Constraint-Related Reinterpretation of Fundamental Physical Equations Can Serve as a Built-In Regularization

Vladik Kreinovich[1], Juan Ferret[2], and Martine Ceberio[1]

Departments of [1]Computer Science and [2]Philosophy
University of Texas at El Paso, El Paso, TX 79968, USA,
{vladik,jferret,mceberio}@utep.edu

Abstract. Many traditional physical problems are known to be *ill-defined:* a tiny change in the initial condition can lead to drastic changes in the resulting solutions. To solve this problem, practitioners *regularize* these problem, i.e., impose explicit constraints on possible solutions (e.g., constraints on the squares of gradients). Applying the Lagrange multiplier techniques to the corresponding constrained optimization problems is equivalent to adding terms proportional to squares of gradients to the corresponding optimized functionals. It turns out that many optimized functionals of fundamental physics already have such squares-of-gradients terms. We therefore propose to re-interpret these equations – by claiming that they come not, as it is usually assumed, from unconstrained optimization, but rather from a constrained optimization, with squares-of-gradients constrains. With this re-interpretation, the physical equations remain the same – but now we have a built-in regularization; we do not need to worry about ill-defined solutions anymore.

Keywords: constraints; fundamental physics; regularization; ill-defined problems

1 Formulation of the Problem

Optimization reformulation of physical equations. Traditionally, laws of physics have been described in terms of differential equations. However, in the 19th century, it turned out that these equations can be reformulated as optimization problems: the actual field is the one that minimizes the corresponding functional (called *action* S). This optimization approach is very useful in many applications (see, e.g., [1]) since there are many efficient algorithms for solving optimization problems.

Decision making and control: ideal situation. In decision making and control applications, in principle, we can similarly predict the result of different decisions, different control strategies. Thus, we can select the decision (or the control strategy) that leads to the most favorable result.

Real-life prediction: limitations. In practice, however, the situation is not so simple. The main problem is that all measurements are only approximate. Even for the most accurate measurements, the measured values of the initial conditions are slightly different from the actual values.

Most prediction problems are *ill-defined* in the sense that small deviations in the initial conditions can cause arbitrary large deviations in the predicted values.

Limitations: example. One of the main reasons why the prediction problem is ill-defined is that no matter how small a sensor is, it always has a finite size. As a result, the sensor does not produce the value $f(x)$ of the measured field f exactly at a given spatial location x; the sensor always captures the "average" value of a signal over a certain neighborhood of the point x – the neighborhood that is occupied by this sensor. Hence, field components with high spatial frequency $f(x) = f_0 \cdot \sin(\omega \cdot x)$ (with large ω) are averaged out and thus, not affected by the measurement result. Therefore, in addition to the measured field $f(x)$, the same measurement result could be produced by a different field $f(x) + f_0 \cdot \sin(\omega \cdot x)$. For many differential equations, future predictions based on this new field can be drastically different from the predictions corresponding to the original field $f(x)$.

How this problem is solved now. To solve the problem, practitioners use *regularization*, i.e., in effect, restrict themselves to the class of solutions that satisfies a certain constraint; see, e.g., [5]. For example, for fields $f(x)$, typical constraints include bounds on the values $\int f^2\,dx$ and $\int f_{,i} \cdot f^{,i}\,dx$, where $\int F\,dx$ means integration over space-time (or, for static problems, over space), $f_{,i} \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}$, and an expression $f_{,i} \cdot f^{,i}$ means summation over all coordinates i.

By imposing bounds on the derivatives, we thus restrict the possibility of high-frequency components of the type $f_0 \cdot \sin(\omega \cdot x)$ and thus, make the problem well-defined.

Limitations. The main limitation of different regularization techniques is that the bounds on the derivatives are introduced *ad hoc*, they do not follow from the physics, and different bounds lead to different solutions.

There is a whole art of selecting an appropriate regularization techniques, and, once a technique is selected, of selecting an appropriate parameter. It is desirable to come up with a more algorithmic way to making the equations well-defined.

2 Main Idea

A mathematical reminder: how to optimize functionals (see, e.g., [2]) As we have mentioned, fundamental physical equations are described in terms of minimizing a functional called *action.* This functional usually has an integral form $S = \int L(f, f_{,i})\,dx$; the corresponding function L is called a *Lagrangian.*

The main idea behind minimizing such functional is similar to the idea of minimizing functions. For functions $f(x_1, \ldots, x_n)$, optima occur when all the partial derivatives are 0s. Similarly, for a functional, an optimum occurs if the *functional derivative* is 0:

$$\frac{\delta L}{\delta f} \stackrel{\text{def}}{=} \frac{\partial L}{\partial f} \cdot \Delta f - \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial f_i}\right) = 0.$$

This is how usual differential equations are derived from the optimization reformulation of the corresponding physical theories.

A mathematical reminder: how constraints are currently taken into account? When we optimize a functional, e.g., $\int f^2\,dx$, under a constraint such as

$$\int f_{,i} \cdot f^{,i}\,dx \le \Delta,$$

then, from the mathematical viewpoint, there are two options:

– It is possible that the optimum of the functional is attained strictly inside the area defined by the constraints. In the above example, it means that the optimum is attained when $\int f_{,i} \cdot f^{,i}\,dx < \Delta$. In this case, all the (functional) derivatives of the original functional are equal to 0. So, in effect, in this case, we have regular physical equations – unaffected by constraints. We have already mentioned that in this case, we often get ill-defined solutions.
– The case when the constrains do affect the solutions is when that the optimum of the functional is attained on the border of the area defined by the constraints. In the above example, it means that the optimum is attained when $\int f_{,i} \cdot f^{,i}\,dx = \Delta$.

Therefore, in cases when constrains are important to impose (and do not just come satisfied "for free" already for the usual solution), the inequality-type constraints are equivalent to equality-type ones.

Optimization under such equality constraints is done by using the usual Lagrange multiplier approach: optimizing a functional F under a constraint $G = g_0$ (i.e., equivalently, $G - g_0 = 0$) is equivalent, for an appropriate real number λ, to an unconstraint optimization of an auxiliary functional $F + \lambda \cdot (G - g_0)$. The value λ must then be found from the constraint $G = g_0$.

In the above example, optimizing a functional $\int f^2\,dx$ under a constraint $\int f_{,i} \cdot f^{,i}\,dx = \Delta$ is equivalent to an unconstrained optimization of the auxiliary functional

$$\int (f^2 + \lambda \cdot f_{,i} \cdot f^{,i})\,dx.$$

Observation. The action functionals corresponding to fundamental physics theories already have a term proportional to $f_{,i} \cdot f^{,i}$ for a scalar field $f(x)$ or proportional to similar terms for more complex fields (vector, tensor, spinor, etc.)

Discussion. At present, this is what physicists are doing:

- They start with the (action) functionals $S = \int L\,dx$ corresponding to fundamental physical phenomena. These action functionals already have terms proportional to $f_{,i} \cdot f^{,i}$.
- Based on these action functionals, physicists derive the corresponding differential equations $\dfrac{\delta L}{\delta f} = 0$.
- A direct solution to the resulting differential equations is ill-defined (too much influenced by noise).
- Thus, instead of directly solving these equations, physicists *regularize* them, i.e., solve them under the constraints of the type $\int f_{,i} \cdot f^{,i}\,dx = \Delta$.

As we have mentioned, from the mathematical viewpoint, the regularization constrains are equivalent to adding terms of the type $f_{,i} \cdot f^{,i}$ to the corresponding Lagrangians. But these Lagrangians already have such terms! So, we arrive at a natural idea.

Idea. Traditionally, in fundamental physics, we assume that we have an *unconstrained* optimization $S = \int L\,dx \to \min$. A natural idea is to assume that in reality, the physical world corresponds to *constrained* optimization $F \to min$ under a constraint $G = g_0$ – and place terms like $f_{,i} \cdot f^{,i}$ into the constraint.

It is simply a re-interpretation. At first glance, the above idea may sound like a sacrilege: a group of non-physicists challenge Einstein's equations? But we are *not* suggesting to change the equations, the differential equations – the only thing that we can check by observation – remain exactly the same. What we propose to change is the *interpretation* of these equations:

- Traditionally, these equations are interpreted via *unconstrained* optimization.
- We propose to interpret them via *constrained* optimization.

What do we gain? One might ask: if we are not proposing new equations, if we are not proposing any new physical theory, then what do we gain?

Our main gain is that we now have a built-in regularization. We do not need to worry about an additional outside regularization step anymore. We can not be sure that our problems are well-defined.

Possible additional gain. There may also be an additional gain, with respect to quantum versions of the fundamental physical theories. In contrast to the non-quantum field theory, in the quantum versions, if we impose the constraints, we do limit quantum solutions – because now, we are requiring the actual field to satisfy the additional constraint, while in the quantum case, all fields are possible (although with different probabilities). In quantum field theory, such absolute constraints are known as *super-selection* rules; see, e.g., [6]. It is known that such rules help to decrease divergence in quantum field theories (i.e., help them avoid these theories leading to meaningless infinite predictions); so maybe super-selection rules coming from our constrains will also be of similar help.

Possible philosophical meaning of our proposal. In addition to a pragmatic meaning (well-foundedness of the problem, possible decrease in divergence, etc.), our proposal may have a deeper philosophical meaning. To discuss such a meaning, let us consider the simplest possible case of a scalar field $f(x)$ corresponding to a particle of rest mass m. In the traditional field theory, its Lagrangian has the form $L = m^2 \cdot f^2 + f_{,i} \cdot f^{,i}$. For this theory, our proposal is, in effect, to make $\int f^2\,dx$ an optimized function, and to introduce a constraint $\int f_{,i} \cdot f^{,i}\,dx = g_0$.

When we apply the Lagrange multiplier to this constrained optimization problem, we get the Lagrangian $L = f^2 + \lambda \cdot f_{,i} \cdot f^{,i}$ whose minimization is equivalent to minimizing $L' = \lambda{-}1 \cdot L = \lambda^{-1} \cdot f^2 + f_{,i} \cdot f^{,i}$. In other words, we recover the original Lagrangian, with $m^2 = \lambda^{-1}$. Now, in contrast to the traditional interpretation, the rest mass m is no longer the original fundamental parameter – it is a Lagrange multiplier that needs to be adjusted to fit the actual fundamental constant g_0 (which should be equal to $\int f_{,i} \cdot f^{,i}\,dx$).

Thus, the particle masses are no longer original fundamental constants – they depend on the fields in the rest of the world. This idea may sound somewhat heretic to a non-physicist, but it is very familiar to those who studied history of modern physics. This general philosophical idea – that all the properties like inertia, mass, etc. depend on the global configuration of the world – was promoted by a 19 century physicist Ernst Mach (see, e.g., [3]), and it was one of the main ideas that inspired Einstein to formulate his General Relativity theory [4], a theory in which what Einstein called *Mach's principle* is, to some extent, satisfied.

In other words, our idea may sound, at first glance, philosophically somewhat heretical, but it seems to be in line with Einstein's philosophical foundations for General Relativity.

Acknowledgments. This work was supported in part by the NSF grants HRD-0734825 and DUE-0926721, and by Grant 1 T36 GM078000-01 from NIH.

References

1. Feynman, R., Leighton, R., Sands, M.: The Feynman Lectures on Physics, Addison Wesley, Boston, Massachusetts (2005)
2. Gelfand, I.M., Fomin, S.V.: Calculus of Variations, Dover Publ., New York (2000)
3. Mach, E.: The Science of Mechanics; a Critical and Historical Account of its Development. Open Court Pub. Co., LaSalle, Illinois (1960)
4. Misner, C.W., Thorne, K.S., Wheeler, J.A.: Gravitation. W.H. Freeman, New York (1973)
5. Tikhonov, A.N., Arsenin, V.Y.: Solutions of Ill-Posed Problems, W. H. Whinston & Sons, Washington, D.C. (1977)
6. Weinberg, S.: The Quantum Theory of Fields. Vol. I. Foundations. Cambridge University Press, Cambridge (1995)