5-2015

# We Live in the Best of Possible Worlds: Leibniz's Insight Helps to Derive Equations of Modern Physics

Vladik Kreinovich
*University of Texas at El Paso*, vladik@utep.edu

Guoqing Liu
*Nanjing University of Technology*, guoqing@njut.edu.cn



Comments:
Technical Report: UTEP-CS-15-38

# We Live in the Best of Possible Worlds: Leibniz's Insight Helps to Derive Equations of Modern Physics

Vladik Kreinovich and Guoqing Liu

**Abstract** To reconcile the notion of a benevolent and powerful God with the actual human suffering, Leibniz proposed the idea idea that while our world is not perfect, it is the best of *possible* worlds. This idea inspired important developments in physics: namely, it turned out that equations of motions and equations which describe the dynamics of physical fields can be deduced from the condition that the (appropriately defined) action functional is optimal. In practice, this idea is not always very helpful in physics applications: to fully utilize this fact, we need to how the action, and there are many possible action functionals. Our idea is to apply Leibniz's insight once again and to assume that (similarly) on the set of all expressions for actions, there is an optimality criterion, and the actual action functional is optimal with respect to this criterion. This idea enables us to derive the standard equations of General Relativity, Quantum Mechanics, Electrodynamics, etc. only from the fact that the corresponding expressions for action are optimal. Thus, the physical equations describing our world are indeed the best possible.

## 1 Introduction

**Leibnitz's idea.** Many religious philosophers have been trying to reconcile the notion of a benevolent and powerful God with the actual human suffering. Leibniz's idea of solving this problem is to conjecture that while our world is not perfect, it is the best of possible worlds; see, e.g., [11].

Vladik Kreinovich
University of Texas at El Paso, El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

Guoqing Liu
Nanjing University of Technology, Nanjing, Jiangsu 210009, P. R. China,
e-mail: guoqing@njut.edu.cn

**Lebniz's idea and physics.** Leibniz's idea of optimality of our world inspired not only interesting philosophical discussions, it also inspired important developments in physics: namely, it turned out that equations of motions and equations which describe the dynamics of physical fields can be deduced from the condition that the action functional

$$S = \int L(x)\,\mathrm{d}x,$$

as determined by the corresponding *Lagrange function* $L(x)$, is optimal; see, e.g., [1, 10].

In other words, there is an optimality criterion on the set of all trajectories, and the actual trajectory is optimal with respect to this criterion.

**Applying Leibniz's idea to physics: the main challenge.** The above application is interesting but not always very very helpful: to find the equations, we need to know the Lagrange function, and there are many possible Lagrange functions. How should we select the most appropriate one?

**Our idea.** Our idea is to apply Leibniz's insight once again and to assume that (similarly) on the set of all Lagrange functions, there is an optimality criterion, and the actual Lagrangian is optimal with respect to this criterion.

**What we get by using this idea: a brief description.** This idea enables us to derive the equations of physics only from the fact that they are optimal. Specifically, under reasonable conditions on the optimality criterion, this approach leads to the standard Lagrange functions for General Relativity, Quantum Mechanics, Electrodynamics, etc.

Thus, the Lagrange functions (and hence equations) of our world are indeed the best possible.

## 2 An Optimality Criterion on the Set of All Lagrange Functions: General Requirements

**What is an optimality criterion.** When we say an *optimality criterion* is defined on the set of all possible Lagrange functions, we mean that on the set of all such functions, there must be a relation $\geq$ describing which Lagrange function is better or equal in quality.

This relation must be transitive (if $L$ is better than $L'$, and $L'$ is better than $L''$, then $L$ is better than $L''$). This relation is not necessarily asymmetric, because we can have two Lagrange functions of the same quality.

**Definition 1.** Let $\mathscr{A}$ be a set; elements of this set will be called *alternatives*. By an *optimality criterion*, we mean a transitive relation $\geq$ on the set $\mathscr{A}$.

**Optimality criterion must be final.** We would like to require that this relation be *final* in the sense that it should define a unique *best* Lagrange function $L_{\text{opt}}$ (i.e., the unique Lagrange function for which $\forall L\,(L_{\text{opt}} \geq L)$. Indeed:

- If none of the Lagrange functions is the best, then this optimality criterion is of no use, so there should be *at least one* optimal family.
- If *several* different Lagrange functions are equally best, that means that this optimality criterion is not sufficient to determine the actual Lagrange function: we must still select between the several "best" ones. As a result, the original optimality criterion was not final: we get a new criterion ($L \geq_{\text{new}} L'$ if either $L \geq_{\text{old}} L'$ in the sense of the old criterion, or if $L \sim_{\text{old}} L'$ and $L$ is better according to some additional criterion), for which the class of optimal Lagrange functions is narrower. We can repeat this procedure until we get a final criterion for which there is only one optimal Lagrange function.

**Definition 2.** We say that an optimality criterion $\geq$ on a set $\mathscr{A}$ is *final* if there exists one and only one *optimal* alternative, i.e., an alternative $a_{\text{opt}}$ for which $\forall a\,(a_{\text{opt}} \geq a)$.

**Optimality criterion must be scale-invariant.** It is reasonable to require that the relation $L \geq L'$ should not change if we simply change the units in which we measure length, i.e., if we change the length *scale*.

In other words, we want the optimality criterion to be *scale-invariant*.

**Definition 3.** Let $G$ be a group of transformations from $\mathscr{A}$ to $\mathscr{A}$. We say that a criterion $\geq$ is *G-invariant* if for every two alternatives $a$ and $a'$, and for every transformation $g \in G$, $a \geq a'$ implies $g(a) \geq g(a')$.

*Comments.*

- Symmetry ideas are known to be very useful in physics and in foundations of physics (see, e.g., [2, 3, 4, 5, 6, 7, 8, 9, 13, 14, 15]), so it is reasonable to apply these ideas to our problem as well.
- In particular, when the optimality criterion $\geq$ is $G$-invariant with respect to a transformation group $G$ describing *scalings*, we will say that $\geq$ is *scale-invariant*

Let us describe how these requirements apply to different fundamental physical fields.

## 3 First Case: Optimal Lagrange Function for Gravitation

**Lagrange function for gravitation: general definition.** A physical field which describes gravitation is the metric field $g_{ij}(x)$. So, a general Lagrange function for gravitation can depend on the values of this field and of its derivatives of all orders. In the absence of the gravitation, $g_{ij} = \eta_{ij} = \text{diag}(1,-1,-1,-1)$, and all partial derivatives of metric are equal to 0: $g_{ij,k} \stackrel{\text{def}}{=} \dfrac{\partial g_{ij}}{\partial x_k} = 0$, $g_{ij,kl} = 0$, etc. It is therefore reasonable to require that the Lagrange function $L$ be *analytical* in terms of the differences between the actual values $g_{ij}(x)$, $g_{ij,k}$, ... of the field and of its derivatives, and the values $\eta_{ij}, 0, \ldots$ corresponding the *null-field* (absence of gravitation). In other words, it is reasonable to require that $L$ is an analytical function of $g_{ij} - \eta_{ij}$, $g_{ij,k}$, $g_{ij,kl}$, ... Thus, we arrive at the following definition:

**Definition 4.** By a *gravitational Lagrange function L*, we mean a generally covariant analytical function of the differences $g_{ij}(x) - \eta_{ij}$ and of the derivatives $g_{ij,k}(x)$, ..., $g_{ij,k\ldots l}(x)$ in the same point $x$:

$$L(x) = L(g_{ij}(x) - \eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x), \ldots). \tag{1}$$

**What does scale invariance mean for gravitation?** When we say that a Lagrange function must be generally covariant, we mean that its value should not depend on the choice of coordinate system (i.e., it should not change if we change a coordinate system). This guarantees that the resulting *action* $S = \int L \cdot \sqrt{-g}\,\mathrm{d}^4x$ will also be generally covariant, and so the resulting field equations will be generally covariant.

In addition to changing coordinates, we can also change the unit of length. From the physical viewpoint, if we change a unit of length, the physical space-time will not change. However, from the mathematical viewpoint, the space changes: if we change the unit of length to a unit which is $\lambda$ times smaller, then the numerical value of the length

$$ds = \sqrt{\sum g_{ij} \cdot dx^i \cdot dx^j} \tag{2}$$

will change to $ds' = \lambda \cdot ds$, i.e., we will get $ds' = \sqrt{\sum g'_{ij} \cdot dx^i \cdot dx^j}$ with a new metric field

$$g'_{ij} = \lambda^2 \cdot g_{ij}. \tag{3}$$

How can we best describe this transformation in physical terms? From the purely mathematical viewpoint, we can simply keep the same coordinate system $x^i$; then the corresponding scaling transformation can be simply described as a transformation (3) for the metric tensor and, correspondingly, a similar transformation

$$g'_{ij,k\ldots l} = \lambda^2 \cdot g_{ij,k\ldots l} \tag{4}$$

for its derivatives.

However, from the physical viewpoint, this description (3), (4) would be rather unnatural, because coordinates are usually assigned based on distances, and therefore, if we change the unit for length, the coordinates should also change accordingly: from $x^i$ to

$$x'^i = \lambda \cdot x^i. \tag{5}$$

In this case, if we change both the metric $ds$ to $ds' = \lambda \cdot ds$ and coordinates from $x^i$ to $x'^i = \lambda \cdot x^i$, then, from (2), we can conclude that $ds' = \sqrt{\sum g_{ij} \cdot dx'^i \cdot dx'^j}$, i.e., that the metric does not change:

$$g'_{ij} = g_{ij}. \tag{6}$$

Correspondingly, due to (5) and (6), the derivatives of the metric get transformed as

$$g'_{ij,k} = \lambda^{-1} \cdot g_{ij,k}; \tag{6a}$$

$$g'_{ij,kl} = \lambda^{-2} \cdot g_{ij,kl}; \tag{6b}$$

etc.

How does the Lagrange function change under this transformation? From the physical viewpoint, action $S = \int L \cdot \sqrt{-g}\, \mathrm{d}^4x$ is energy$\times$time. We are considering a relativistic theory, and moreover, we are following the tradition of gravitation theory in using the units in which distance and time are measured by the same unit, i.e., in which the speed of light $c$ is equal to 1 (and so, $h_{ij} = \mathrm{diag}(1, -1, -1, -1)$). In such units, energy $E = m \cdot c^2$ is described in the same units as mass, and time in the same units as distance, so action changes as mass$\times$distance.

If we change a unit of length, how will the corresponding unit of mass change? To describe this change, it is sufficient to look at the known approximate gravitational theory: Newtonian gravitation. In Newtonian gravitation, the force $F = m \cdot a$ with which a body of mass $M$ attracts a body of mass $m$ is proportional to

$$m \cdot a = G \cdot \frac{m \cdot M}{r^2},$$

hence

$$a = \frac{G \cdot M}{r^2}. \tag{7}$$

When we change a unit of length (and the corresponding unit of time), we get $r' = \lambda \cdot r$, $t' = \lambda \cdot t$, $a' = r'/(t')^2 = \lambda^{-1} \cdot a$ and therefore, to preserve the above relation (7), we must have $M' = \lambda \cdot M$.

So, the mass (hence, the energy) transforms as $M \to \lambda \cdot M$; we already know that time $t$ transforms as $t \to t' = \lambda \cdot t$. Hence, the action (energy$\times$time) transforms as $S \to S' = \lambda^2 \cdot S$, and therefore, the Lagrange function $L$, which is defined as the density of the action, i.e., as $L \sim S/r^4$, is transformed as

$$L \to L' = (\lambda^2/\lambda^4) \cdot L = \lambda^{-2} \cdot L.$$

Hence, after scaling, the old Lagrange function (1) transforms, in the new units, into the expression

$$L'(x) = \lambda^{-2} \cdot L(g_{ij}(x) - \eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x), \ldots).$$

This expression describes $L'$ as a function of values $g_{ij}$, $g_{ij,k}$, $g_{ij,kl}$, $\ldots$, expressed in the old units. We want to get an expression of $L'$ in terms of $g'_{ij}$, $g'_{ij,k}$, $g'_{ij,kl}$, $\ldots$, i.e., in terms of the field values and derivatives expressed in new units. From (6), (6a), (6b), etc., we can conclude that $g'_{ij} = g_{ij}$, $g'_{ij,k} = \lambda \cdot g_{ij,k}$, $g'_{ij,kl} = \lambda^2 \cdot g_{ij,kl}$, etc. Therefore, in the new unit, the Lagrange function is expressed as:

$$L' = g_\lambda(L) = \lambda^{-2} \cdot L(g_{ij} - \eta_{ij}, \lambda \cdot g_{ij,k}, \lambda^2 \cdot g_{ij,kl}, \ldots). \tag{8}$$

So, for gravitational Lagrange functions, scale transformation means going from $L$ to $L' = g_\lambda(L)$, and scale-invariance means invariance with respect to such transformations.

**Main result for the gravity field.** Now, we are ready to describe our main result for gravitation.

**Theorem 1.** *For every scale-invariant final optimal criterion on the set of all gravitational Lagrange functions, the optimal Lagrange function has the form $L = b \cdot R$, where $b$ is a constant, and $R$ is the scalar curvature.*

In other words, for any reasonable optimality criterion, General Relativity is the best of all possible Lagrange functions. To be more precise, of all Lagrange functions in which gravitation is described by a single field: metric tensor field $g_{ij}$; alternative gravitation theories are described in Sections 6 and 7.

**Proof.** This proof is similar to the proofs from [13]; the second part is similar to the proofs from [6, 7].

$1^\circ$. Let us first show that the optimal Lagrange function $L_{\rm opt}$ is itself scale-invariant, i.e., that for every $\lambda > 0$, $g_\lambda(L_{\rm opt}) = L_{\rm opt}$.

Indeed, let $\lambda > 0$ be an arbitrary positive number. Since $L_{\rm opt}$ is optimal, for every other Lagrange function $L$, we have $L_{\rm opt} \geq g_{1/\lambda}(L)$. Since the optimality criterion $\geq$ is invariant, we conclude that $g_\lambda(L_{\rm opt}) \geq g_\lambda(g_{1/\lambda}(L)) = L$. Since this is true for every Lagrange function $L$, the Lagrange function $g_\lambda(L_{\rm opt})$ is also optimal. But since our criterion is final, there is only one optimal Lagrange function and therefore, $g_\lambda(L_{\rm opt}) = L_{\rm opt}$. In other words, the optimal Lagrange function is indeed invariant.

$2^\circ$. Let us now show that $L = b \cdot R$.

From Part 1 of this proof, we conclude that $g_\lambda(L) = L$, i.e., that

$$\lambda^{-2} \cdot L(g_{ij} - \eta_{ij}, \lambda \cdot g_{ij,k}, \lambda^2 \cdot g_{ij,kl}, \ldots) = L(g_{ij} - \eta_{ij}, g_{ij,k}, g_{ij,kl}, \ldots). \quad (9)$$

Let us consider an arbitrary point $A$ and normal coordinates in it (see, e.g., [12]). It is known that in some neighborhood of $A$, $g_{ij}(B) = h_{ij} +$ some analytical function of $B^i - A^i$ with coefficients which polynomially depend on curvature tensor $R_{ijkl}(A)$ and its covariant derivatives of arbitrary orders. Therefore $g_{ij}(A)$ and every derivative $g_{ij,i_1 \ldots i_p}(A)$ are also such polynomial functions. If we substitute these expressions into $L$, then $L$ will become an analytical function of the curvature tensor $R_{ijkl}$ and of its covariant derivatives $R_{ijkl;m}, R_{ijkl;mn}, \ldots$, i.e. a sum of infinitely many monomials of the variables $R_{ijkl;m}, R_{ijkl;mn}, \ldots$ :

$$L = L(g_{ij} - \eta_{ij}, R_{ijkl}, R_{ijkl;m}, R_{ijkl;mn}, \ldots).$$

Let us express (9) in terms of these new variables. With respect to scale transformations,

$$R_{ijkl} \to R'_{ijkl} = \lambda^{-2} \cdot R_{ijkl}, \text{ and}$$

$$R_{ijkl;i_1\ldots i_p} \to R'_{ijkl;i_1\ldots i_p} = \lambda^{-(2+p)} \cdot R_{ijkl;i_1\ldots i_p}.$$

Therefore,

$$R_{ijkl} = \lambda^2 \cdot R'_{ijkl},$$

$$R_{ijkl;i_1\ldots i_p} = \lambda^{2+p} \cdot R'_{ijkl;i_1\ldots i_p},$$

and (9) turns into

$$\lambda^{-1} \cdot L(g_{ij} - \eta_{ij}, \lambda^2 \cdot R_{ijkl}, \lambda^3 \cdot R_{ijkl;m}, \lambda^4 \cdot R_{ijkl;mn}, \ldots) =$$
$$L(g_{ij} - \eta_{ij}, R_{ijkl}, R_{ijkl;m}, R_{ijkl;mn}, \ldots). \tag{10}$$

Expressions on both sides of (10) are sums of similar monomials. Since the two analytical functions coincide, this means that all the coefficients at the corresponding monomials must coincide.

Each monomial in the right-hand side does not depend on $\lambda$; the corresponding monomial in the left-hand side of (10) is multiplied by $\lambda^{2n_R+n_D-2}$, where $n_R$ is a total number of all curvature tensors and their covariant derivatives in this monomial, and $n_D$ is a total number of all differentiation indices in it. Since the coefficients must coincide, we conclude that the function $L$ can only have monomials with

$$2n_R + n_D = 2.$$

Both numbers $n_R$ and $n_D$ are non-negative integers, so there are only two possibilities for $2n_R + n_D = 2$: when $n_R = 1$ and $n_D = 0$, and when $n_R = 0$ and $n_D = 2$.

In the first case, $L$ is a linear function of $R_{ijkl}$, so, since $L$ is generally covariant, we have $L = b \cdot R$.

In the second case, there is no curvature tensor in $L$ , and covariant differentiation is applied only to $g_{ij}$, therefore the result is zero ($g_{ij;kl} = 0$).

So, $L = b \cdot R$. The theorem is proven.

*Comments.*

1. We have shown that *if* a Lagrange function is optimal with respect to some scale-invariant final optimality criterion, then it is $L = b \cdot R$, but we have not yet proven the existence of such criteria. The following simple example proves this existence: we can define an optimality criterion according to which $R > L$ for any $L \neq R$, and $L \sim L'$ for every two $L, L' \neq R$. This criterion is clearly scale-invariant and final.
2. The requirements that the Lagrange function $L$ is analytical and that the optimality criterion is scale-invariant are both essential:
   - If we do not require that $L$ is analytical, then we can have the Lagrange function $L_0 = \sqrt{R_{ij} \cdot R^{ij}}$, and an optimality criterion according to which $L_0 > L$ for any $L \neq L_0$, and $L \sim L'$ for every two $L, L' \neq L_0$. This criterion is scale-invariant and final, and the corresponding optimal Lagrange function is $L_0 \neq b \cdot R$.

- If we do not require that the optimality criterion is scale-invariant, then we can take a Lagrange function $L_1 = R + R^2$, and an optimality criterion according to which $L_1 > L$ for any $L \neq L_1$, and $L \sim L'$ for every two $L, L' \neq L_1$. This criterion is final, and the corresponding optimal Lagrange function is $L_0 \neq b \cdot R$.

**Fundamentality principle.** In the previous text, we used transformational properties of $L$ with respect to scaling, which were deduced from physical arguments. In the present section we show that we can eliminate these arguments, if we use the following *fundamentality principle*:

A phenomenon is called *fundamental* if it can be explained without using other phenomena. In our case, it means that transformation law for $L$ must be chosen in such a way that field equations are uniquely determined by optimality requirement, i.e. $L$ must be determined uniquely modulo multiplicative constant.

To be more precise, we define scale transformations as

$$L' = g_\lambda(L) = \lambda^{-d} \cdot L(g_{ij} - \eta_{ij}, \lambda \cdot g_{ij,k}, \lambda^2 \cdot g_{ij,kl}, \ldots). \tag{8a}$$

for some unspecified value $d$.

**Proposition 1.** *The only value $d$ for which all Lagrange functions which are optimal with respect to scale-invariant final optimal criteria lead to the same dynamical equations is $d = 2$.*

**Proof.** Similarly to Theorem 1, we conclude that for every $d$, the optimal Lagrange function is a sum of terms for which $2n_R + n_D = d$. The value $d$ is a sum of two non-negative integers, so $d \geq 0$.

The Lagrange function is a scalar, so the total number of indices in every term $P$ is even, hence $n_D$ is even. So, $d$ must also be even.

If $d = 0$, then $n_R = n_D = 0$, hence $L = \text{const}$, and there are no variational equations at all.

If $d \geq 4$, then we can take terms $L = (R_{ij}R^{ij} + bR_{ijkl}R^{ijkl} + cR^2) \cdot R^{(d-4)/2}$. For different $b$ and $c$, these Lagrange functions lead to different variational equations, and each of these function $L_0$ is optimal with respect to some scale-invariant final optimality criterion: namely, a criterion in which $L_0 > L$ for all $L \neq L_0$, and $L \sim L'$ for all $L, L' \neq L_0$.

Thus, only for $d = 2$, we get the desired uniqueness. The proposition is proven.

## 4 Second Case: Electromagnetic Field (in Curved Space)

**Lagrange function for electromagnetic field: general definition.** Electromagnetic field is described by a vector potential $A_i(x)$; its source is the 4-current $j^i$ which satisfies the charge conservation law $j^i_{;i} = 0$. In classical electrodynamics, the vector potential does not have a direct physical meaning, only $F_{ij} = A_{i,j} - A_{j,i}$; therefore, it is normally assumed that the Lagrange function should be invariant under *gauge transformations* $A_i \to A_i - f_{,i}$ which preserve $F_{ij}$ for an arbitrary function $f(x)$.

**Definition 5.** By a *Lagrange function for electromagnetic field* $L$, we mean a generally covariant analytical function of the differences $g_{ij}(x) - \eta_{ij}$, of $A_i(x)$, $j^i(x)$, and of the derivatives $g_{ij,k}(x), g_{ij,kl}(x), \ldots$, $A_{i,k}(x), A_{i,kl}(x), \ldots$, $j^i_{,k}, j^i_{,kl}(x), \ldots$, in the same point $x$:

$$L(x) = L(g_{ij}(x) - \eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x), \ldots,$$

$$A_i(x), A_{i,k}(x), A_{i,kl}(x), \ldots, j^i(x), j^i_{,k}(x), j^i_{,kl}(x), \ldots),$$

for which the variational equations are gauge-invariant.

**What does scale invariance mean for electromagnetic field?** In the Newtonian approximation, the force $F = q \cdot Q/r^2$ between the two charges is described by the same formula as the (gravitational) force between the two masses; therefore, if we want to preserve this approximation, then when we change the unit of length, we must transform charges in exactly the same way as masses, i.e., as $q \to q' = \lambda \cdot q$. Thus, the 4-current $j^i$ (charge/length$^3$) should transform as $j^i \to j'^i = \lambda^{-2} \cdot j^i$. In Newtonian approximation, the electromagnetic potential is $Q/r$, so the potential $A_i$ should not change under scale transformations. So, the expression

$$L'(x) = \lambda^{-2} \cdot L(g_{ij}(x) - \eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x), \ldots,$$

$$A_i(x), A_{i,k}(x), A_{i,kl}(x), \ldots, j^i(x), j^i_{,k}(x), j^i_{,kl}(x), \ldots).$$

leads to the following scale transformation:

$$L' = g_\lambda(L) = \lambda^{-2} \cdot L(g_{ij} - \eta_{ij}, \lambda \cdot g_{ij,k}, \lambda^2 \cdot g_{ij,kl}, \ldots,$$

$$A_i, A_{i,k}, A_{i,kl}, \ldots, \lambda^2 \cdot j^i, \lambda^3 \cdot j^i_{,k}, \lambda^4 \cdot j^i_{,kl}, \ldots).$$

**Main result for electromagnetic field.**

**Theorem 2.** *For every scale-invariant final optimal criterion on the set of all Lagrange functions for electromagnetic field, the optimal Lagrange function has the form* $L = b \cdot R + c \cdot F_{ij} \cdot F^{ij} + d \cdot A_i \cdot j^i$ *for some constants* $b$, $c$, *and* $d$.

Thus, the Lagrange function corresponding to standard Maxwell's equations is indeed optimal.

**Proof.** Similarly to the proof of Theorem 1, we conclude that the optimal Lagrange function is scale-invariant, that it is an analytical function of the metric field, of curvature, of vector potential, of 4-current, and of their covariant derivatives, and, therefore, that it can only contain monomials which do not depend on $\lambda$. On the other hand, each monomial is proportional to $\lambda^{2n_R + n_D + 2n_J - 2}$, where $n_R$ and $n_D$ are defined as in the proof of Theorem 1, and $n_J$ is the total number of currents and its derivatives in this monomial. Thus, we must have $2 = 2n_R + n_D + 2n_J$. Since all three numbers $n_R$, $n_D$, and $n_J$ are non-negative integers, we have three possibilities:

- $n_R = 1$, $n_D = n_J = 0$;
- $n_R = n_D = 0$, $n_J = 1$;
- $n_R = n_J = 0$, $n_D = 2$.

In the first case, $L$ contains either $R$, or the product of $R_{ijkl}$ and terms $A_i$; this product leads to the terms in variational equations which are not gauge invariant, so it cannot be in $L$.

In the second case, due to the fact that $L$ is a scalar, the total number of indices of all tensors (whose product constitutes the monomial) must be even; therefore the total number $n_A$ of potentials and its derivatives in this monomial must be odd. If $n_A = 1$, the only possibility is $P = d \cdot j_i \cdot A^i$. If $n_A \geq 3$, the result of varying is not gauge invariant.

In the third case, $n_A$ must also be even. If $n_A = 0$, then $P = g_{ij;kl} = 0$. If $n_A = 2$, then gauge invariance leads to $P = c \cdot F_{ij} \cdot F^{ij}$, and if $n_A \geq 4$, the result of varying is not gauge invariant. The theorem is proven.

## 5 Third Case: (Non-Relativistic) Quantum Mechanics

**Lagrange function for non-relativistic quantum mechanics: general definition.** We want to obtain a Lagrange function describing the dynamics of a particle of mass $m$, described by a (complex-valued) wave function $\psi(x,t)$, in a field with a potential energy function $V(x,t)$. Since the Lagrange function must be real-valued, it can also depend on the complex conjugate values $\psi^*(x,t)$.

This Lagrange function should be rotation-invariant. There is one more invariance specific for non-relativistic quantum mechanics. Namely, it is known that in quantum mechanics, we can add a constant phase to all the values of $\psi(x,t)$ without changing the physical meaning. Thus, the Lagrange function should be *phase-invariant*, i.e., invariant with respect to the transformation

$$\psi(x,t) \to \exp(\mathrm{i} \cdot \alpha) \cdot \psi(x,t)$$

for any real constant $\alpha$.

**Definition 6.** By a *Lagrange function for non-relativistic quantum mechanics $L$*, we mean a phase-invariant rotation-invariant real-valued analytical function of the mass $m$, its inverse $m^{-1}$, fields $\psi(x,t)$, $\psi^*(x,t)$, and $V(x,t)$, and their derivatives of arbitrary orders with respect to time and spatial coordinates:

$$L(m, m^{-1}, \psi(x,t), \psi_{,k}(x,t), \dot{\psi}(x,t), \ldots, \psi^*(x,t), \psi^*_{,k}(x,t), \dot{\psi}^*(x,t), \ldots,$$

$$V(x,t), V_{,k}(x,t), \dot{V}(x,t), \ldots)$$

**What does scale invariance mean for non-relativistic quantum mechanics?** In (relativistic) gravitation, there is a direct connection between units of space and time.

In non-relativistic case, there is no such direct connection, so we can independently change the unit for space $x^i \to x'^i = \lambda \cdot x^i$ and a unit of time $t \to t' = \mu \cdot t$. It is reasonable to require that the optimality criterion on the set of all Lagrange functions for non-relativistic quantum mechanics be invariant with respect to both scaling transformations.

How do $L$, $\psi(x,t)$, and $V(x,t)$ change under these transformations? A specific feature of quantum measurements is that simple experiments enable us to obtain a unit of action $\hbar$; therefore action $S = \int L(x,t)\,\mathrm{d}^3x\,\mathrm{d}t$ must be invariant with respect to scale transformations. Hence, $L(x,t)$ (which is action/(volume$\times$time)) must transform as $L \to L' = \lambda^{-3} \cdot \mu^{-1} \cdot L$.

Similarly, since action is energy$\times$time, and action is invariant, the potential energy $V(x,t)$ must transform as $V \to V' = \mu^{-1} \cdot V$.

Energy is mass$\times$velocity$^2$; we know how energy is transformed and how velocity is transformed; therefore, for mass, we get $m \to m' = \lambda^{-2} \cdot \mu \cdot m$.

The transformation law for the wave function $\psi(x,t)$ can be deduced from its physical meaning: the integral $\int |\psi|^2\,\mathrm{d}V$ is a probability and is therefore independent (invariant) on the choice of length or time units, i.e. invariant. So, $|\psi|^2 \sim 1/\text{length}^3$, hence, $|\psi|^2 \to \lambda^{-3} \cdot |\psi|^2$, and $\psi \to \psi' = \lambda^{-3/2} \cdot \psi$.

Therefore, the expression

$$L'(x,t) = \lambda^{-3} \cdot \mu^{-1} \cdot L(m, m^{-1}, \psi(x,t), \psi_{,k}(x,t), \dot{\psi}(x,t), \ldots,$$
$$\psi^*(x,t), \psi^*_{,k}(x,t), \dot{\psi}^*(x,t), \ldots, V(x,t), V_{,k}(x,t), \dot{V}(x,t), \ldots)$$

leads to

$$L' = g_{\lambda,\mu}(L) = \lambda^{-3} \cdot \mu^{-1} \cdot L(\lambda^2 \cdot \mu^{-1} \cdot m, \lambda^{-2} \cdot \mu \cdot m^{-1},$$
$$\lambda^{3/2} \cdot \psi, \lambda^{5/2} \cdot \psi_{,k}, \lambda^{3/2} \cdot \mu \cdot \dot{\psi}, \ldots, \lambda^{3/2} \cdot \psi^*, \lambda^{5/2} \cdot \psi^*_{,k}, \lambda^{3/2} \cdot \mu \cdot \dot{\psi}^*, \ldots,$$
$$\mu \cdot V, \lambda \cdot \mu \cdot V_{,k}, \mu^2 \cdot \dot{V}, \ldots).$$

**Main result for non-relativistic quantum mechanics.** Now, we are ready to present the main result of this section.

**Theorem 3.** *For every scale-invariant final optimal criterion on the set of all Lagrange functions for non-relativistic quantum mechanics, the optimal Lagrange function has the form*

$$L = \mathrm{i} \cdot b \cdot \left( \psi \cdot \frac{\partial \psi^*}{\partial t} - \psi^* \cdot \frac{\partial \psi}{\partial t} \right) + \frac{c}{m} \cdot (\nabla \psi \cdot \nabla \psi^*) + d \cdot V \cdot \psi \cdot \psi^* + L_0,$$

*where b, c, and d are real constants, and $L_0$ is an expression which does not contribute to variational equations.*

This Lagrange function leads to Schrödinger equation which is, thus, optimal.

**Proof.** Let us first fix $m$ and consider only transformations which preserve $m$, i.e., transformations for which $\mu = \lambda^2$. For these transformations,

$$L' = g_\lambda(L) = \lambda^{-5} \cdot L(\lambda^{3/2} \cdot \psi, \lambda^{5/2} \cdot \psi_{,k}, \lambda^{7/2} \cdot \dot{\psi}, \ldots,$$

$$\lambda^{3/2} \cdot \psi^*, \lambda^{5/2} \cdot \psi^*_{,k}, \lambda^{7/2} \cdot \dot{\psi}^*, \ldots, \lambda^2 \cdot V, \lambda^3 \cdot V_{,k}, \lambda^4 \cdot \dot{V}, \ldots).$$

Similarly to the proof of Theorem 1, we conclude that the optimal Lagrange function is scale-invariant, and therefore, that it can only contain monomials which do not depend on $\lambda$. On the other hand, each monomial is proportional to $\lambda^{(3/2)\cdot n_\psi + 2n_V + n_S + 2n_T - 5}$, where $n_\psi$ is the total number of terms $\psi$, $\psi^*$, and their derivatives, $n_V$ is the total number of $V$ and its derivatives, $n_S$ is the total number of spatial differentiations, and $d_T$ is the total number of differentiations with respect to time. Thus, we must have $(3/2) \cdot n_\psi + 2n_V + n_S + 2n_T = 5$. Since all four numbers $n_\psi$, $n_V$, $n_S$, and $n_T$ are integers, we must have $n_\psi$ even. Since all these integers are non-negative, we have the following options:

- $n_\psi = 2$, $n_V = 1$, $n_S = n_T = 0$;
- $n_\psi = 2$, $n_V = 0$, $n_S = 2$, $n_T = 0$;
- $n_\psi = 2$, $n_V = 0$, $n_S = 0$, $n_T = 1$;
- $n_\psi = 0$ and $2n_V + n_S + 2n_T = 5$.

In the first case, we get a product of $V$ and two terms of type $\psi$ and $\psi^*$; the only way to make it real-valued is to have $V \cdot \psi \cdot \psi^*$. Another possibility would be

$$V \cdot (\psi^2 + (\psi^*)^2),$$

but the corresponding variational equations are not phase-invariant.

In the second case, we have two derivatives of two functions $\psi$. Due to the requirement that $L$ is real-valued, one of them must be $\psi$, and another one $\psi^*$. Due to rotation-invariance, we have two possibilities: $\psi_{,i} \cdot \psi^*_i$ and $\psi \cdot \Delta \psi^*$; the second term differs from the first one by a full derivative, so we can assume that we get the first term, and add the full derivative to $L_0$.

In the third case, we have two functions $\psi$ and $\psi^*$ and one time derivative. This leads to the corresponding term in $L$.

In the fourth case, the monomial does not depend on $\psi$ at all, so it does not contribute to the variational equations at all; so all terms of these type go directly to $L_0$.

We have almost proved the theorem, except for the dependence on $m$. To do that, we can take the expression that we have obtained so far, substitute the dependence on $m$, and explicitly require that the result be invariant with respect to all scaling transformation. This will enable us to find the exact dependence on $m$. The theorem is proven.

*Comments.*

1. If in the formulation of Theorem 3, we allow $L$ to depend also on the cosmological field $\Lambda$ and on its derivatives, then we'll obtain the Lagrange function which

can be obtained from that of Theorem 3 by a nonessential change $V \to V + \text{const}$. This result implies that *the cosmological lambda term does not influence non-relativistic effects.*

2. The wave function $\psi$ is not directly observable. Therefore, it may seem natural, instead of using $\psi(x,t)$, to use a directly observable probability density $\rho(x,t)$. We can repeat the same arguments as above and try to get a Lagrange function depending on $m$, $m^{-1}$, $\rho(x,t)$, $V(x,t)$, and their derivatives of different orders that is optimal with respect to some scale-invariant optimality criterion. A similar proof can describe the corresponding Lagrange functions; it turns out that they do not lead to any dynamics at all, because the only possible term containing time derivative is $\dot{\rho}$, which is a full derivative. Therefore, *our approach explains why we cannot restrict ourselves to directly observable quantities in the formulation of quantum mechanics*.

## 6 First Auxiliary Result: Gravitation With a $\Lambda$-Term

**Lagrange function for gravitation with a $\Lambda$-term: general definition.** Gravitation theory with a $\Lambda$-term is not invariant with respect to scale transformations, because it contains a fixed unit of length $\sqrt{\Lambda^{-1}}$. But if we consider $\Lambda$ not as a constant, but as a new field, transforming according to the law

$$\Lambda \to \Lambda' = \lambda^{-2} \cdot \Lambda, \tag{11}$$

then we get a possibly scale-invariant situation. So, we arrive at the following definition:

**Definition 7.** By a *Lagrange function for gravitation with a $\Lambda$-term $L$*, we mean a generally covariant analytical function of the differences $g_{ij}(x) - \eta_{ij}$, field $\Lambda(x)$, and of the derivatives $g_{ij,k}(x), \ldots, g_{ij,k\ldots l}(x), \ldots, \Lambda_{,k}(x), \ldots, \Lambda_{,k\ldots l}(x), \ldots$ at the same point $x$:

$$L(x) = L(g_{ij}(x) - \eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x), \ldots, \Lambda(x), \Lambda_{,k}(x), \Lambda_{,kl}(x), \ldots).$$

**What does scale invariance mean for gravitation with a $\Lambda$-term?** Under scale transformations, the new field $\Lambda$ gets transformed according to the formula (11). Therefore, $\Lambda = \lambda^2 \cdot \Lambda'$, and hence, the expression

$$L'(x) = \lambda^{-2} \cdot L(g_{ij}(x) - \eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x), \ldots, \Lambda(x), \Lambda_{,k}(x), \Lambda_{,kl}(x), \ldots)$$

leads to

$$L' = g_\lambda(L) =$$
$$\lambda^{-2} \cdot L(g_{ij} - \eta_{ij}, \lambda \cdot g_{ij,k}, \lambda^2 \cdot g_{ij,kl}, \ldots, \lambda^2 \cdot \Lambda, \lambda^3 \cdot \Lambda_{,k}, \lambda^4 \cdot \Lambda_{,kl}, \ldots). \tag{12}$$

**Main result for gravitation with a $\Lambda$-term.** Now, we are ready for the main result.

**Theorem 4.** *For every scale-invariant final optimal criterion on the set of all Lagrange functions for gravitation with a $\Lambda$-term, the optimal Lagrange function has the form $L = b \cdot R + a \cdot \Lambda$ for some constants $a$ and $b$.*

If we rename $\Lambda' = (a/b) \cdot \Lambda$, we get the standard Einstein's theory $L = b(R + \Lambda')$, which is, thus, optimal.

**Proof.** Similarly to the proof of Theorem 1, we conclude that the optimal Lagrange function is scale-invariant, that it is an analytical function of the metric field, of the curvature, of the field $\Lambda$, and of their covariant derivatives, and, therefore, that it can only contain monomials which do not depend on $\lambda$. On the other hand, each monomial is proportional to $\lambda^{2n_R+n_D+2n_\Lambda-2}$, where $n_R$ and $n_D$ are defined as in the proof of Theorem 1, and $n_\Lambda$ is the total number of $\Lambda$ and its derivatives in this monomial. Thus, we must have $2 = 2n_R + n_D + 2n_\Lambda$. Since all three numbers $n_R$, $n_D$, and $n_\Lambda$ are non-negative integers, we have either $n_\Lambda = 0$ (then $P = b \cdot R$), or $n_\Lambda = 1$, in which case $n_R = n_D = 0$ and $P = a \cdot \Lambda$. The theorem is proven.

## 7 Second Auxiliary Result: Scalar-Tensor Gravitation

**Lagrange function for scalar-tensor gravitation: general definition.** The main idea of the scalar-tensor theory is that the gravitational constant $G$ which relates the gravitational force $F$ to masses ($F = G \cdot m \cdot M/r^2$) is not necessarily a constant, it may change with time, i.e., in other words, it represent a new physical field. Traditionally, the inverse value $\varphi = 1/G$ is used in such theories; to make a comparison with the existing theories easier, we will use this requirement.

Since $\varphi$ is not necessarily a small number, we can assume that the Lagrange function is analytically depending not only on $\varphi$, and on the derivatives of $\varphi$, but also on $\varphi^{-1}$. So, we arrive at the following definition:

**Definition 8.** By a *Lagrange function for scalar-tensor gravitation $L$*, we mean a generally covariant analytical function of the differences $g_{ij}(x) - \eta_{ij}$, field $\varphi(x)$, its inverse $\varphi^{-1}(x)$, and of the derivatives $g_{ij,k}(x)$, $g_{ij,kl}(x)$, $\ldots$, $\varphi_{,k}(x)$, $\varphi_{,kl}(x)$, $\ldots$, in the same point $x$:

$$L(x) = L(g_{ij}(x) - \eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x), \ldots, \varphi(x), \varphi^{-1}(x), \varphi_{,k}(x), \varphi_{,kl}(x), \ldots).$$

**What does scale invariance mean for scalar-tensor gravitation?** In metric-only gravitation, $G$ was a constant and therefore, when the unit of length changes, the unit of mass must change accordingly. In the scalar-tensor gravitation, $G$ is no longer a constant, and therefore, we can independently change a unit of length $x^i \to x'^i = \lambda \cdot x^i$ *and* a unit of mass $m \to m' = \mu \cdot m$. In this case, the Lagrange function, whose physical meaning is energy$\times$time/length$^4$, transforms as $L \to L' = \mu \cdot \lambda^{-3} \cdot L$. Due to the definition of $\varphi$ as $1/G$, where $m\cdot a = G\cdot m\cdot M/r^2$ and $G = a\cdot r^2/M$, where $a = r/t^2$, we have $G \to G' = \lambda\cdot\mu^{-1}\cdot G$, and $\varphi \to \varphi' = \lambda^{-1}\cdot\mu\cdot\varphi$.

Therefore, the expression

$$L'(x) = \lambda^{-3}\cdot\mu\cdot L(g_{ij}(x)-\eta_{ij}, g_{ij,k}(x), g_{ij,kl}(x),\ldots,\varphi(x),\varphi^{-1}(x),\varphi_{,k}(x),\ldots)$$

leads to

$$L' = g_{\lambda,\mu}(L) = \lambda^{-3}\cdot\mu\cdot L(g_{ij}-\eta_{ij},\lambda\cdot g_{ij,k},\lambda^2\cdot g_{ij,kl},\ldots,$$
$$\lambda\cdot\mu^{-1}\varphi,\lambda^{-1}\cdot\mu\varphi,\lambda^2\cdot\mu^{-1}\cdot\varphi_{,k},\lambda^3\cdot\mu^{-1}\cdot\varphi_{,kl},\ldots). \qquad (13)$$

**Main result for scalar-tensor gravitation.** Now, we are ready for the main result of this section.

**Theorem 5.** *For every scale-invariant final optimal criterion on the set of all Lagrange functions for scalar-tensor gravitation, the optimal Lagrange function has the form*

$$L = a\cdot\varphi\cdot\left(R-\omega\cdot\frac{\varphi_{,i}\cdot\varphi^{,i}}{\varphi^2}\right)+L_0,$$

*where $a$ and $\omega$ are constants, and $L_0$ is an expression which does not contribute to variational equations.*

Thus, we get Brans-Dicke scalar-tensor theory (see, e.g., [12]), which is, therefore, optimal.

**Proof.** Similarly to the proof of Theorem 1, we conclude that the optimal Lagrange function is scale-invariant, that it depends only on the metric, curvature, scalar field, and their covariant derivatives, and, therefore, that it can only contain monomials which do not depend on $\lambda$ and $\mu$. On the other hand, each monomial is proportional to $\lambda^{2n_R+n_D+n_\varphi-n_{-\varphi}-3}\cdot\mu^{n_{-\varphi}-n_\varphi+1}$, where $n_R$ and $n_D$ are defined as in the proof of Theorem 1, $n_\varphi$ is the total number of $\varphi$ and its derivatives in this monomial, and $n_{-\varphi}$ is the total number of terms $\varphi^{-1}$.

Thus, we must have $3 = 2n_R+n_D+n_\varphi-n_{-\varphi}$ and $-1 = n_{-\varphi}-n_\varphi$. Adding these two equalities, we get $2 = 2n_R+n_D$, hence either $n_R = 1$ and $n_D = 0$, or $n_R = 0$ and $n_D = 2$. In both cases, we have $n_\varphi - n_{-\varphi} = 1$.

In the first case, the monomial can only contain $\varphi$, $R_{ijkl}$, and no derivatives. The only covariant term of this type is $a\cdot\varphi\cdot R$.

In the second case, we do not have any curvature terms, and we have two derivatives which can be only applied to $\varphi$. Thus, we have two options: $\varphi_{,i}\cdot\varphi^{,i}/\varphi$ and $\varphi^{;i}_{,i}$. The term corresponding to the second option differs from the term corresponding to the first option by a full derivative; therefore we can replace this term by the term of the first option without changing the variational equations. The theorem is proven.

**Acknowledgements** This work was supported in part by the Chinese Scholarship Council (CSC) and by the US National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

## References

1. R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures On Physics*, Addison-Wesley, Redwood City, CA, 1989.
2. D. R. Finkelstein, *Quantum Relativity: A Synthesis of the Ideas of Einstein and Heisenberg*, Springer-Verlag, Berlin-Heidelberg, 1997.
3. A. Finkelstein, O. Kosheleva, and V. Kreinovich, "Astrogeometry, error estimation, and other applications of set-valued analysis", *ACM SIGNUM Newsletter*, 1996, Vol. 31, No. 4, pp. 3–25.
4. A. Finkelstein, O. Kosheleva, and V. Kreinovich, "Astrogeometry: towards mathematical foundations", *International Journal of Theoretical Physics*, 1997, Vol. 36, No. 4, pp. 1009–1020.
5. A. Finkelstein, O. Kosheleva, and V. Kreinovich, "Astrogeometry: geometry explains shapes of celestial bodies", *Geombinatorics*, 1997, Vol. VI, No. 4, pp. 125–139.
6. A. M. Finkelstein and V. Kreinovich, "Derivation of Einstein's, Brans-Dicke and other equations from group considerations," *On Relativity Theory. Proceedings of the Sir Arthur Eddington Centenary Symposium, Nagpur India 1984*, Vol. 2, Y. Choque-Bruhat and T. M. Karade (eds), World Scientific, Singapore, 1985, pp. 138–146.
7. A. M. Finkelstein, V. Kreinovich, and R. R. Zapatrin, "Fundamental physical equations uniquely determined by their symmetry groups," *Lecture Notes in Mathematics*, Springer-Verlag, Berlin-Heidelberg-N.Y., Vol. 1214, 1986, pp. 159–170.
8. *Group theory in physics: proceedings of the international symposium held in honor of Prof. Marcos Moshinsky, Cocoyoc, Morelos, Mexico, 1991*, American Institute of Physics, N.Y., 1992.
9. V. Kreinovich, "Derivation of the Schroedinger equations from scale invariance," *Theoretical and Mathematical Physics*, 1976, Vol. 8, No. 3, pp. 282–285.
10. L. D. Landau and E. M. Lifschitz, *The Classical Theory of Fields*, Butterworth-Heinemanm, Oxford, UK, 1987.
11. G. W. Leibniz, *New Essays on Human Understanding*, Cambridge University Press, Cambridge, UK, 1996.
12. W. Misner, K. Thorne, and J. A. Wheeler, *Gravitation*, Freeman Co., San Francisco, 1973.
13. H. T. Nguyen and V. Kreinovich, *Applications of Continuous Mathematics to Computer Science*, Kluwer, Dordrecht, 1997.
14. P. J. Olver, *Equivalence, invariants, and symmetry*, Cambridge University Press, Cambridge, N.Y., 1995.
15. *Symmetries in physics: proceedings of the international symposium held in honor of Prof. Marcos Moshinsky, Cocoyoc, Morelos, Mexico, 1991*, Springer-Verlag, Berlin, N.Y., 1992.