implicit in the concept of hyperbolicity). “Initial data” for fields (or knowledge of the fields now) involves past data on particle motions (or knowledge of the entire history of the two particles).

That is, Zeh made two mistakes: he put his faith in the Newtonian paradigm (“initial data are enough”), and he neglected the details of the actual calculations involved (“that is the mathematician’s job”), which would have made manifest the need for past data. One should clearly understand that this is not about Zeh as a person: he only personifies and represents a mistake made by practically the entire physics community, a mistake which has persisted for over a century.

2.5 Coupled ODEs and PDEs

One further point needs to be noted. The ODEs of motion (Heaviside-Lorentz force law + Newton’s second law) determine the motion of the charges *if the fields are known*. Maxwell’s equations (PDEs) determine the fields *if the motion of the charges are known*. So, to solve for the motion of two charges, we have to solve both *together*. So, what we have here is a *coupled* system of ODEs and PDEs, not an independent pair of ODEs and PDEs.

Unlike Peano’s existence theorem for ODEs, or the Cauchy-Kowaleski theorem for PDEs, there is still no corresponding formal mathematical theorem for a *coupled* system of ODEs and PDEs, as of now. (Though it is easy to prove such a theorem, I won’t do it, since, I now work with a different philosophy of mathematics,[2] and no longer think formal mathematical proofs are important.)

I mention this only to bring out that the coupled system of ODEs and PDEs is distinct from the two systems of ODEs and PDEs considered individually. It is necessary to hammer home the point about the coupled system, since most current texts on electromagnetic theory do not consider this coupling at all.

It is this coupled system of ODEs and PDEs which is equivalent to FDEs. Briefly:

$$\text{ODEs} + \text{PDEs} = \text{FDEs}.$$

We can solve either the coupled system of ODEs and PDEs or we can solve FDEs. However, FDEs are much easier to solve numerically, and well-tested computer programs to do so have been readily available for a quarter century.[3] (Also, as stated earlier, there are formally proven theorems that solutions of FDEs exist with past data.) Whichever way we look at it, it is, at present, preferable to solve FDEs instead of a coupled system of ODEs and PDEs, though the two are equivalent.

Either way, the paradigm shift is there, it is real, and it comes not from any new physical hypothesis but from a better understanding of the mathematics—an understanding which has been sadly missing for a whole century.

3 The classical hydrogen atom

So, what actual difference does that make to physics? Let us again consider our two charges, but this time in the explicit context of the classical hydrogen atom. Almost exactly a century ago, it was declared that classical electrodynamics does not work for this case, and this argument would be familiar to most physics students from high school.

3.1 The text-book story

Under the weight of the Newtonian paradigm, the atom was regarded as a sort of a miniature solar system, for the inverse-square law Coulomb force is just like the inverse-square law Newtonian gravitational force. Usually a further simplification is made: the proton is regarded, like the sun, as infinitely massive, hence unmoving at the centre. This reduces the 2-body problem to a *1-body case* of central orbits. Physics students are very familiar with central orbits which they study as part of their undergraduate course. According to the theory of central orbits, circular (or elliptic orbits) are stable with an inverse square law central force.[4]

Beyond this point, students learn, the analogy between classical electrodynamics and Newtonian gravitation fails: an accelerating charged particle gives out electromagnetic radiation. And, motion in a central orbit involves constant acceleration. Thus, it is argued, the classical central orbits are not stable in the electrodynamic case but must decay because of the associated radiation damping. The electron would constantly lose energy and eventually spiral into the nucleus. Hence the conclusion that classical electrodynamics cannot correctly describe the hydrogen atom.

3.2 The full electrodynamic force

It is odd how physicists have accepted this argument for over a century, when it is immediately obvious that something is wrong with it. The Coulomb potential propagates instantaneously like the Newtonian gravitational potential, but the actual Lienard-Wiechert (L-W) electromagnetic potentials propagate only at the speed of light. To correct this, we must use the full electrodynamic force, not just the Coulomb force.

The full electrodynamic force is obtained as follows. Recall that the electric and magnetic fields are obtained by differentiating the scalar potential V and the vector potential $\vec{A}$:

$$\begin{aligned}\vec{E} &= -\nabla V - \frac{\partial \vec{A}}{\partial t},\\ \vec{B} &= \nabla \times \vec{A}.\end{aligned} \tag{1}$$

We need to apply this to the L-W potentials

$$\begin{aligned}V(\vec{r},\, t) &= \frac{1}{4\pi\epsilon_0}\, \frac{qc}{(Rc - \vec{R}\cdot\vec{v})}\Bigg|_{\text{ret}},\\ \vec{A}(\vec{r},\, t) &= \frac{\vec{v}}{c^2}\, V(\vec{r},\, t)\Bigg|_{\text{ret}}.\end{aligned} \tag{2}$$

Recall that $\vec{R} = \vec{r} - \vec{r}_p(t_r)$ is the vector connecting the spacetime point $(\vec{r},\, t)$ to the position $\vec{r}_p(t_r)$ of charge q at retarded time t_r, which satisfies $c^2(t - t_r)^2 = R^2$. Recall also that the velocity too must be evaluated at retarded time, $\vec{v} = \dot{\vec{r}}_p(t_r)$.

Carrying out the differentiation is a long and tedious process, found in standard texts.[5] This gives us the following expressions for the electric and magnetic fields.

$$\vec{B}(\vec{r},t) = \frac{1}{c}\hat{\vec{R}} \times \vec{E}(\vec{r},t),$$
$$\vec{E}(\vec{r},t) = \frac{q}{4\pi\epsilon_0}\frac{R}{(\vec{R}\cdot\vec{u})^3}\Big[(c^2-v^2)\vec{u} + \vec{R}\times(\vec{u}\times\vec{a})\Big]. \tag{3}$$

Here, we have introduced $\vec{u} = c\hat{\vec{R}} - \vec{v}$, and $\vec{a}$ is the acceleration at retarded time: $\vec{a} = \ddot{\vec{r}}_p(t_r)$.

If the (retarded) acceleration is zero, we are left with only the first (velocity dependent) term for $\vec{E}$. If the retarded velocity too is zero, this term reduces to a term similar to the Coulomb force:

$$\vec{E}(\vec{r},t) = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2}\hat{\vec{R}}. \tag{4}$$

(It is exactly the Coulomb force if the charge q is static for all time.)

From the above expressions (3) for $\vec{E}$ and $\vec{B}$, we can calculate the force on a charge q_1 moving with velocity $\vec{v}_1$ using the Heaviside-Lorentz force law: $\vec{F} = q_1(\vec{E} + \vec{v}_1 \times \vec{B})$. This gives the expression:

$$\vec{F} = \frac{qq_1}{4\pi\epsilon_0}\frac{R}{(\vec{R}\cdot\vec{u})^3}\Big\{\Big[(c^2-v^2)\vec{u} + \vec{R}\times(\vec{u}\times\vec{a})\Big] + \frac{\vec{v}_1}{c}\times\Big[\hat{\vec{R}}\times[(c^2-v^2)\vec{u} + \vec{R}\times(\vec{u}\times\vec{a})]\Big]\Big\}. \tag{5}$$

3.3 Full electrodynamic force leads to FDEs

Using the expression (5) for the full electrodynamic force leads to FDEs because of the retarded quantities involved. If we use only the Coulomb force, that would lead to ODEs. So, approximating the full electrodynamic force by the Coulomb force amounts to approximating FDEs by ODEs. That is an error-prone process, as already explained in the first part of this article.

On the other hand, if we did not solve those FDEs, how do we know that circular orbits are stable? Can we just say, like Zeh, that we know this "on physical grounds" and that it is the mathematician's job to prove us right?

It is not very difficult to numerically solve the retarded FDEs for the non-relativistic case without radiation damping. Nevertheless, for the classical hydrogen atom, the first solution was published by me only in 2004 while resolving the Groningen debate. Here are some of the problems that arise.

3.4 How should past data be prescribed?

The first problem is the question of prescribing past data. How exactly should that be done? Because the issue of past data has not been considered in physics until now, there are no clear guidelines available. However, the above connection between FDEs and PDEs suggests a way: prescribing past data on particle motions is equivalent to prescribing fields on a spacelike Cauchy hypersurface. Allowing arbitrary Cauchy data corresponds to allowing arbitrary past motions of the two particles. That is, we don't try to explain how those particle motions came about.

Fortunately, unlike the PDE case, for which one must (on existing theory) prescribe the *entire* past history, for the FDE case it is usually enough to prescribe only a short portion of the past history of the two particles, since we are typically interested in the solution for a short period of space and time. The same can be done with ODEs+PDEs, but there is no formal proof for that, as of now.

However, various questions arise. How did that past history come about? Is it physically realizable? These are questions which the mathematical theory of FDEs and PDEs cannot answer, and which physicists have not raised or answered till now. Prescribing past data arbitrarily may lead to a discontinuity in the solution or (more usually) its derivatives. This is not a major technical problem, since most existing numerical codes are equipped to handle such discontinuities. The discontinuity may be physically understood as arising from suddenly "switching on" the interaction. Nevertheless, this is an unsatisfactory state of affairs. Can the past data be prescribed in such a way as to avoid this discontinuity? The answer to this question is not known at present.

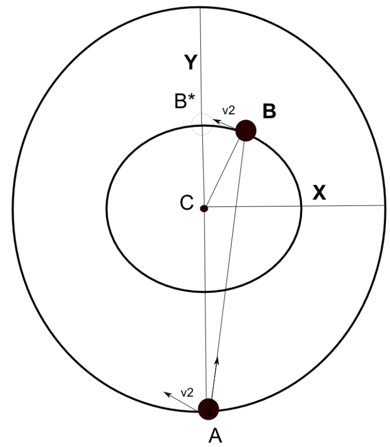

Figure 2: **E**ffect of retardation in the 2-body problem. Classically, both particles A and B rotate around a common centre of mass, C. The force acting on A now involves the past or retarded position of B and not its instantaneous position ($B*$). This force does not pass through the instantaneous centre of mass and hence results in a torque, called the "delay torque".

4 The delay torque

Anyway, without going into all the details of that solution, we can heuristically see the problems involved.

Examine Fig. 2. First, note the difference from central orbits (or the 1-body problem). For the *two* body problem, even in classical Newtonian gravity, both bodies revolve around a common centre of mass. To put matters in another way the motion of planets is *not* heliocentric, but is better understood as barycentric.

The next step is to take into account the retardation, since the full electrodynamic force does not travel instantaneously but only at the speed of light.

Consider now just the first term in the first square bracket in (5) which corresponds to the "Coulomb" component of the electric field due to q. This component acts in the direction of the vector $\vec{u}$, given by $\vec{u} = c\hat{\vec{R}} - \vec{v}$. If $\frac{v}{c}$ is small this is approximately the same direction as $\hat{\vec{R}}$. That is the force on A (charge q_1) due to B (charge q) will point approximately towards the "last seen" position of B, not its instantaneous position. Consequently, the force will not pass through the instantaneous centre of mass (barycentre) of the two particles and it will exert a torque on particle A. That means that circular orbits will *not* be stable even in the *absence* of radiation damping!

Of course, we should take into account that the force acts in the direction of the vector $\vec{u}$ and not just the vector $\vec{R}$. But will that resolve the problem in an obvious way? Will that "correct" the force, so that it points towards the instantaneous centre of mass? It cannot because there is no physical way to ascertain the exact instantaneous position of B, hence no way to ascertain the instantaneous centre of mass: at best the full force may point towards the estimated or extrapolated instantaneous centre of mass. That is of little consequence, for that process of extrapolation would fail for more complex motions. We will consider this argument again, in the context of gravity.

4.1 Choice of gauge

Have we neglected anything else? What if we were to use the Coulomb gauge instead of the Lorenz gauge? Choices of gauge typically confuse physics students, for the scalar potential in the Coulomb gauge is just the same old Coulomb potential which propagates instantaneously. However, that is deceptive: changing the gauge does *not* change the electric or magnetic fields. It does not change the full force in (5). So the torque, too, does not depend upon the gauge.

4.2 Delay torque initially accelerates the particle

A glance at Fig. 2 shows that the torque will *accelerate* the electron. We can confirm this by evaluating the delay torque numerically for an electron initially in a classically stable orbit: the torque (initially) accelerates the electron. That is, *in the absence of radiation damping*, an electron initially in a classically stable orbit (on the Newtonian paradigm) tends to fall *out* of the atom!

The exact motion may be complicated, but the point here is only this. The whole argument given by Bohr for the instability of the classical hydrogen atom, an argument repeated for a century in physics texts, is defective because it assumed (on the Newtonian paradigm) that classical central orbits are stable in the absence of radiation damping. (To reiterate, it does not matter whether the conclusion is still valid, for even if one gets the right answers for the wrong reasons, that is not science.)

Further, we need to take into account motions more complex than simple circular orbits, for it may be that a more complex past motion (such as an oscillation superposed on a circular orbit) leads to stable solutions. Whether or not that is so, is an open question at present. We cannot decide that without solving the FDEs. Unfortunately, this has not been done for the past century.

5 Summary

Thus, for the 2-body problem of classical electrodynamics, the stability of orbits, even in the absence of radiation damping, is a complex problem which has not been properly studied so far. The whole physics community just went along with a wrong solution based on the Newtonian paradigm.

A natural question arises. If, in the absence of radiation damping, the delay torque makes circular orbits unstable, so that the electron tends to fall *out* of the atom, then will stability be somehow restored by reintroducing radiation damping? In short, are there motions for which the delay torque and the radiation damping cancel (either exactly or on an average)? We will examine this question in the next part which connects radiation damping to FDEs.

References

[1] C. K. Raju. The electrodynamic 2-body problem and the origin of quantum mechanics. *Foundations of Physics*, 34(6):937–962, 2004. http://arxiv.org/pdf/quant-ph/0511235v1.

[2] C. K. Raju. *Cultural Foundations of Mathematics*. Pearson Logman, 2007.

[3] E. Hairer, S. P. Norsett, and G. Wanner. *Solving Ordinary Differential Equations*. Springer, revised ed. 1991 edition, 1987.

[4] J. L. Synge and B. A. Griffith. *Principles of Mechanics*. McGraw Hill, 3rd edition, 1959. Equation 18.320.

[5] David J. Griffith. *Introduction to Electrodynamics*. Prentice Hall, India, 1999.