University of Texas at El Paso

From the SelectedWorks of Juan Ferret

January 2010

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

Contact Author

Start Your Own SelectedWorks

Notify Me of New Work

Available at: http://works.bepress.com/juan_ferret/11

1

EQUATIONS WITHOUT EQUATIONS: CHALLENGES ON A WAY TO A MORE ADEQUATE FORMALIZATION OF CAUSALITY REASONING IN PHYSICS, BY ROBERTO ARAIZA, VLADIK KREINOVICH, AND JUAN FERRET

1.1 Need to formalize causality reasoning in physics

Our interest in causality reasoning in physics was enhanced by the need to formalize reasoning in physics. In medicine, geophysics, and in many other applications areas, expert systems use automated expert reasoning to help the users; see, e.g., (2).

We expect that similar systems may be helpful in general theoretical physics as well. To design such systems, we must formalize physicists' reasoning inside an automated computer system. To do this, we must be able to describe physicists' reasoning in precise terms.

Physicists' reasoning has a clearly formalized part: indeed, physical theories are usually formulated in terms of differential equations

$$\frac{dx}{dt} = F(x) \tag{1.1}$$

(in general, partial differential equations) that describe how the corresponding fields (and/or physical quantities) x change with time t.

From the purely mathematical viewpoint, the situation seems to be straightforward:

- from the observations, we find the initial conditions $x(t_0)$ at some moment of time t_0;
- we then solve the differential equation (1.1) and find the values $x(t)$ for all moments of time t.

In mathematical physics, the problem of finding the solution to the differential equation with the known initial conditions is called the *Cauchy problem*.

In physical terms, this mathematical description reflects the basic idea of causality: the future dynamic of the system is caused by the initial state $x(t_0)$.

The above description does not capture the known fact that not all solutions to the equation (1.1) are physically meaningful. For example, when a cup breaks into pieces, the corresponding trajectories of molecules make physical sense. If we now reverse all the velocities, we get pieces spontaneously assembling themselves

into a cup. This reverse process is clearly physically impossible. However, since Newton's equations remain valid when we change the time direction $t \to -t$ (i.e., are, in physical terms, *T-invariant*), the reverse process also satisfies the Newton's equations.

To provide an adequate description of physicists' reasoning, we must be able to capture not only the equations, but also the fact that some solutions of these equations are not physical. A usual physical explanation of this fact is that, e.g., the "time-reversed" solution is non-physical because its initial conditions are "degenerate". Specifically, this means that if we modify the initial conditions even slightly, the pieces will no longer get together.

So, for a solution to be physically meaningful, not only the equations must be satisfied, but also the initial conditions must be "non-degenerate". It is known that the notion of non-degeneracy in physics can be adequately described by the use of Kolmogorov complexity and algorithmic randomness; see, e.g., (3; 4).

However, there is another important challenge that it not easy to resolve: that the separation between equations and initial conditions depends on the way equations are presented.

In this paper, we will illustrate this dependence on the example of simple physical equations.

1.2 First example: Schrödinger's equation

As our first example, let us consider Schrödinger's equation, the basic equation of non-relativistic quantum mechanics:

$$\mathrm{i} \cdot \hbar \cdot \frac{\partial \Psi}{\partial t}(x,t) = -\frac{\hbar^2}{2 \cdot m} \cdot \nabla^2 \Psi(x,t) + V(x) \cdot \Psi(x,t). \tag{1.2}$$

This equation describes the dynamic of a wave function $\Psi(x,t)$ that characterizes the state of a particle of mass m in the potential field $V(x)$.

In this representation, the potential $V(x)$ is a part of the equation, and the initial conditions are the values $\Psi(x,t_0)$ of the wave function at the moment t_0.

On the other hand, we can undertake the following simple transformation of the equation (1.2). First, by moving the right-hand side term not containing $V(x)$ to the left-hand side and dividing both sides by $\Psi(x,t)$, we can represent the potential field $V(x)$ as a function of the wave function $\Psi(x,t)$ and its derivatives:

$$V(x) = \frac{\mathrm{i} \cdot \hbar}{\Psi} \cdot \frac{\partial \Psi}{\partial t} + \frac{\hbar^2}{2 \cdot m} \cdot \frac{\nabla^2 \Psi}{\Psi}. \tag{1.3}$$

The left-hand side $V(x)$ of this equation does not depend on time. So, if we differentiate the right-hand side of the equation (1.3) by time t, we should get 0:

$$\frac{\partial}{\partial t}\left(\frac{\mathrm{i} \cdot \hbar}{\Psi} \cdot \frac{\partial \Psi}{\partial t} + \frac{\hbar^2}{2 \cdot m} \cdot \frac{\nabla^2 \Psi}{\Psi}\right) = 0. \tag{1.4}$$

The original Schrödinger's equation is first order in terms of time; the new equation (1.4) is second order with respect to time.

We have shown that every solution of the Schrödinger's equation for any potential field $V(x)$ satisfies this new equation (1.4). Let us show that this new equation is actually *equivalent* to the Schrödinger's equation in the sense that

- every solution of the Schrödinger's equation for any $V(x)$ satisfies this new equation, and
- every solution of the new equation satisfies Schödinger's equation for some potential field $V(x)$.

Indeed, if the equation (1.4) is satisfied, i.e., the time derivative of the corresponding expression

$$\frac{\mathrm{i}\cdot\hbar}{\Psi}\cdot\frac{\partial\Psi}{\partial t}+\frac{\hbar^2}{2\cdot m}\cdot\frac{\nabla^2\Psi}{\Psi}. \tag{1.5}$$

is equal to 0, this means that this expression cannot depend on time t, it can only depend on the spatial coordinates x. If we denote this dependence by $V(x)$:

$$V(x)\stackrel{\text{def}}{=}\frac{\mathrm{i}\hbar}{\Psi}\cdot\frac{\partial\Psi}{\partial t}+\frac{\hbar^2}{2m}\cdot\frac{\nabla^2\Psi}{\Psi}, \tag{1.6}$$

then, by multiplying both sides of this equation by Ψ and moving the terms containing the second derivative to the other side, we get the Schrödinger's equation (1.2).

From the purely mathematical viewpoint, the equations (1.2) and (1.6) are equivalent. However, from the viewpoint of causality reasoning, these equations are different:

- In the original Schrödinder's equation, the potential field $V(x)$ was a part of the equation. So, the "non-degeneracy" restrictions on initial conditions did not limit possible fields.
- In contrast, in the new equation (1.4), initial conditions, in effect, include $V(x)$. Thus, we are no longer allowing arbitrary potential fields: the "non-degeneracy" ("randomness") condition must now include $V(x)$ as well.

It is worth mentioning that the transformation leading to the equation (1.4) is not simply a purely mathematical trick. The new equation can be used to find the wave functions by solving the corresponding Cauchy problem:

- In the Schrödinger's equations, once we know the initial values $\Psi(x,t_0)$ of the wave function for all locations x, we can predict its values $\Psi(x,t)$ at all future moments of time.
- Similarly, once we know the initial values of the wave function $\Psi(x,t_0)$ and its first time derivative $\dfrac{\partial\Psi}{\partial t}(x,t_0)$ for all locations x, we can predict its values $\Psi(x,t)$ at all future moments of time t.

Comment. Readers must be cautioned that while theoretically, this reconstruction is possible, its practical applicability may be limited. Indeed, the corresponding equation (1.4) requires that divide by $\Psi(x,t)$. Thus, in areas where the

value of the wave function $\Psi(x,t)$ is close to 0, we will need to measure this value with a very high accuracy to get practical results.

1.3 Towards second example: general physical theories

In the previous section, we discovered an interesting feature of the Schrödinger's equation, a specific equation describing a single physical field $\Psi(x,t)$. The exact derivation of our result used specific features of this equation. It is therefore natural to ask: how important are these features for our derivation? Is our result true only for the Schrödinger equation – or is it also true for an arbitrary theory describing a scalar field?

To answer this question, we need to recall how general physical theories are described. Starting from the 17 century Newton's mechanics, physical theories have been usually described in terms of differential equations. For example, Newton's mechanics is described by the equation

$$m \cdot \ddot{x} = F, \tag{1.7}$$

where $\ddot{x} \stackrel{\text{def}}{=} \dfrac{d^2x}{dt^2}$ and $F = -\nabla V$ is a force field corresponding to the potential field $V(x)$.

In the 18 century, an important discovery was made: that most physical theories can be reformulated in terms of the *least action principle*, i.e., as a requirement that a certain functional S (called *action*) attain its smallest possible value: $S \to \min$; see, e.g., (1). For example, Newton's mechanics (1.7) is equivalent to minimizing the functional $S = \int L\,dt \to \min$, where

$$L = \frac{1}{2} \cdot m \cdot (\dot{x})^2 - V(x)$$

is the difference between the kinetic energy $\dfrac{1}{2} \cdot m \cdot (\dot{x})^2$ and the potential energy $V(x)$ of the system.

The least action principle, originally proposed by Pierre Louis Moreau de Maupertuis, enables us to derive optical trajectories (such as Snell's law) from a simple requirement that the light trajectories minimize travel time. This principle enables us to come up with good approximate solutions to electrostatic problems (1).

The least action formulation is, at present, the main way of presenting physical theories.

For example, the most widely used quantization technique – the technique of Feynman integrals over trajectories – describes the (complex-valued) amplitude of a transition between the two states as the normalized sum $\dfrac{1}{N} \cdot \sum \exp\left(\mathrm{i} \cdot \dfrac{S}{\hbar}\right)$, where N is a normalization constant, the sum is taken over all trajectories connecting the two states, and S is the value of the action along the trajectory.

Since few non-physicists are well familiar with the least action formulation of physical theories, let us briefly (and informally) describe the main formulas for the interested readers.

How can we find the minimum of a functional? This question is similar to the question of finding the minimum of a function, the question which was successfully solved by calculus. Specifically, for a function $f(x)$ of one variable, its derivative $\dfrac{df}{dx}$ is defined as the limit of the ratio $\dfrac{f(x+\Delta x)-f(x)}{\Delta x}$ when $\Delta x \to 0$. Informally, the limit means that when the value Δx is small (i.e., close to 0), then the derivative is close to this ratio:

$$\frac{df}{dx} \approx \frac{f(x+\Delta x)-f(x)}{\Delta x}. \tag{1.8}$$

Multiplying both sides of this approximate equality by Δx and moving the term proportional to Δx to the other side, we get

$$f(x+\Delta x) \approx f(x) + \frac{df}{dx} \cdot \Delta x. \tag{1.9}$$

When the function $f(x)$ attains its minimum at a point x_0, then, according to this formula:

- the derivative $\dfrac{df}{dx}$ cannot be positive, since then we have $f(x_0+\Delta x) < f(x_0)$ for negative $\Delta x < 0$, which contradicts to our assumption that the values $f(x_0)$ is the smallest; and
- the derivative $\dfrac{df}{dx}$ cannot be negative, since then we have $f(x_0+\Delta x) < f(x_0)$ for positive $\Delta x > 0$.

Thus, the only possibility is to have $\dfrac{df}{dx} = 0$. This is how calculus recommends to search for the minimum of a function of one variable: by finding the values where the derivative is equal to 0.

For a function of several variables $f(x) = f(x_1, \ldots, x_n)$, if we change the variable x_1, we similarly get

$$f(x_1+\Delta x_1, x_2, \ldots, x_n) \approx f(x_1, \ldots, x_n) + \frac{\partial f}{\partial x_1} \cdot \Delta x_1. \tag{1.10}$$

When we then change the second variable x_2, we get

$$f(x_1+\Delta x_1, x_2+\Delta x_2, x_3, \ldots, x_n) \approx f(x_1+\Delta x_1, x_2, \ldots, x_n) + \frac{\partial f}{\partial x_2} \cdot \Delta x_2, \tag{1.11}$$

i.e., substituting (1.10) into the right-hand side of (1.11), that

$$f(x_1+\Delta x_1, x_2+\Delta x_2, x_3, \ldots, x_n) \approx$$

$$f(x_1, x_2, \ldots, x_n) + \frac{\partial f}{\partial x_1} \cdot \Delta x_1 + \frac{\partial f}{\partial x_2} \cdot \Delta x_2. \tag{1.12}$$

By changing all the other variables one by one, we thus conclude that

$$f(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n) \approx$$

$$f(x_1, x_2, \ldots, x_n) + \frac{\partial f}{\partial x_1} \cdot \Delta x_1 + \ldots + \frac{\partial f}{\partial x_n} \cdot \Delta x_n. \tag{1.13}$$

The same argument as for the function of one variable shows that minimum is attained when all the coefficients at Δx_i are equal to 0, i.e., when $\dfrac{\partial f}{\partial x_i} = 0$ for all i.

A functional $S(f)$ is a function whose inputs are also functions $f(x)$. Each function $f(x)$ requires infinitely many parameters to describe – e.g., its values $f(x)$ at different locations x. Thus, a functional can be viewed as a function of infinitely many variables $f(x)$ – a family of variables which are described by a continuous parameter x. It turns out that for many functionals, we can have a formula similar to (1.13) – with the only difference that instead of a finite sum, we have a natural continuous limit of the sum – an integral:

$$S(f + \Delta f) \approx S(f) + \int \frac{\delta S}{\delta f}(x) \cdot \Delta f(x)\, dx, \tag{1.14}$$

for appropriate "variational derivatives" $\dfrac{\delta S}{\delta f}$. If one of the values of the variational derivative is non-zero, i.e., positive or negative, then we can have $S(f + \Delta f) < S(f)$ for appropriate Δf. Thus, at the point where the minimum is attained, all these derivatives are equal to 0:

$$\frac{\delta S}{\delta f} = 0. \tag{1.15}$$

In physics, the action S usually takes the form $S = \int L\, dx$ for a "Lagrange" function L that depends on the fields φ, ..., and their derivatives $\varphi_{,i} \stackrel{\text{def}}{=} \dfrac{\partial \varphi}{\partial x_i}$. For such functional, the *Euler-Lagrange equations* describe the minimum:

$$\frac{\delta S}{\delta f} = \frac{\partial L}{\partial f} - \sum_i \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial f_{,i}}\right) = 0. \tag{1.16}$$

(For interested readers, a simple derivation of this formula is given in the Appendix.)

To simplify formulas, we will use "Einstein's rule" according to which repeated indices mean summation: e.g., $f_{,i}f_{,i}$ means $\sum\limits_i f_{,i}f_{,i}$. This rule enables us to somewhat simplify the formula (1.16) into

$$\frac{\delta S}{\delta f} = \frac{\partial L}{\partial f} - \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial f_{,i}}\right) = 0. \tag{1.17}$$

In particular, for the case of a a single scalar field φ, this equation takes the form

$$\frac{\partial L}{\partial \varphi} - \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial \varphi_{,i}}\right) = 0. \tag{1.18}$$

Now, we are ready to describe the second example.

1.4 Second example: general scalar field

A general scalar theory, with a single scalar field φ, is characterized by a Lagrange function $L = L(\varphi, \varphi_{,i})$ that depends on this field and on its derivatives $\varphi_{,i}$.

In the 3-D case, it is reasonable to consider rotation-invariant Lagrangian functions L. Rotation invariance means that L depends only on the (square of the) length

$$\varphi_{,i}\varphi^{,i} = (\varphi_{,1})^2 + (\varphi_{,2})^2 + (\varphi_{,3})^2$$

of the vector $\varphi_{,i}$, not on its orientation.

Similarly, in 4-D case, it is reasonable to require that the Lagrangian L be invariant w.r.t. Lorentz transformations (4-D "rotations"). Under this requirement, we can similarly conclude that $L = L(\varphi, a)$, where

$$a \stackrel{\text{def}}{=} \varphi_{,i}\varphi^{,i} = (\varphi_{,0})^2 - (\varphi_{,1})^2 - (\varphi_{,2})^2 - (\varphi_{,3})^2\,.$$

By a general scalar theory, we will therefore understand a theory with a Lagrangian of this type.

Our main result is that there exists a third order equation such that:

- if φ satisfies the Euler-Lagrange equation for *some* Lorentz-invariant L,
- then φ satisfies this new equation.

The proof of this result is given in the next section.

From the purely mathematical viewpoint, the new equation is a consequence of the standard Euler-Lagrange equations. However, from the viewpoint of causality reasoning, the new equation is drastically different from the usual ones. Indeed:

- In the original description of the scalar theories, the Lagrangian $L(\varphi, a)$ is incorporated in the equation, so "non-degeneracy" restrictions on initial conditions did not limit possible Lagrangians.
- In contrast, in the new equation, initial conditions, in effect, include $L(\varphi, a)$. Thus, we are no longer allowing arbitrary Lagrangians: the "non-degeneracy" ("randomness") condition must now include $L(\varphi, a)$ as well.

Similarly to the case of the Schrödinger equation, the transformation leading to the new equation is not simply a purely mathematical trick. The new equation can be used to find the scalar function $\varphi(x)$ by solving the corresponding Cauchy problem:

- In the original second-order Euler-Lagrange equations, once we know the initial values $\varphi(t_0, x_1, \ldots, x_n)$ of the scalar function and its first time derivative $\dfrac{\partial \varphi}{\partial t}(t_0, x_1, \ldots, x_n)$ for all spatial locations $(x_1, \ldots, x_n)$, we can predict its values $\varphi(t, x_1, \ldots, x_n)$ at all future moments of time t.
- Similarly, in the new 3rd order equation, once we know the initial values of the scalar function $\varphi(t_0, x_1, \ldots, x_n)$ and its first two time derivatives $\dfrac{\partial \varphi}{\partial t}(t_0, x_1, \ldots, x_n)$ and $\dfrac{\partial^2 \varphi}{\partial t^2}(t_0, x_1, \ldots, x_n)$, we can predict its values $\varphi(t, x_1, \ldots, x_n)$ at all future moments of time t.

The new "equation" does not contain L at all – and is still able to predict how the scalar field changes with time. In this sense, we have "equations without equations":

- on the one hand, we *have* equations – in the sense that prediction is possible, but
- on the other hand, we *do not have* equations – in the sense that no specific Lagrange function is given.

This situation is somewhat similar to Wheeler's cosmological ideas of "mass without mass" and "charge without charge"; see, e.g., (5). As an example of charge without charge, we can consider a space-time with a wormhole. It is possible to set up an electric field in this curved space-time whose lines enter one of the wormhole's mouths and exit at the other one. While there are no charged particles in this space-time model, there is an electric attraction to one of the mouths and repulsion by the other – as if these mouths have opposite charges.

1.5 Scalar fields: proof of the main result

For the Lagrangian $L = L(\varphi, a)$, where $a \stackrel{\text{def}}{=} \varphi_{,i}\varphi^{,i}$, the Euler-Lagrange equations take the form $\dfrac{\partial L}{\partial \varphi} - \partial_i \dfrac{\partial L}{\partial \varphi_{,i}} = 0$. By using the chain rule, we conclude that

$$\frac{\partial L(\varphi, a)}{\partial \varphi_{,i}} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial \varphi_{,i}} = \frac{\partial L}{\partial a} \cdot 2 \cdot \varphi^{,i}. \tag{1.19}$$

Thus, the Euler-Lagrange equations take the form

$$L_{,\varphi} - \partial_i(2 \cdot L_{,a} \cdot \varphi_{,i}) = 0, \tag{1.20}$$

where we denoted $L_{,a} \stackrel{\text{def}}{=} \dfrac{\partial L}{\partial a}$. Using chain rule again, we get

$$L_{,\varphi} - 2 \cdot L_{,a} \cdot \Box\, \varphi - 2 \cdot L_{,a\varphi} \cdot (\varphi_{,i}\varphi^{,i}) - 4 \cdot L_{,aa} \cdot \varphi_{,ij}\varphi^{,i}\varphi^{,j} = 0, \tag{1.21}$$

where $L_{,\varphi} \stackrel{\text{def}}{=} \dfrac{\partial L}{\partial \varphi}$, $\Box\varphi \stackrel{\text{def}}{=} \varphi_{,i}^{,i}$, $L_{,a\varphi} \stackrel{\text{def}}{=} \dfrac{\partial^2 L}{\partial a \partial \varphi}$, and $L_{,aa} \stackrel{\text{def}}{=} \dfrac{\partial^2 L}{\partial a^2}$.

In general, L depends on a, so $L_{,a} \neq 0$. Thus, by moving the term $L_{,a} \cdot \Box\,\varphi$ to the right-hand side and dividing both sides by $2L_{,a}$, we conclude that

$$\Box\,\varphi = \frac{L_{,\varphi}}{2 \cdot L_{,a}} - \frac{L_{,a\varphi}}{L_{,a}} \cdot (\varphi_{,i}\varphi^{,i}) - 2 \cdot \frac{L_{,aa}}{L_{,a}} \cdot \varphi_{,ij}\varphi^{,i}\varphi^{,j}. \tag{1.22}$$

Here, the values $L_{,\varphi}$, $L_{,a}$, $L_{,a\varphi}$, and $L_{,aa}$ at a point x depend only on the values of φ and a at this same point. So,

- if at two points, we have the same values of φ, $a = \varphi_{,i}\varphi^{,i}$, and $b \stackrel{\text{def}}{=} \varphi_{,ij}\varphi^{,i}\varphi^{,j}$,
- then we have same values of $c = \Box\,\varphi$ at these two points.

In particular, this means that if we have dx^k for which $\varphi_{,k}dx^k = 0$, $a_{,k}dx^k = 0$, and $b_{,k}dx^k = 0$, then replacing x by $x + dx$ will not change the values of φ, a, and b, and thus, the value of c will not change, i.e., we should have $c_{,k}dx^k = 0$.

In geometric terms, the condition $\varphi_{,k}dx^k = 0$ means that the vectors $\varphi_{,k}$ and dx^k are orthogonal: $dx^k \perp \varphi_{,k}$. In these terms, the above condition means that for every vector dx^k:

- if $dx^k \perp \varphi_{,k}$, $dx^k \perp a_{,k}$, and $dx^k \perp b_{,k}$,
- then $dx^k \perp c_{,k}$.

It is known that this property is equivalent to $c_{,k}$ being in the 3-plane generated by the vectors $\varphi_{,k}$, $a_{,k}$, $b_{,k}$. In other words, the four vectors $\varphi_{,k}$, $a_{,k}$, $b_{,k}$, and $c_{,k}$ lie in the same 3-plane. In algebraic terms, this means these four vectors are linearly dependent, i.e., that if we form a matrix with these vectors as rows, then its determinant is 0:

$$\varepsilon_{ijkl} \cdot \varphi_{,i} \cdot a_{,j} \cdot b_{,k} \cdot c_{,l} = 0, \tag{1.23}$$

where $\varepsilon_{ijkl} = 0$ if some indices are equal and is ± 1 else (depending on the parity).

The equation (1.23) is the desired third order partial differential equation. Knowing this equation, we can predict future evolution – without knowing the actual Lagrangian L.

1.6 Possible relation to dimension of space-time

Our conclusion is based on the idea that the four vectors $\varphi_{,k}$, $a_{,k}$, $b_{,k}$, and $c_{,k}$ lie in a 3-D plane. If the dimension of space-time is 3 or smaller, this is always true. Thus, the "equations without equations" are only possible when dimension is ≥ 4. Maybe this explains why our space-time is 4-dimensional?

What about the case when we have two scalar fields φ and ψ? Here, we can similarly conclude that the preservation of 11 quantities

$$\varphi \quad \psi$$

$$\begin{array}{ccc} \varphi_{,i}\varphi^{,i} & \psi_{,i}\psi^{,i} & \varphi_{,i}\psi^{,i} \\ \varphi_{,ij}\varphi^{,i}\varphi^{,j} & \varphi_{,ij}\varphi^{,i}\psi^{,j} & \varphi_{,ij}\psi^{,i}\psi^{,j} \\ \psi_{,ij}\varphi^{,i}\varphi^{,j} & \psi_{,ij}\varphi^{,i}\psi^{,j} & \psi_{,ij}\psi^{,i}\psi^{,j} \end{array}$$

means that $\Box\varphi$ and $\Box\psi$ are also preserved. Thus, 12 vectors – gradients of the above quantities and of $\Box\varphi$ – must be in the same 11-D space. This requirement is always true in spaces of dimension ≤ 11. Thus, for two scalar fields, equations without equations are only possible in dim ≥ 12. This inequality is in good accordance with the known fact that a consistent quantum field theory is only possible when dim ≥ 11.

1.7 Conclusion

Not all mathematical solutions to physical equations are physically meaningful: e.g., if we reverse all the molecular velocities in a breaking cup, we get pieces self-assembling into a cup. The resulting initial conditions are "degenerate": once we modify them, self-assembly stops. So, in a physical solution, the initial conditions must be "non-degenerate".

A challenge in formalizing this idea is that it depends on the representation. Example 1: we can use the Schrödinger equation $\mathrm{i}\cdot\hbar\cdot\dfrac{\partial\Psi}{\partial t} = -\dfrac{\hbar^2}{2\cdot m}\cdot\nabla^2\Psi + V(x)\cdot\Psi$ to represent $V(x) = F(\Psi,\ldots)$ as a function of $\Psi(x)$ and its derivatives. The new equation $\dfrac{dF}{dt} = 0$ is equivalent to the Schrödinger equation, but now $V(x)$ is in the initial conditions.

Example 2: for a scalar field φ, we describe a new equation which is satisfied if φ satisfies the Euler-Lagrange equations for some Lagrangian $L(\varphi, \varphi_{,i}\varphi^{,i})$. So, similarly to Wheeler's cosmological "mass without mass", we have "equations without equations".

Thus, when formalizing physical equations:

- we must not only describe them in *a* mathematical form,
- we must also select *one* of the mathematically equivalent forms.

Acknowledgements

This work was supported in part by the National Science Foundation grant HRD-0734825, by Grant 1 T36 GM078000-01 from the National Institutes of Health, and by Grant MSM 6198898701 from MŠMT of Czech Republic.

The authors are thankful to all the participants of the Fall 2008 Meeting of the Texas and Four Corner Sections of the American Physical Society, El Paso, Texas, October 17–18, 2008, for valuable suggestions.

References

Feynman, R., Leighton, R. and Sands, M. (2005). *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts.

Giarratano, J. C. and Riley, G. (2005). *Expert Systems, Principles and Programming*. Course Technology, Boston, Massachusetts.

Kreinovich, V., and Finkelstein, A. M. (2004). Towards applying computational complexity to foundations of physics", *Notes of Mathematical Seminars of St. Petersburg Department of Steklov Institute of Mathematics*. **316**, 63–110; reprinted in *Journal of Mathematical Sciences*, 2006, **134**, 2358–82.

Li, M. and Vitányi, P. (2008.) *An Introduction to Kolmogorov Complexity and Its Applications*. Springer Verlag, New York, 2008.

Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation*, W. H. Freeman, New York.

Appendix: Informal derivation of Euler-Lagrange equations

In general, the Lagrange function has the form $L(f, f_{,i})$, i.e., the form

$$L(f, f_{,0}, \ldots, f_{,n}).$$

For small Δf, we have

$$L(f+\Delta f, f_{,0}+\Delta f_{,0}, \ldots, f_{,n}+\Delta f_{,n}) \approx L(f, f_{,i}) + \sum_{i=0}^{n} \frac{\partial L}{\partial f} \cdot \Delta f + \frac{\partial L}{\partial f_{,i}} \cdot \Delta f_{,i}. \quad (1.24)$$

Thus, we have

$$\int L(f + \Delta f, f_{,i} + \Delta f_{,i})\, dx \approx$$

$$\int L(f, f_{,i})\, dx + \int \frac{\partial L}{\partial f} \cdot \Delta f\, dx + \sum_{i=0}^{n} \int \frac{\partial L}{\partial f_{,i}} \cdot \Delta f_{,i}\, dx, \quad (1.25)$$

i.e.,

$$S(f + \Delta f) \approx S(f) + \int \frac{\partial L}{\partial f} \cdot \Delta f\, dx + \sum_{i=0}^{n} \int \frac{\partial L}{\partial f_{,i}} \cdot \Delta f_{,i}\, dx. \quad (1.26)$$

Each of the integrals $\int \frac{\partial L}{\partial f_{,i}} \cdot \Delta f_{,i}\, dx$ can be computed by using the formula $(A \cdot B)' = A' \cdot B + A \cdot B'$ which implies that

$$\int_{-\infty}^{\infty} A \cdot B'\, dx = -\int_{-\infty}^{\infty} A' \cdot B\, dx + A \cdot B|_{-\infty}^{\infty}. \quad (1.27)$$

In our case, we have $A = \frac{\partial L}{\partial f_{,i}}$ and $B = \Delta f$. Since physical fields usually tend to 0 at infinity, the last term disappears and therefore, we conclude that

$$\int \frac{\partial L}{\partial f_{,i}} \cdot \Delta f_{,i}\, dx = -\int \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial f_{,i}} \right) \cdot \Delta f\, dx. \quad (1.28)$$

Substituting this expression into the formula (1.26), we conclude that

$$S(f+\Delta f) \approx S(f) + \int \frac{\partial L}{\partial f} \cdot \Delta f \, dx - \sum_{i=0}^{n} \int \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial f_{,i}} \right) \cdot \Delta f \, dx, \tag{1.29}$$

i.e., that

$$S(f+\Delta f) \approx S(f) + \int \frac{\delta S}{\delta f} \cdot \Delta f \, dx, \tag{1.30}$$

where

$$\frac{\delta S}{\delta f} = \frac{\partial L}{\partial f} - \sum_{i=0}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial f_{,i}} \right). \tag{1.31}$$

The formula has been derived.

Example. For the Newtonian Lagrange function

$$L(x, \dot{x}) = \frac{1}{2} \cdot m \cdot (\dot{x})^2 - V(x), \tag{1.32}$$

the Euler-Lagrange equation leads to

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = 0, \tag{1.33}$$

i.e., to

$$-\frac{\partial V}{\partial x} - \frac{d}{dt} (m \cdot \dot{x}) = 0, \tag{1.34}$$

or, equivalently, to Newton's equations

$$m \cdot \ddot{x} = -\frac{\partial V}{\partial x}. \tag{1.35}$$