Functional differential equations. 1: A new paradigm in physics

C. K. Raju

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

(Submitted 22-07-2013)

Abstract

This article explains the basics of functional differential equations (FDEs). FDEs differ fundamentally from the ordinary and partial differential equations. As such, FDEs present a new paradigm in physics. Since FDEs arise naturally through Maxwellian electrodynamics, this involves no new physical hypothesis, but only a revised mathematical understanding of existing physics.

1 Introduction

1.1 ODEs and PDEs

Physicists have long been familiar with ordinary differential equations (ODEs). Newton's second "law" of motion reduces physical problems of classical mechanics to the mathematical problem of solving ODEs.

For example, as every physics student knows, a force proportional and opposite to displacement leads to the ODE of simple harmonic motion:

$$y'' = -k^2 y. \tag{1}$$

Here, $y \equiv y(t)$ denotes the position of the particle at time t, primes denote derivatives and k is a constant.

The second order ODE (1) has a unique solution if we prescribe the initial value of y and its first derivative y'.

$$\begin{aligned} y(0) &= y_0, \\ y'(0) &= y_1. \end{aligned} \tag{2}$$

In general, ODEs have a unique solution from (appropriate) initial data. Physically, this feature of ODEs means that the future states of a classical system are determined if we prescribe its initial state: namely the initial position and velocity.

In general, if we use Hamilton's form of the ODEs for a classical system of n particles, to fix the initial state we must prescribe the initial values of the positions and canonically conjugate momenta for each particle. This is the Newtonian paradigm: *the future states of a physical system are determined by its initial state alone.*

Less easy, but common in physics, are partial differential equations (PDEs). In classical physics, these arise when, instead of a system of n discrete particles, one works with continua—that is, fluids or fields. Some well-known PDEs are the Navier-Stokes equations for fluid flow, Maxwell's equations for electromagnetic fields, and the Hilbert-Einstein equations of general relativity. In quantum physics, even a single particle is described by PDEs such as the Schrödinger equation.

1.2 FDEs

Little known, however, are the functional differential equations (FDEs) which differ from both ODEs and PDEs. The mathematical theory of FDEs has been around for some time now, but their fundamental importance for physics was understood only relatively recently. Hence, these are not yet common in physics texts.

Let us start with a very simple example of an FDE. Consider

$$y'(t) = y(t-1). \tag{3}$$

This is also called a retarded FDE or a delay differential equation, since it involves a time-lag (retardation) or delay: instead of $y(t)$ we have $y(t-1)$ on the right hand side. What difference does that make? To see this, suppose we try to solve this FDE like we solve ODEs. That is, we give the initial value $y(0)$, and ask for the value of $y(1)$. Can we obtain it?

1.3 Retarded FDEs need past data

Symbolically we can write:

$$y(1) = y(0) + \int_0^1 y'(t)dt \tag{4}$$

$$= y(0) + \int_0^1 y(t-1)dt. \tag{5}$$

But to actually carry out the integration we need to know the values of the integrand $y(t-1)$ in (5). To know $y(t-1)$ for all $t \in [0, 1]$ is to know $y(t)$ for all $t \in [-1,\ 0]$. That is, unlike the ODE $y' = y$ for which the *initial* data at $t = 0$ suffices, for the retarded FDE (3) we need *past* data or the values of the function $y(t) \equiv \phi(t)$ over an entire past interval $[-1,\ 0]$.

If, instead of an initial function $\phi(t)$, only the initial value $y(0)$ is prescribed, then we can assume the past values in infinitely many different ways, so we can find an infinity of different solutions. Three such solutions are

depicted in the graph: all these solutions have the same prescribed initial value of $y(0)$ but different past values (see Fig. 1).

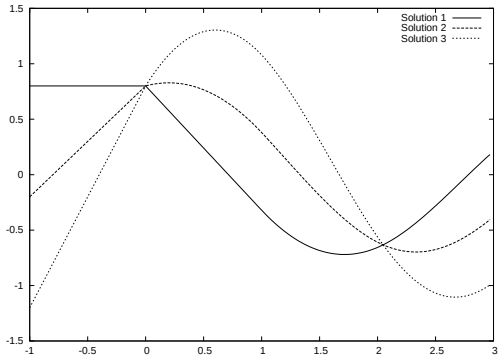

Figure 1: **FDEs need past data:** Three different solutions of the FDE (3) with the same initial data but different past data.

The equation (3) relates the rate of change of y *now* to its *past* values, i.e., the FDE models a history-dependent situation. So, intuitively, it makes sense that to solve it we need to know the relevant past data or past history. It is obvious from (5) that the FDE (3) has a unique solution if the initial function $\phi(t)$ is continuous. (However, this solution may fail to be differentiable at $t = 0$, as in Solution 1 of Fig. 1.) This can be generalised into a formal mathematical theorem[1] that FDEs admit a unique local solution from past data under appropriate conditions.

1.4 Time asymmetry of retarded FDEs

FDEs differ from ODEs in another fundamental way. ODEs are instantaneous: Newton's second "law" relates force *now* to acceleration (or the second derivative of position) *now*. Since ODEs relate to what happens at *one* instant of time, they do not discriminate between past and future. Instead of regarding (2) as "initial" data, we can just as well regard it as "final" data and solve (1) backward in time.

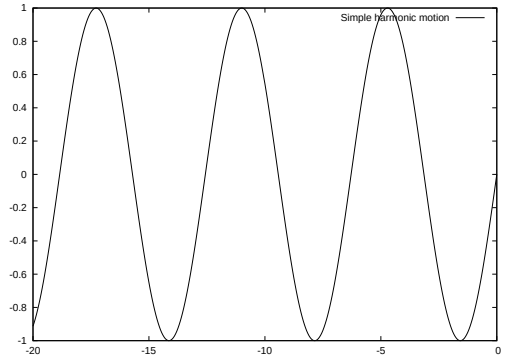

Figure 2: **An ODE solved backward.** ODEs are time symmetric; they can be solved either forward or backward in time. The above graph shows a solution of the simple harmonic oscillator calculated backward from prescribed values of $y(0)$ and $y'(0)$ regarded as final data. This can be done exactly as we calculate forward solutions with the same values regarded as initial data. The solution was obtained with my software CALCODE which solves ODEs in either direction in time.

The retarded FDE (3), however, is time *asymmetric*: it relates $y'(t)$ (or the rate of change of y now) to its past values, namely $y(t-1)$. It can only be solved forward in time. We can determine $y(t)$ for future values ($t \geq 0$) given past data ($y(t)$ for $t \leq 0$), but not the other way round: past values ($y(t)$ for $t \leq 0$) cannot be determined from future data ($y(t)$ for $t \geq 0$). In the following example, different past histories all converge to the same future, hence from a knowledge of the future, one cannot ascertain the past.

1.5 An example

Consider the FDE

$$y'(t) = b(t)y(t-1), \tag{6}$$

where b is a continuous function which vanishes outside [0, 1], and satisfies

$$\int_{-\infty}^{\infty} b(t)\ dt = \int_0^1 b(t)\ dt = -1. \tag{7}$$

For example,

$$b(t) = \begin{cases} 0 & t \leq 0, \\ -1 + \cos 2\pi t & 0 \leq t \leq 1, \\ 0 & t \geq 1. \end{cases} \tag{8}$$

For $t \leq 0$, the FDE (6) reduces to the ODE $y'(t) = 0$, so that, for $t \leq 0$, $y(t) = k$ for some constant k $(= y(0))$.

Now, for $t \in [0,\ 1]$,

$$\begin{aligned} y(t) &= y(0) + \int_0^t y'(s)ds \\ &= y(0) + \int_0^t b(s)y(s-1)ds \\ &= y(0) + y(0)\int_0^t b(s)ds, \end{aligned} \tag{9}$$

since $y(s-1) \equiv k = y(0)$ for $s \in [0, 1]$. Hence, using (7), $y(1) = 0$, no matter what k was. However, since $b(t) = 0$ for $t \geq 1$, the FDE (6) again reduces to the ODE $y'(t) = 0$, for $t \geq 1$, so that $y(1) = 0$ implies $y(t) = 0$ for all $t \geq 1$.

Hence, the past of a system modeled by (6) cannot be retrodicted from a knowledge of the entire future; for if the future data (i.e., values of the function for *all* future times $t \geq 1$) are prescribed using a function ϕ that is different from 0 on $[1,\ \infty]$, then (6) admits no backward solutions for $t \leq 1$. If, on the other hand, $\phi \equiv 0$ on $[1,\ \infty]$, then there are an infinity of distinct backward solutions. Fig. 3 shows three such solutions. In either case, knowledge of the entire future furnishes no information about the past.

With retarded FDEs we can infer future from past, but not, in general, past from future. They model a situation where we have *more* information about the past than the future. To put matters differently, we can say there is loss of information towards the future. This is the same as saying that retarded FDEs model a situation where there is an increase of entropy towards the future, for information is the negative of entropy, as I have explained in detail, in an earlier paper.[2] With ODEs,

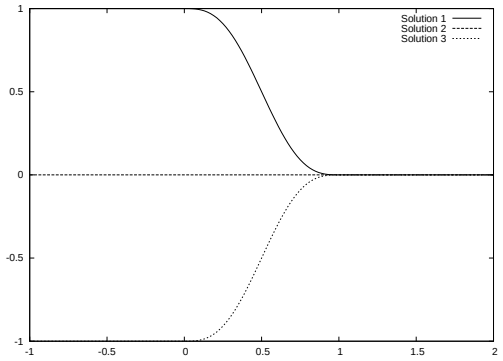

Figure 3: **FDEs are time asymmetric:** Three different solutions of a retarded FDE with the different pasts have the same future, so that retarded FDEs cannot be solved backward. That is, past cannot be inferred from knowledge of future.

entropy must stay constant.

2 The paradigm shift

These features—the need for past data, and time asymmetry—mean that the physics of a system modeled by FDEs differs fundamentally from the physics of a system modeled by ODEs (classical mechanics). This led to my claim that FDEs involve a shift away from the Newtonian paradigm.

A "paradigm shift" refers to a fundamental new idea. People often resist such new ideas, and try to hang on to the old ways of thinking. More specifically, physicists accustomed to ODEs and PDEs resist the changes necessitated by FDEs, and tend to fall back into the old ways associated with ODEs and PDEs.

2.1 Converting FDEs to ODEs is a mistake

This has led to a common mistake: physicists (including Einstein) tried to approximate FDEs by ODEs. Such an approximation seems plausible because the retarded FDEs which arise in physics typically involve "small" delays. That is, we have FDEs of the type

$$y'(t) = y(t-\tau) \tag{10}$$

where $\tau > 0$ is "small". Suppose we "Taylor" expand $y(t-\tau)$ in powers of the delay τ. This allows us to approximate $y(t-\tau)$ by the values of y and its derivatives at t. Inserting this approximation into the FDE (10) converts it to a higher-order ODE. Is this a valid approximation? Will the solutions of the resulting ODE approximate the solutions of the original FDE?

No. Not, in general. This should be obvious by now, because of the qualitative differences between FDEs and ODEs brought out above. However, let us see yet another explicit counter-example.

2.2 Incorrectness of approximating FDEs by ODEs

Consider the equation

$$y'(t) = y(t-\tau) - y(t), \tag{11}$$

where $\tau > 0$ is a small constant. If we expand the right hand side in "Taylor's" series in

powers of the delay τ, and truncate after two terms, we obtain

$$y'(t) = \{y(t) - \tau y'(t) + \frac{\tau^2}{2!} y''(t)\} - y(t). \quad (12)$$

This simplifies to

$$y''(t) - \frac{2(1+\tau)}{\tau^2} y'(t) = 0. \quad (13)$$

This is a linear ODE with constant coefficients, which every physics student should know how to solve. The solution is

$$y_{\text{approx}} = c_1 + c_2 e^{st}, \quad (14)$$

where $s = \frac{2(1+\tau)}{\tau^2} > 0$ since, by assumption, $\tau > 0$. Here, c_1, c_2 are constants determined by the initial data. Hence, if $c_2 \neq 0$, the solution grows exponentially in the future.

On the other hand one may obtain solutions to (11) as follows. Substituting, $y = e^{zt}$ (z complex) in (11), and cancelling e^{zt} from both sides, we find that z must satisfy

$$z = e^{-z\tau} - 1, \quad \text{or} \quad z + 1 = e^{-z\tau}. \quad (15)$$

This is the so-called characteristic quasi-polynomial equation. For large $|z|$ we can approximately replace $z + 1$ by z, so this reduces approximately to

$$z = e^{-z\tau}, \quad \text{or} \quad ze^{z\tau} = 1. \quad (16)$$

The above quasi-polynomial equation admits an infinity of roots. Without going into the detailed derivation, these roots (hence also the large modulus roots of (15)) are approximately given by[3]

$$z_k \approx -\frac{1}{\tau} \ln(2k - \frac{1}{2})\frac{\pi}{\tau} + i(2k - \frac{1}{2})\frac{\pi}{\tau}, \quad (17)$$

where k is an integer. Since $\tau > 0$ has been assumed small, the above roots z_k of the quasi-polynomial all lie in a left half-plane $\text{Re}(z_k) < 0$. Each root z_k corresponds to a solution of (11) of the form $y = e^{z_k t}$, or to an oscillation with amplitude $e^{\text{Re}(z_k)}$. If $\text{Re}(z_k) < 0$, this solution must be exponentially damped. Hence, there are an infinity of solutions of (11) which are exponentially damped oscillations, contrary to the exponentially increasing solutions (14) of the approximating ODE (13).

Thus, expanding FDEs to obtain ODEs, by means of a "Taylor's" series, may lead, as above, to spurious solutions with the completely opposite behaviour. Note, also, in passing, that the retarded FDE (11) admits an infinity of distinct complex solutions which is impossible for the approximating ODE (14), or any ODE.

Electrodynamics provides a common situation where FDEs arise in physics, and where such a mistaken conversion of FDEs to ODEs is common.

3 The motion of two charges

3.1 Preliminaries

Consider *two* interacting charges which are otherwise isolated. According to present-day physics the cosmos consists mostly of charged particles, such as the electron, proton, etc. So the interaction of two charges should be simple problem which every physicist should know how to solve!

However, physics students are taught in high school that this basic problem requires quantum mechanics; they are taught that it *cannot* be solved using classical electrodynamics. Exactly why not?

What I will now show is this: the belief that classical electrodynamics is inadequate may be right or wrong, but it is based on bad reasoning. (And even if we arrive at the right answer for wrong reasons that is not science.) Classical electrodynamics actually leads to FDEs which physicists mistakenly converted to ODEs to draw the conclusion that classical electrodynamics does not work for the atom.

The FDEs for the classical electrodynamic 2-body problem are explicitly written down in [4](relativistic case) and [5] (non-relativistic case). Here let us intuitively understand why at all FDEs are needed. On classical electrodynamics, the two charges interact through electromagnetic fields. Each charge moves in the electromagnetic field of the other.

3.2 Heaviside-Lorentz force

A charge moving in an electromagnetic field experiences a Heaviside-Lorentz force given by the expression:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{18}$$

From this force, we can determine the motion of the charge by applying Newton's law of motion. These two laws together determine the motion of each charge if the electromagnetic field of the other is known.

3.3 Maxwell's equations

The electromagnetic field of each charge is determined using Maxwell's equations:

$$\begin{aligned} \nabla \cdot \vec{E} &= \frac{\rho}{\epsilon_0}, \\ \nabla \cdot \vec{B} &= 0, \\ \nabla \times \vec{E} &= -\frac{\partial B}{\partial t}, \\ \nabla \times \vec{B} &= \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial E}{\partial t}. \end{aligned} \tag{19}$$

We must solve these PDEs for each charge.

How to do so is explained in any standard physics text[6]. The fields $\vec{E}$ and $\vec{B}$ are calculated as the derivatives of a scalar potential V and a vector potential $\vec{A}$.

$$\begin{aligned} \vec{E} &= -\nabla V - \frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \nabla \times \vec{A} \end{aligned} \tag{20}$$

With this, the middle two of Maxwell's equations are automatically satisfied. The choice of potential is not unique, and we can add an extra condition (called a gauge condition) to simplify the remaining two equations. In the Lorenz gauge, the first and the last of Maxwell's equations turn into inhomogeneous wave equations for the scalar and vector potential respectively.

3.4 The Lienard-Wiechert potentials

We can solve the inhomogeneous wave equation for any charge distribution, provided we know the solutions for a point charge, or a δ

function charge density. (Such solutions are also known as Green functions, and mathematicians call them fundamental solutions.) We can calculate these Green functions by taking a Fourier transform (the Fourier transform of the delta function or distribution is 1), solving the resulting algebraic equation, and then applying the inverse Fourier transform.

The potentials V and $\vec{A}$ in this case are known as the Lienard-Wiechert (L-W) potentials. Unlike the Newtonian gravitational potential, which propagates instantaneously, these L-W potentials propagate only with the speed of light c. Thus, the field acting on charged particle A *now* ($t = 0$) depends upon the motion of B at a *different* time ($t \neq 0$).

Which different time? That depends on which kind of the L-W potential we use, for the L-W potentials are of two kinds—retarded and advanced. In the retarded case, the field acting on a charge A now depends upon the motion of B in the *past* ($t \leq 0$). In the advanced case it depends upon the motion of B in the *future* ($t \geq 0$). Usually, only the retarded potentials are considered on grounds of "causality". Let us go along with that for the moment.

The retarded L-W potentials are given by the expressions:

$$V(\vec{r},\, t) = \frac{1}{4\pi\epsilon_0}\, \frac{qc}{(Rc - \vec{R}\cdot\vec{v})}\,\Bigg|_{\rm ret},$$
$$\vec{A}(\vec{r},\, t) = \frac{\vec{v}}{c^2}\, V(\vec{r},\, t)\Bigg|_{\rm ret}. \qquad (21)$$

This give us the potentials (hence the fields) at any position $\vec{r}$ at any time t, due to a charge q. Here, $\vec{R} = \vec{r} - \vec{r}_p(t_r)$, where $\vec{r}_p(t_r)$ denotes the position of the charge q at the *retarded time* t_r. The subscript "ret" emphasizes that $\vec{v}$, the velocity of the charge q, is also to be evaluated at retarded time t_r, so that $\vec{v} = \dot{\vec{r}}_p(t_r)$. The retarded time t_r satisfies

$$c^2(t - t_r)^2 = R^2, \qquad (22)$$

where c is the speed of light.

3.5 Retarded time

Exactly what is this retarded time? The equation (22) is actually the equation of the null cone. Suppose we look at particle B from the position of A at time $t = 0$. The light wave from B which reaches us just now ($t = 0$) started off in the past, from the "last seen" position of B. Light waves travel along the null cone, so the last-seen position of B is its retarded position. The corresponding time is the retarded time. The position, velocity etc. of B at *retarded time* is what we must use to evaluate the L-W potentials and calculate the fields acting on A now (and vice versa).

Geometrically, the retarded time is obtained as follows. We construct the backward null cone with vertex at the worldline of A at $t = 0$, and find the point, $t = t_B$ say, at which it intersects the worldline of B. That is equivalent to saying that the forward null cone with vertex at $t = t_B$ on the worldline of B intersects the worldline of A at $t = 0$. Although the point t_B is only in the *relative* past of $t = 0$, recall that the solution of the FDEs needs past data, or knowledge of the past worldline. Hence knowledge of the absolute past is needed to solve the resulting equations.

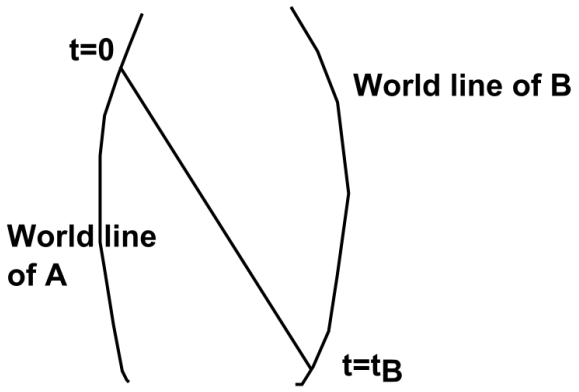

Figure 4: **R**etarded time. The retarded time is obtained by drawing the backward null cone from the present position of A and finding the point at which it intersects the world line of B.

This means that for a system of two charged particles, the future motion is not decided by their states at the present moment alone—we need to know their past motions. Since, charged particles (electrons, protons, ...) are everywhere that means we have a paradigm shift in physics: according to *existing* physics. We have *not* introduced any new hypothesis, but have just done the math correctly. The conclusion is that, contrary to common belief, the future states (of a classical system) are *not* decided by its present state alone; we require knowledge of its past.

4 The Groningen debate

This conclusion about a paradigm shift in physics was first published in this very journal in 1992, as part of a series of 10 articles "On Time" which were published between 1990 and 1994. The intention was to carry on the conversation which used to take place under the tree near the Pune university canteen.

However, many scientists today assess the validity of a scientific claim, not by applying their mind to the claim, but just by judging the "prestige" of the publisher who published it. Though this regrettable practice of judging scientific truth by social prestige and authority is contrary to the spirit of science, it is widely prevalent. Accordingly, I republished this conclusion about a paradigm shift (and the whole related series of papers) as a book, *Time: Towards a Consistent Theory*, with a "prestigious" foreign publisher (Kluwer, now Springer), in 1994,[7] since every scientist knows that only Western publications are prestigious!

Nevertheless, this claim about a paradigm shift ran into heavy fire, when I mentioned it at a meeting ("Retrocausality Day") in Groningen, in 1999. One of the participants, H. D. Zeh, a professor from Heidelberg, objected strenuously, asserting that no paradigm shift was needed. The whole meeting was stalled by the debate on this one point: is there or is there not a paradigm shift?

Zeh's argument was this: the motion of two charges is completely decided by two equations.

1. The Heaviside-Lorentz force (18), which force gives us a system of ODEs for the motion of each charge, by Newton's second law of motion.
2. Maxwell's equations (19) which deter-

mine the fields E and B generated by each particle (to be plugged into (18)). These are PDEs.

Hence, Zeh argued, to solve for the motion of the two charges we only need to solve ODEs and PDEs, never any FDEs. Further, Zeh argued, for either ODEs or PDEs initial data is adequate. Therefore, Zeh concluded, there is no need for past data, and consequently no paradigm shift.

My immediate response was that I had actually solved the electrodynamic 2-body problem using FDEs. This solution required past data, and if Zeh had a different way of obtaining the solution, he ought to show it. Zeh replied that he was sure of his stand on "physical grounds", and that actually solving the equations was the mathematician's job!

The puzzle which emerged was this: my formulation of the electrodynamic 2-body problem in terms of FDEs ultimately used nothing more than the very same Heaviside-Lorentz force (ODEs) and Maxwell's equations (PDEs). So how could two such very different conclusions emerge from the *same* underlying physics? Is there, or is there not a paradigm shift involved? Is there, or is there not a need for past data in physics?

Though the participants in the meeting came from many prestigious universities around the world, no one had a ready answer. The debate remained unresolved during the Groningen meeting. We will see the resolution in the next part.

References

[1] R. D. Driver. *Introduction to Differential and Delay Equations*. Springer, Berlin, 1977. The existence theorem there does not, however, apply to FDEs with unbounded delays, which *may* arise in electrodynamics.

[2] C. K. Raju. Thermodynamic time. In *Time: Towards a Consistent Theory*, chapter 4, pages 79–101. Kluwer Academic, Dordrecht, 1994.

[3] L. E. El'sgol'ts and S. B. Norkin. *Introduction to the theory of differential equations with deviating arguments*. Holden-Day, San Francisco, 1966. Typo corrected in equations (28) and (29) on p. 74.

[4] C. K. Raju. Electromagnetic time. *Physics Education (India)*, 9(2):119–128, 1992. http://arxiv.org/pdf/0808.0767v1.

[5] 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.

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

[7] C. K. Raju. *Time: Towards a Consistent Theory*, volume 65 of *Fundamental Theories of Physics*. Kluwer Academic, Dordrecht, 1994.