Equations Without Equations: Challenges on a Way to a More Adequate Formalization of Reasoning in Physics

Roberto Araiza, Vladik Kreinovich,
and Juan Ferret
University of Texas, El Paso, TX 79968, USA
raraiza@gmail.com, vladik@utep.edu

1. Need to Formalize Reasoning in Physics

- *Fact:* in medicine, geophysics, etc., expert systems use automated expert reasoning to help the users.
- *Hope:* similar systems may be helpful in general theoretical physics as well.
- *What is needed:* describe physicists' reasoning in precise terms.
- *Reason:* formalize this reasoning inside an automated computer system.
- *Formalized part of physicists' reasoning:* theories are formulated in terms of PDEs (or ODEs) $\frac{dx}{dt} = F(x)$.
- *Meaning:* these equations describe how the corresponding fields (or quantities) x change with time t.

2. Mathematician's View of Physics and Its Limitations

- *Mathematician's view:* we know the initial conditions $x(t_0)$ at some moment of time t_0.
- We solve the corresponding Cauchy problem and find the values $x(t)$ for all t.
- *Limitation:* not all solutions to the equation $\frac{dx}{dt} = F(x)$ are physically meaningful.
- *Example 1:* when a cup breaks into pieces, the corresponding trajectories of molecules make physical sense.
- *Example 2:* when we reverse all the velocities, we get pieces assembling themselves into a cup.
- *Fact:* this is physically impossible.
- *Fact:* the reverse process satisfies all the original (T-invariant) equations.

3. Physicists' Explanation

- *Reminder:* not all solutions to the physical equation are physically meaningful.
- *Explanation:* the "time-reversed" solution is non-physical because its initial conditions are "degenerate".
- *Clarification:* once we modify the initial conditions even slightly, the pieces will no longer get together.
- *Conclusion:* not only the equations must be satisfied, but also the initial conditions must be "non-degenerate".
- *Two challenges* in formalizing this idea:
  - how to formalize "non-degenerate";
  - the separation between equations and initial conditions depends on the way equations are presented.
- *First challenge:* can be resolved by using Kolmogorov complexity and randomness.

4. First Example: Schrödinger's Equation

- *Example:* Schrödinger's equation

$$i\hbar \cdot \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \cdot \nabla^2 \Psi + V(\vec{r}) \cdot \Psi.$$

- *In this representation:* the potential V is a part of the equation, and $\Psi(\vec{r}, t_0)$ are initial conditions.

- *Transformation:*
  - we represent $V(\vec{r})$ as a function of Ψ and its derivatives,
  - differentiate the right-hand side by time, and
  - equate the derivative w.r.t. time to 0.

- *Result:*

$$\frac{\partial}{\partial t}\left(\frac{i\hbar}{\Psi} \cdot \frac{\partial \Psi}{\partial t} + \frac{\hbar^2}{2m} \cdot \frac{\nabla^2 \Psi}{\Psi}\right) = 0.$$

5. First Example (cont-d)

- *Reminder:*

$$\frac{\partial}{\partial t}\left(\frac{\mathrm{i}\hbar}{\Psi}\cdot\frac{\partial\Psi}{\partial t}+\frac{\hbar^2}{2m}\cdot\frac{\nabla^2\Psi}{\Psi}\right)=0.$$

- *Mathematically:* the new equation (2nd order in time) is equivalent to the Schrödinger's equation:
  - every solution of the Schrödinger's equation for any $V(\vec{r})$ satisfies this new equation, and
  - every solution of the new equation satisfies Schödinger's equation for some $V(\vec{r})$.
- *Observation:* in the new equation, initial conditions, in effect, include $V(\vec{r})$.
- *Conclusion:* "non-degeneracy" ("randomness") condition must now include $V(\vec{r})$ as well.

6. Towards 2nd Example: General Physical Theories

- *Traditional description of physical theories:* in terms of differential equations.
- *Example (17 cent.):* Newton's mechanics $m \cdot \dfrac{d^2x}{dt^2} = F$.
- *Important discovery (18 cent.):* most physical theories can be reformulated as $S \to \min$ for "action" S.
- *Example:* Newton's mechanics is equivalent to $S = \int L\,dt \to \min$, where $L = \dfrac{1}{2} \cdot m \cdot \dot{x}^2 + V(x)$.
- *For functions $f(x_1, \ldots, x_n)$:* minimum when $f(x + dx) \approx f(x)$, so $\dfrac{\partial f}{\partial x_i} = 0$ for all i.
- *For functions of functions ("functionals"):* minimum when $S(f + \delta f) \approx S(f)$, so $\dfrac{\delta S}{\delta f}(x) = 0$ for all x.

7. Euler-Lagrange Equations

- *Reminder:* physical theories can be formulated in terms of the minimal action principle $S \to \min$.
- Here, $S = \int L\,dx$ for a "Lagrange" f-n L that depends on the fields φ, ..., and their derivatives $\varphi_{,i} \stackrel{\text{def}}{=} \dfrac{\partial \varphi}{\partial x_i}$.
- *Euler-Lagrange equations:* when $S = \int L\,dx$,

$$\frac{\delta S}{\delta f} = \frac{\partial L}{\partial f} - \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial f_{,i}}\right) = 0.$$

- *Comment:* we use "Einstein's rule" that repeated indices mean summation: e.g., $f_{,i} f_{,i}$ means $\sum\limits_i f_{,i} f_{,i}$.
- *For a single scalar field φ:*

$$\frac{\partial L}{\partial \varphi} - \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial \varphi_{,i}}\right) = 0.$$

8. Second Example: General Scalar Field

- *General scalar theory:* $L = L(\varphi, \varphi_{,i})$.
- *3-D case:* it is reasonable to consider rotation-invariant Lagrangian functions L.
- *Conclusion:* L depends only on the length $\varphi_{,i}\varphi^{,i}$ of the vector $\varphi_{,i}$,not on its orientation.
- *4-D case:* L should be invariant w.r.t. Lorentz transformations (4-D "rotations").
- *Conclusion:* $L = L(\varphi, a)$, where $a \stackrel{\text{def}}{=} \varphi_{,i}\varphi^{,i}$.
- *Traditional formulation:* every Lagrangian is possible, but initial conditions $\varphi(x, t_0)$ must be non-degenerate.
- *Our result:* there exists a 3rd order equation such that:

$$\varphi \text{ satisfies this equation} \Leftrightarrow$$
$$\varphi \text{ satisfies Euler-Lagrange equation for } \textit{some } L.$$

9. Scalar Field: Proof

- *Reminder:* $L = L(\varphi, a)$, where $a \stackrel{\text{def}}{=} \varphi_{,i}\varphi^{,i}$.
- *Euler-Lagrange equations:* $\dfrac{\partial L}{\partial \varphi} - \partial_i \dfrac{\partial L}{\partial \varphi_{,i}} = 0$.
- *Using chain rule:* $\dfrac{\partial L(\varphi, a)}{\partial \varphi_{,i}} = \dfrac{\partial L}{\partial a} \cdot \dfrac{\partial a}{\partial \varphi_{,i}} = \dfrac{\partial L}{\partial a} \cdot 2\varphi^{,i}$.
- *Conclusion:* $L_{,\varphi} - \partial_i(2L_{,a} \cdot \varphi_{,i}) = 0$.
- *Using chain rule again*, we get

  $$L_{,\varphi} - 2L_{,a} \cdot \Box\,\varphi - 2L_{,a\varphi} \cdot (\varphi_{,i}\varphi^{,i}) - 4L_{,aa} \cdot \varphi_{,ij}\varphi^{,i}\varphi^{,j} = 0,$$

  where $\Box\varphi \stackrel{\text{def}}{=} \varphi^{,i}_{,i}$.
- *Conclusion:*
  - if at two points, we have the same values of φ, $\varphi_{,i}\varphi^{,i}$, and $\Box\,\varphi$,
  - then we have same values of $\varphi_{,ij}\varphi^{,i}\varphi^{,j}$.

10. Scalar Field: Proof (cont-d)

- *Reminder:* if at two points, we have the same values of φ, $a = \varphi_{,i}\varphi^{,i}$, and $b \stackrel{\text{def}}{=} \Box\,\varphi$, then we have same values of $c \stackrel{\text{def}}{=} \varphi_{,ij}\varphi^{,i}\varphi^{,j}$.
- *Particular case:* if we have dx^k for which $\varphi_{,k}\cdot dx^k = 0$, $a_{,k}\cdot dx^k = 0$, and $b_{,k}\cdot dx^k = 0$, then $c_{,k}\cdot dx^k = 0$.
- *In geom. terms:* if $dx^k \perp \varphi_{,k}$, $dx^k \perp a_{,k}$, and $dx^k \perp b_{,k}$, then $dx^k \perp c_{,k}$.
- *Conclusion:* $\varphi_{,k}$, $a_{,k}$, $b_{,k}$, and $c_{,k}$ lie in the same 3-plane.
- *In algebraic terms:* the determinant is 0:

$$\varepsilon_{ijkl}\cdot\varphi_{,i}\cdot a_{,j}\cdot b_{,k}\cdot c_{,l} = 0,$$

  where $\varepsilon_{ijkl} = 0$ if some indices are equal and is ± 1 else.
- We get a 3-rd order equation; so, we can predict future evolution – w/o knowing L.

11. Scalar Field: Discussion and Conclusions

- *Observation:* the new "equation" does not contain L at all.
- *Fact:* a field φ satisfies the new equation $\Leftrightarrow$ it satisfies the Euler-Lagrange equations for *some* L.
- *Observation:*
  - similarly to Wheeler's cosmological "mass without mass" and "charge without charge",
  - we now have "equations without equations".
- *Conclusion:* when formalizing physical equations:
  - we must not only describe them in *a* mathematical form,
  - we must also select *one* of the mathematically equivalent forms.

12. Interesting Relation to Dimension of Space-Time

- *Reminder:* our conclusion is based on the idea that four vectors lie in a 3-D plane.
- *Observation:* if the dimension of space-time is 3 or smaller, this is always true.
- *Conclusion:* "equations without equations" are only possible when dimension is ≥ 4.
- *Speculation:* maybe this explains why our space-time is 4-D?

13. Interesting Relation to Dimension of Space-Time (cont-d)

- *What about 2 scalar fields φ and ψ:* here, preservation of 10 quantities
$$\varphi, \psi, \varphi_{,i}\varphi^{,i}, \psi_{,i}\psi^{,i}, \varphi_{,i}\psi^{,i}, \varphi_{,ij}\varphi^{,i}\varphi^{,j}, \varphi_{,ij}\varphi^{,i}\psi^{,j}, \varphi_{,ij}\psi^{,i}\psi^{,j},$$
$$\psi_{,ij}\varphi^{,i}\varphi^{,j}, \psi_{,ij}\varphi^{,i}\psi^{,j}, \psi_{,ij}\psi^{,i}\psi^{,j}$$
means that $\Box\varphi$ and $\Box\psi$ are the same.
- *Conclusion:* 11 vectors (gradients of the above quantities) and $(\Box\varphi)_{,k}$ must be in the same 11-D space.
- *Observation:* this requirement is always true in spaces of dimension ≤ 11.
- *Conclusion:* for 2 scalar fields, equations w/o equations are possible in dim ≥ 12.
- *Is this physical?* yes: consistent quantum field theory is only possible when dim ≥ 11.

14. Acknowledgments

This work was supported in part:

- by National Science Foundation grant HRD-0734825, and EAR-0225670 and DMS-0532645 and
- by Grant 1 T36 GM078000-01 from the National Institutes of Health.

15. Physicists Assume that Initial Conditions and Values of Parameters are Not Abnormal

- To a mathematician, the main contents of a physical theory is its equations.
- Not all solutions of the equations have physical sense.
- *Ex. 1:* Brownian motion comes in one direction;
- *Ex. 2:* implosion glues shattered pieces into a statue;
- *Ex. 3:* fair coin falls heads 100 times in a row.
- *Mathematics:* it is possible.
- *Physics* (and common sense): it is not possible.
- *Our objective:* supplement probabilities with a new formalism that more accurately captures the physicists' reasoning.

16. A Seemingly Natural Formalizations of This Idea

- *Physicists:* only "not abnormal" situations are possible.
- *Natural formalization: idea.*
  - If a probability $p(E)$ of an event E is small enough,
  - then this event cannot happen.
- *Natural formalization: details.* There exists the "smallest possible probability" p_0 such that:
  - if the computed probability p of some event is larger than p_0, then this event can occur, while
  - if the computed probability p is $\le p_0$, the event cannot occur.
- *Example:* a fair coin falls heads 100 times with prob. 2^{-100}; it is impossible if $p_0 \ge 2^{-100}$.

17. The Above Formalization of the Notion of "Typical" is Not Always Adequate

- *Problem:* every sequence of heads and tails has exactly the same probability.
- *Corollary:* if we choose $p_0 \geq 2^{-100}$, we will thus exclude all sequences of 100 heads and tails.
- However, anyone can toss a coin 100 times.
- This proves that some such sequences are physically possible.
- *Similar situation:* Kyburg's lottery paradox:
  - in a big (e.g., state-wide) lottery, the probability of winning the Grand Prize is very small;
  - a reasonable person should not expect to win;
  - however, some people do win big prizes.

18. New Idea

- *Example:* height:
  - if height is ≥ 6 ft, it is still normal;
  - if instead of 6 ft, we consider 6 ft 1 in, 6 ft 2 in, etc., then $\exists h_0$ s.t. everyone taller than h_0 is abnormal;
  - we are not sure what is h_0, but we are sure such h_0 exists.
- *General description:* on the universal set U, we have sets $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq \ldots$ s.t. $\cap A_n = \emptyset$.
- *Example:* $A_1 =$ people w/height ≥ 6 ft, $A_2 =$ people w/height ≥ 6 ft 1 in, etc.
- A set $T \subseteq U$ is called a *set of typical (not abnormal) elements* if

$$\forall \text{ definable sequence of sets } A_n \text{ for which } A_n \supseteq A_{n+1} \text{ for all } n \text{ and } \cap A_n = \emptyset, \exists N \text{ for which } A_N \cap T = \emptyset.$$

19. Coin Example

- Universal set $U = \{\mathrm{H}, \mathrm{T}\}^{\mathbb{N}}$
- Here, A_n is the set of all the sequences that start with n heads and have at least one tail.
- The sequence $\{A_n\}$ is decreasing and definable, and its intersection is empty.
- Therefore, for every set T of typical elements of U, there exists an integer N for which $A_N \cap T = \emptyset$.
- This means that if a sequence $s \in T$ is not abnormal and starts with N heads, it must consist of heads only.
- *In physical terms:* it means that

  a random sequence (i.e., a sequence that contains both heads and tails) cannot start with N heads.

- This is exactly what we wanted to formalize.

20. Possible Practical Use of This Idea: When to Stop an Iterative Algorithm

- *Situation* in numerical mathematics:
  - we often know an iterative process whose results x_k are known to converge to the desired solution x,
  - but we do not know when to stop to guarantee that

$$d_X(x_k, x) \le \varepsilon.$$

- *Heuristic approach:* stop when $d_X(x_k, x_{k+1}) \le \delta$ for some $\delta > 0$.
- *Example:* in physics, if 2nd order terms are small, we use the linear expression as an approximation.

21. Result

- Let $\{x_k\} \in S$, k be an integer, and $\varepsilon > 0$ a real number.
- We say that x_k is $\varepsilon-$*accurate* if $d_X(x_k, \lim x_p) \le \varepsilon$.
- Let $d \ge 1$ be an integer.
- By a *stopping criterion*, we mean a function $c : X^d \to R_0^+ = \{x \in R \,|\, x \ge 0\}$ that satisfies the following two properties:
  - If $\{x_k\} \in S$, then $c(x_k, \ldots, x_{k+d-1}) \to 0$.
  - If for some $\{x_n\} \in S$ and k, $c(x_k, \ldots, x_{k+d-1}) = 0$, then $x_k = \ldots = x_{k+d-1} = \lim x_p$.
- *Result:* Let c be a stopping criterion. Then, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that
  - if $c(x_k, \ldots, x_{k+d-1}) \le \delta$, and the sequence $\{x_n\}$ is not abnormal,
  - then x_k is $\varepsilon-$accurate.