Noname manuscript No.
(will be inserted by the editor)

Joule's 19th Century Energy Conservation Meta-Law and the 20th Century Physics (Quantum Mechanics and Relativity): 21st Century Analysis

Vladik Kreinovich · Olga Kosheleva

Received: December 22, 2019 / Accepted: date

Abstract Joule's Energy Conservation Law was the first "meta-law": a general principle that all physical equations must satisfy. It has led to many important and useful physical discoveries. However, a recent analysis seems to indicate that this meta-law is inconsistent with other principles – such as the existence of free will. We show that this conclusion about inconsistency is based on a seemingly reasonable – but simplified – analysis of the situation. We also show that a more detailed mathematical and physical analysis of the situation reveals that not only Joule's principle remains true – it is actually strengthened: it is no longer a principle that all physical theories *should* satisfy – it is a principle that all physical theories *do* satisfy.

Keywords Joule · Energy Conservation Law · Free will · General Relativity · Planck's constant

1 Introduction

Joule's Energy Conservation Law: historically the first meta-law. Throughout the centuries, physicists have been trying to come up with equations and laws that describe different physical phenomenon. Before Joule, however, there were no general principles that restricted such equations.

James Joule showed, in [13–15], that different physical phenomena are inter-related, and that there is a general principle covering all the phenom-

This work was supported in part by the US National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

V. Kreinovich and O. Kosheleva
University of Texas at El Paso, El Paso, Texas 79968, USA
Tel.: +915-747-6951
Fax: +915-747-5030
E-mail: vladik@utep.edu, olgak@utep.edu

ena. Specifically, he showed that energy can be transformed from one type to another – e.g., from mechanical energy to heat, and that in all these transactions, the overall energy is conserved.

This Energy Conservation Law became the first "meta-law", the first general principle that restricts possible physical theories.

Other important meta-laws. Since energy conservation, several other general principles ("meta-laws") were proposed. Probably the most well known is the ideas of (relative) *simplicity* that can be traced back to Einstein (and probably earlier); see, e.g., [25] for a deep analysis of this idea.

Many such meta-laws are mentioned and analyzed in [27]. For example, this book states that all fundamental physical equations should be *reversible* – in the sense that:

- if we can, in principle, go from state A to state B,
- then there should be a way to go from state B to state A.

To the best of our knowledge, all known fundamental physical equations do have this property: e.g., Newton's equations of classical physics, Shroedinger's equations of quantum physics, or Einstein's General Relativity equations that describe the gravitational field. This is a very interesting fact, a fact that makes it difficult to explain why on the *macro*-level, many physical processes are *irreversible*; e.g.:

- a cup can fall down break into pieces, but
- no physical process can transform these pieces back into the original cup state.

We are just mentioning other meta-laws, so that a reader will know that they exist. There are already many papers and books analyzing these meta-laws and their possible physical consequences, as well as many interesting related open problems. A simple overview of all these meta-laws would take several new papers. So, in this paper, we concentrate on the meta-law of energy conservation.

Joule's meta-law in the 20 century: it led to important physical discoveries. Joule's meta-principle turned out to be very helpful for working physics: by restricting possible physical theories, it helped find the correct ones.

A classical example of this help is the discovery of neutrinos; see, e.g., [7, 29]. Specifically, it has been known that while combined neutrons and protons form stable atomic nuclei, stand-alone neutrons are not stable: they decay into protons, emitting electrons in the process. The puzzling problem was that the total energy of the resulting proton and electron is smaller than the energy of the original neutron – which seemingly contradicted to the energy conservation law. To preserve this law, Fermi conjectured that a yet unknown particle – which he called "small neutron" (*neutrino* in Italian) was capturing ("stealing") the missing energy, in what was later called an *urca*-process, after the Russian slang word for a small thief. And yes, neutrinos were later found

– so the Energy Conservation Law not only survived, once again it proved to be very helpful for working physics.

Joule's meta-law in the 20 century: it naturally follows from symmetries. In the 20th century, Noether's Theorem showed that energy conservation is indeed a fundamental principle – since it follows from the natural idea that in fundamental physical equations, there is no fixed moment of time, and that thus all physical equations should not change if we simply select a different starting point for measuring time [7, 20, 29].

What we do in this paper: detailed analysis of Joule's meta-law reveals unexpected subtleties. At first glance, Joule's principle triumphed:

– it is useful: it has led to important empirical discoveries,
– it is natural: it can be derived from the intuitively clear fundamental properties of time.

What more can we expect from a physical idea?

However, in this paper, we show that from the viewpoint of 20th century physics – quantum mechanics and relativity theory — the situation with energy conservation is not so simple and not so straightforward. Let us briefly outline this paper's logic and structure.

At first glance, Joule's principle triumphs: in quantum physics, energy is always preserved. Since historically, quantum ideas came first – around 1900 – we start this paper with analyzing how quantization affects energy conservation. First, in Section 2, following [18], we analyze how the very *existence* of quantization affects *non-quantum* physical effects.

This analysis is based on the fact that, according to modern physics, the world is quantum, but not every possible classical theory has a quantum analog, such analogs exist only for Lagrangian theories – this is why nowadays, even non-quantum physical theories are usually formulated in this way.

Via the above-mentioned Noether's theorem, we can explicitly describe, for each such theory, an expression for energy – which is always conserved.

On second glance, the situation is not so straightforward: some quantity is preserved, but is this quantity really energy? As we show in Section 2, there is a serious problem with the above result.

To illustrate this problem, we provide an example of a simple system in which energy – in the physical sense of this word – is clearly *not* conserved, but by using the formalism, we get some well-preserved quantity. Moreover, we show that for *any* physical systems – whether a physical energy is preserved there or not – the quantity defined by Noether's theorem is preserved.

This shows that a more mathematical Noether-theorem-based notion of "energy" is not always the same as the physical energy.

On third glance, maybe energy is never preserved? The situation becomes even more complicated if, instead of classical *approximations* to quantum theories, we consider the *actual* fully *quantum* theories. In Section 3, following [17] and we show that, for such theories, an intuitive notion of free will always leads to *non*-conservation of energy. This conclusion may seem to "kill" Joule's meta-law.

Maybe we cannot even check whether energy is conserved? In Section 4, following [19], we show that in the quantum case, there is another problem with energy conservation. Namely, in quantum physics, it is not easy even to describe the possibility of possible energy non-conservation: this possibility leads us *outside* traditional quantum physics, to alternative theories in which Planck's constant turns into a new physical field.

Relativity saves the day: energy is preserved if we also take relativistic effects into account. Up to now, we only considered *quantum* effects. However, to have a fully adequate description of the real world, we also need to take *relativistic* effects into account.

In Section 5, we show that taking these effects into account saves Joule's meta-law. Even if we start with a system in which energy is not conserved, then in the relativistic version of this system, we get energy conservation: a very strong gravitational field compensates for the decreased energy of the original system.

In other words, in the relativistic version of the original theory, energy does not disappear, it simply gets transformed into the gravitational energy – just like in Joule's experiments, mechanical energy and heat energy got transformed into each other.

So not only Joule's principle becomes valid again – this principle is strengthened. It is no longer a principle that all physical theories *should* satisfy – it is a principle that all physical theories *do* satisfy.

2 How the Possibility of Quantization Affects Energy Conservation in a Non-Quantum Theory

In this section, we analyze how the possibility of quantization affects a non-quantum theory.

In Section 2.1, we recall that before quantum physics, physical systems were usually described in terms of differential equations, and we provide two simple examples of such systems:

- a system in which energy is conserved, and
- a system with friction, in which energy is not conserved.

In Section 2.2, we recall that not every classical physical theory has a quantum analog – such an analog exists only for theories described by the *minimal action principle*. Physics textbooks usually claim that for such theories, according to Noether's theorem, energy is always conserved.

Our point is that *some* quantity is indeed conserved – a theorem is a theorem, but is this quantity energy? In Section 2.3, we consider the example of a system with friction, a system is which energy – in the physical sense of this word – is clearly *not* conserved. For this system, the Noether-theorem-based quantity *is* conserved, but clearly this quantity is *not* the same as what physicists would call energy. Moreover, in Section 2.4, we show that a similar conserved quantity exists for *any* dynamical system – irrespective of whether for this system, the actual energy is preserved or not.

2.1 Systems Described by Differential Equations: For Some Systems, the Energy Is Preserved; for Some, Energy Is Not Preserved

General idea. In some physical systems, the total energy is *conserved.* For example, when the kinetic energy of a particle decreases, the potential energy increases accordingly.

For some physical systems, energy is *not* conserved. For example, if we consider a particle moving with friction, the energy of the particle itself is not conserved: it is transformed into thermal energy of the surrounding medium. So, if we consider the particle on its own, without taking the surrounding medium into account, then the total energy of this particle will not be preserved: it will decrease with time.

Let us give precise examples of these two types of systems. Starting with Newton, physical theories were described by differential equations. Let us therefore provide two examples of corresponding differential equations:

- an equation that describe a system in which energy is *preserved*, and
- an equation for which energy is *not* preserved.

A system in which energy is conserved. Let us provide an example of a physical system which is *conservative* in the sense that its total energy is preserved. This example is a particle in a potential field $V(x) = V(x_1, x_2, x_3)$. Its dynamics is described, in Newtonian mechanics, by Newton's equations

$$m \cdot \ddot{x}_i = -\frac{\partial V}{\partial x_i}, \tag{1}$$

where $\dot{x}_i$, as usual, denotes time derivative. For this particle, the overall energy

$$E = \frac{1}{2} \cdot m \cdot \sum_{i=1}^{3} (\dot{x}_i)^2 + V(x) \tag{2}$$

is conserved: when the kinetic energy $\frac{1}{2} \cdot m \cdot \sum_{i=1}^{3} (\dot{x}_i)^2$ decreases, the potential energy $V(x)$ increases appropriately, and vice versa.

A system in which energy is not conserved. A classical example of a physical system for which energy is not conserved is a system with friction. Its simplest case is when we do not even have any potential field, i.e., when the dynamical equations have the form

$$m \cdot \ddot{x}_i = -k \cdot \dot{x}_i, \tag{3}$$

for some friction coefficient k. This equation can be further simplified into

$$\ddot{x}_i = -k_0 \cdot \dot{x}_i, \tag{4}$$

for $k_0 \stackrel{\text{def}}{=} \dfrac{k}{m}$.

A system that follows this equation slows down, its velocity (and hence, its kinetic energy) exponentially decreases with time – without being transferred into any other type of energy.

Comment. In real life, this non-conservation of energy means that the system described by the equation (4) is not closed: the energy lost in this system is captured by other objects. For friction, it is very clear where this energy goes: it gets transformed into the thermal energy, i.e., into kinetic energy of individual molecules in the surrounding medium. So, it we consider a system consisting of the particle and of the surrounding medium, the energy will be preserved. However, if we consider only the decelerating particle by itself, the energy decreases.

2.2 Possibility of Quantization Seems to Lead to Energy Conservation

It is not easy to quantize a system of differential equations. Newton's physics was originally formulated in terms of differential equations. Differential equations is now most physical theories were formulated for centuries after that.

The problem with this description is that, according to modern physics, the correct picture of the physical world comes from quantum mechanics. From this viewpoint, even when we observe non-quantum phenomena, these phenomena are simply an approximation of the corresponding quantum theory. In other words, every non-quantum physical theory is *quantizable* – i.e., is a non-quantum limit case of an appropriate quantum theory.

How can we build the corresponding quantum theory based on the classical one? This is very difficult to do if we start with a system described by differential equations. In general, a quantum analogue of each physical quantity is an operator: there are operators describing coordinates x, operators describing component of the momentum, etc. In principle, we can replace each quantity in the original differential equation with an appropriate operator.

This works perfectly for simple cases, e.g., for Newton's equations that connect coordinates x_i and components $p_i \stackrel{\text{def}}{=} m \cdot \dfrac{dx_i}{dt}$ of the momentum vector. In the first approximation, the usual quantum mechanics is described by the usual Newton's equations

$$\frac{dx_i}{dt} = \frac{1}{m} \cdot p_i, \quad \frac{dp_i}{dt} = f_i, \tag{5}$$

with the only difference that instead of scalars x_i and p_i, we now consider operators. This description was first found by P. Ehrenfest (see, e.g., [7]).

The problem is that many physical equations are *non-linear*, so, in addition to linear terms, they have, e.g., terms proportional to the product $q \cdot q'$ of two quantities. In the classical differential equations, the quantities are described by numbers, and for numbers, the result of multiplication does not depend on the other: the product $q \cdot q'$ is equal to the product $q' \cdot q$. In contrast, operators are, in general, not commutative, so we get two different quantum theories depending on whether we take $q \cdot q'$ or $q' \cdot q$.

In general, there are many such choices. How can we select one of them? Usually, we have several related theories between which we want to preserve relations, and we have general properties that we want to preserve (e.g., some symmetries). If we start with a system of differential equations, it is very difficult to select proper quantization that would preserve the desired relations and/or properties – and sometimes even impossible to make such a selection.

Minimal action principle. So how can we form a quantum version of a classical theory? A solution to this problem was proposed by R. Feynman. This solution is based on the fact that most physical theories can be equivalently described in terms of the *minimal action* principle: the actual dynamics of particles and fields is the one that minimizes a special physical quantity called *action* S.

For particles, action has the form $S = \int L(x(t), \dot{x}(t))\,dt$, where the function $L(x(t), \dot{x}(t))$ is known as the *Lagrangian.* For example, for the Newtonian particle in a potential field $V(x)$, the Lagrangian has the form

$$L = \frac{1}{2} \cdot m \cdot \sum_{i=1}^{3} (\dot{x}_i)^2 - V(x). \qquad (6)$$

For fields $f(x)$, ..., the action S has a similar form

$$S = \int L(f(x), \ldots, f_{,i}(x), \ldots)\,dx,$$

where $f_{,i}$ denotes the corresponding partial derivative $f_{,i} \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}$.

At present, most new physical theories are proposed in the form of the corresponding Lagrangian.

Feynman's integration over trajectories: a general approach to quantization. The Nobelist Richard Feynman proposed a general quantization procedure known as integration-over-trajectories approach. In this approach, the amplitude $\psi_{A,B}$ of a transition from a state A to the state B is proportional to the "sum" (integral) of the expression $\exp\left(\mathrm{i} \cdot \dfrac{S}{\hbar}\right)$ over all trajectories leading from A to B. The probability to observe the transition into different states B is proportional to the squared absolute value of this amplitude $|\psi_{A,B}|^2$; see, e.g., [7, 20].

This explains the ubiquity of the Lagrangian approach. Because of the need to restrict ourselves to quantizable theories, the Lagrangian approach is now ubiquitous in physics.

Let us recall how energy conservation is treated in this approach.

Energy conservation in the Lagrangian approach: what physics textbooks say. Once we know the Lagrangian, we can use Euler-Lagrange equations to derive the corresponding differential equations. For particles, these equations take the form

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) = 0. \tag{7}$$

One can easily check that for the Newtonian Lagrangian (6), we get exactly Newton's equations (1). For fields, the equations take the form

$$\frac{\partial L}{\partial f} - \sum_i \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial f_{,i}}\right) = 0. \tag{8}$$

In the Lagrange approach, the energy of a particle is formally defined as

$$E_L \stackrel{\text{def}}{=} \sum_i \dot{x}_i \cdot \frac{\partial L}{\partial \dot{x}_i} - L. \tag{9}$$

One can check that for the Newtonian Lagrangian (6), we get the standard expression (2) for energy. A similar expression defines energy for a field theory [7, 20]:

$$E_L = \sum_f \sum_i f_{,i} \cdot \frac{\partial L}{\partial f_{,i}} - L. \tag{10}$$

It can be proven that for each Lagrangian, this quantity is always preserved, in the usual sense of this word: its value does not change with time.

Where are we in relation to Joule's principle. Let us summarize where we are in relation to energy conservation law. At first glance, we are good: quantization necessitates the Lagrangian formalism, and this formalism always leads to energy conservation.

However, as we will show in the next section, the situation is not so simple: yes, there is a conserved quantity – that for Newton's mechanics coincides with energy – but does this quantity always coincide with what physics mean by energy? Our answer is No, and here is an example.

2.3 A Physical Quantity Is Preserved, but This Quantity Is Not What Physics Means by Energy: An Example

Description of the simple example. Let us consider the simplest possible example of a physical system in which, from the physical viewpoint, energy is not conserved: a 1-D particle with friction, whose dynamics is described by the equation

$$\ddot{x}(t) = -k_0 \cdot \dot{x}(t). \tag{11}$$

In time, this particle slows down, its kinetic energy decreases, and since there is no potential energy or thermal energy here, this means that its overall energy decreases too.

This is when k_0 is positive. If k_0 is negative, then, vice versa, the particle by itself starts accelerating. So, we can make a perpetuum mobile out of it:

- let it accelerate, then
- let it do some useful work (e.g., rotate a wheel) while decelerating, then
- let us accelerate again, etc.

This is exactly a type of physical theory that Joule's principle is trying to prohibit.

What we will do. We will show that this system can be described by a Lagrangian and thus, for this system, "energy" (as defined in the Lagrangian formalism) is well conserved. This will show that – at least on this example – the Lagrangian formalism does not adequately convey the physical meaning of energy conservation.

Comment. In the next subsection, we show that this inadequacy is not a freaky property of this particular simple system: a generic dynamical system can be described by an appropriate Lagrangian.

Towards finding an appropriate Lagrangian. The classical Newtonian Lagrangian (6) is a sum of two terms: a term depending only on $\dot{x}_i$ and a term depending only on x_i. Let us look for a similar type Lagrangian for describing the equation (11), i.e., let us look for a Lagrangian of the type

$$L = a(\dot{x}) + b(x), \tag{12}$$

for some functions $a(\dot{x})$ and $b(x)$. For this Lagrangian, Euler-Lagrange equations (7) lead to

$$b'(x) - \frac{d}{dt}a'(\dot{x}) = 0, \tag{13}$$

where $b'(x)$ and $a'(\dot{x})$, as usual, indicated derivatives of the corresponding functions. By applying the chain rule to the formula (13), we get

$$b'(x) - a''(\dot{x}) \cdot \ddot{x} = 0. \tag{14}$$

We want to find a Lagrangian that leads to differential equation (11). For this Lagrangian, the formula (14) will be true when we substitute the expression (11) for the acceleration $\ddot{x}$. As a result, we get the following formula

$$b'(x) + k_0 \cdot a''(\dot{x}) \cdot \dot{x} = 0, \tag{15}$$

i.e., equivalently,

$$k_0 \cdot a''(\dot{x}) \cdot \dot{x} = -b'(x) \tag{16}$$

for all possible values x and $\dot x$.

The left-hand side of the formula (16) does not depend on $\dot x$, and its right-hand side does not depend on x. Since these two sides are equal, this means that this expression cannot depend neither on x nor on $\dot x$ and is, therefore, a constant. Let us denote this constant by C. Then, from the condition that the right-hand side is equal to this constant, we conclude that $b'(x) = -C$, hence $b(x) = -C \cdot x + C_0$. The constant term C_0 in the Lagrangian does not affect the corresponding equations (7) and can thus be safely ignored. So, we have $b(x) = -C \cdot x$.

Similarly, from the condition that the left-hand side of the formula (16) is equal to the constant C, we conclude that

$$k_0 \cdot a''(y) \cdot y = C, \tag{17}$$

where, for simplicity, we denoted $y \stackrel{\text{def}}{=} \dot x$. From (17), we conclude that

$$a''(y) = \frac{C}{k_0 \cdot y}. \tag{18}$$

Integrating over y, we get

$$a'(y) = \frac{C}{k_0} \cdot \ln(y) + C_0, \tag{19}$$

and, integrating once again, that

$$a(y) = \frac{C}{k_0} \cdot y \cdot \ln(y) + C_0 \cdot y + C_1. \tag{20}$$

Ignoring the constant C_1 and taking into account that $L(x, \dot x) = a(\dot x) + b(x)$ and that $b(x) = -C \cdot x$, we get the following expression for the desired Lagrangian:

Resulting Lagrangian. The system (11) can be described by the Lagrangian

$$L(x, \dot x) = \frac{C}{k_0} \cdot \dot x \cdot \ln(\dot x) + C_0 \cdot \dot x - C \cdot x. \tag{21}$$

Comment. One can easily check that for this Lagrangian, Euler-Lagrange equations (7) indeed lead to the equations (11).

Resulting expression for conserved "energy". Here,

$$\frac{\partial L}{\partial \dot x} = \frac{C}{k_0} \cdot (\ln(\dot x) + 1) + C_0.$$

Thus, applying the usual formula (9) to the Lagrangian (21), we get the expression

$$E_L = \dot x \cdot \frac{\partial L}{\partial \dot x} - L = \frac{C}{k_0} \cdot \dot x + C \cdot x. \tag{22}$$

One can easily check that this "energy" is indeed conserved. Indeed, here

$$\frac{dE_L}{dt} = \frac{d}{dt}\left(\frac{C}{k_0} \cdot \dot{x} + C \cdot x\right) = \frac{C}{k_0} \cdot \ddot{x} + C \cdot \dot{x}. \tag{23}$$

Substituting the expression $\ddot{x} = -k_0 \cdot \dot{x}$ into this formula, we indeed get

$$\frac{dE_L}{dt} = 0.$$

Where are we in relation to Joule's principle. At first glance, as we have mentioned earlier, the possibility of quantization seems to boost energy conservation law: this possibility implies that the physical theory must be described in Lagrangian terms, and for such theories, energy (properly defined) is always preserved.

Alas, the above simple example shows that already for the simplest possible system in which physical energy is not preserved, this mathematically defined "energy" is not really the same as physical energy. In other words, some quantity *is* preserved; in Newtonian physics, this quantity indeed coincides with energy, but in general, this preserved quantity is *not* what a physicist would call energy.

A natural question. A natural question is: how general is this phenomenon? Maybe only for the above simple system we get a quantity which differs from energy, but other systems, the above mathematical definition of energy makes perfect physical sense? In the next subsection, we show that the friction example *is* typical: for every dynamical system, whether it preserves energy or not, Lagrangian formalism produces some preserved quantity – which, for systems in which energy is not preserved, cannot coincide with physical energy.

2.4 For Any Theory in Which Energy Is Not Preserved, the Lagrangian-Based Conserved Quantity is *Not* Energy: A General Result

What we do in this subsection. Let us show that a similar Lagrangian reformulation – and thus, the existence of a preserved "energy" – is possible for a generic dynamical system

$$\ddot{x}_i = f_i(x_1, \ldots, x_n, \dot{x}_1, \ldots, \dot{x}_n), \quad i = 1, \ldots, n. \tag{24}$$

A simple multi-D case. Let us start with a multi-D analog of a system with friction, in which the differential equations have the form

$$\ddot{x}_i(t) = -k_0 \cdot \dot{x}_i(t). \tag{25}$$

This system can be described, e.g., by a Lagrangian

$$L = \sum_i \frac{1}{k_0} \cdot \dot{x}_i \cdot \ln(\dot{x}_i) - \sum_i x_i. \tag{26}$$

General case. In the general case, differential equations (7) take the form

$$L_{,x_i} - \frac{d}{dt} L_{,\dot{x}_i} = 0, \tag{27}$$

where $L_{,z}$ denotes partial derivative. By using the chain rule to differentiate the expression $L_{,\dot{x}_i}(x_j, \dot{x}_j)$, we get

$$L_{,x_i} - \sum_j L_{,\dot{x}_i x_j} \cdot \dot{x}_j - \sum_j L_{,\dot{x}_i \dot{x}_j} \cdot \ddot{x}_j = 0. \tag{28}$$

Substituting $\ddot{x}_i = f_i$ into this formula and using notations $y_i = \dot{x}_i$, we get

$$L_{,x_i} - \sum_j L_{,y_i x_j} \cdot y_j - \sum_j L_{,y_i \dot{y}_j} \cdot f_j(x_1, \ldots, x_n, y_1, \ldots, y_n) = 0. \tag{29}$$

Our objective is to define a function $L(x_1, \ldots, x_n, y_1, \ldots, y_n)$ of $2n$ variables for which the second-order partial differential equation (29) holds.

Let us show how we can construct such a function. Let us take, e.g., $L(x_1, \ldots, x_n, 0, \ldots, 0) = 0$ when all the derivatives y_i are equal to 0. Then, we extend it to the case when $y_1 \neq 0$ and $y_2 = \ldots = y_n = 0$. With respect to y_1, (29) becomes a simple second order equation

$$\frac{\partial^2 L}{\partial y_1^2} \cdot f_1 + \frac{\partial^2 L}{\partial y_1 \partial \ldots} \cdot \ldots + \ldots = 0,$$

from which one can explicitly obtain such an extension – e.g., by Euler-style step-by-step integration. Then, we can extend this function along y_2, etc. At the end, we get a function defined for all possible values of x_i and y_j.

Where are we in relation to Joule's principle. We are back at square one. It looks like, contrary to what some textbooks claim, the possibility of quantization – and the resulting existence of a conserved quantity that these textbooks call "energy" – does not help us decide whether *physical* energy is preserved or not.

Even if in a theory, physical energy is clearly not preserved – to the extent that we can build a perpetuum mobile machine – the Lagrangian formalism still produces a preserved quantity, which, in this case, *cannot* coincide with physical energy.

So, based only on the possibility of quantization, we cannot decide whether energy is conserved or not.

What now? Since we cannot make any conclusion about energy conservation based only on the *possibility* of quantization, let us analyze what we can say about Joule's principle based on the *actual* quantization.

3 Quantum Physics, Free Will, and Energy Conservation: A Seeming Contradiction

Free will: main idea. In non-quantum physics, equations enable us to predict all the future states based on the current state of the world. This means that the initial state of the world uniquely pre-determines what will happen in the future. In particular, it pre-determines what we say, where we go, what we do.

In quantum mechanics, we *cannot predict* the exact *events*, but we *can predict* their *probabilities* – i.e., frequencies of different outcomes. For example, in quantum physics, the initial state of the world uniquely pre-determines the frequencies with which people will behave this way or that way – and thus, e.g., pre-determines the situations involving big groups of people, such as wars, revolutions, or other mass movements.

On the other hand, we all know that we can make decision that change the state of the world – even if, for most of us, a little bit. This intuitive idea of free will permeates all our life, all our activities – and it seems to contradict the above-described uniquely pre-determined nature of all the world's processes – including human behavior; see, e.g., [2–5, 8, 9, 11, 12, 22, 24, 26, 28, 30–32] and references therein.

This is a problem that has bothered philosophers for centuries, and many of them – including Einstein himself – believed in a simple way to avoid the above contradiction: that:

– yes, we have an illusion of a free will, but
– in reality, all our thoughts and desires are also uniquely pre-determined.

This idea is known as *compatibilism*; see, e.g., [1]. According to this idea, after a well-lived life, we can die happily thinking that we have made many important decisions that changed the world for the better – but in reality, our decisions and our changes were pre-determined from the very beginning, the same way the solar eclipses can be predicted many centuries from now.

Einstein, by the way, was very serious about this idea: not only he claimed it in his papers, he followed it in everyday life. For example, while he could not swim at all, he would often, to the horror of his wife and his friends, go yachting alone, and his argument was: if it is pre-determined that I drown, there is nothing I can do to change this future event.

This worked for Einstein, and we understand that it is a logically consistent approach, but we do not feel the same way. We lived too long in the Soviet Union, where Hegel's "freedom is the recognition of necessity" formula – picked up by Engels and Lenin and others – was used to justify our non-freedom and our oppression, to explain to us that "freedom" about which the Western societies like to brag is largely an illusion. This experience taught us to value *true* freedom, to believe that, in spite of all – sometimes successful – manipulation attempts, there *is* also a true freedom, that we *can* make different decisions, that our decisions are *not* uniquely pre-determined.

This was our own path to this idea, but many other people have the same intuition and the same belief.

The belief in true free will is perfectly compatible with classical (non-quantum) physics. Let us start with simple physical systems, such as point particles, whose state $s(t)$ at any given moment of time t can be described by describing the values of finitely many quantities $s_1(t), \ldots, s_n(t)$. For example, in the original Newton's approximate description of celestial bodies as points, to describe the state of each body, it is sufficient to describe the current values x_1, x_2, and x_3 of its three spatial coordinates, three components v_1, v_2, and v_3 of the current velocity, and the body's mass m. In electrodynamics of point particles, we need to add electric change q to the list of these quantities. To describe a system of several interacting points, we need to describe the quantities describing each of these points.

Dynamical equations describe how each of these quantities change:

$$\frac{ds_i}{dt} = f_i(s_1, s_2, \ldots).$$

If we take freedom of will into account, then the change in the state $\frac{ds_i}{dt}$ is no longer uniquely determined by the current state $s(t)$. So, to determine the desired change, we also need to describe the values of some other quantities $w_1, \ldots$, which we can set arbitrarily because of our freedom of will:

$$\frac{ds_i}{dt} = f_i(s_1, s_2, \ldots, w_1, \ldots).$$

There are no differential equations for describing how the quantities w_k change, since we can change them at will.

In quantum physics, the situation is drastically different. In quantum physics, the situation is different. In quantum physics, the state of the world at any given moment of time t is described by a wave function $\psi(t)$, and the change in this state is described by Schrödinger's equation

$$\mathrm{i} \cdot \hbar \cdot \frac{d\psi}{dt} = H\psi. \tag{30}$$

In this equation, the change is determined by the Hamilton operator H that describes the total energy of the system.

So, if we want to allow free will (in the above sense), if we want the ability to change the derivative $\frac{d\psi}{dt}$, we have to be able to change the Hamilton operator.

This leads to non-conservation of energy. In quantum case, as we have concluded, freedom of will means that we can modify the Hamilton operator, the operator that described the total energy of the system. What does it mean that the Hamilton operator changes? It means for the some states, the energy value changes. Thus, in effect, in quantum physics, freedom of will means that, by exercising our will, we can change the total energy of the system. In other words, *in quantum physics, free will seems to lead to non-conservation of energy.*

Where are we in relation to Joule's principle. From the viewpoint of free will and quantum physics, it looks like energy is *not* conserved, it looks like Joule's meta-law is a thing of the past.

Comment. A possible objection to this may be that so far, all experimental evidence has confirmed energy conservation law. However, this does not invalidate our claim. Indeed, as we have mentioned earlier, in classical (pre-quantum) physics, freedom of will does not necessarily lead to energy non-conservation.

Thus, energy non-conservation caused by the freedom of will is a purely quantum effect, that disappears in the classical limit, when the Planck's constant $\hbar$ tends to 0. So, this purely quantum effect should be proportional to $\hbar$ and thus, it should be reasonably small. This smallness explains why this effect have not been observed: in our usual free-will decisions, we control macro-size objects, objects for which the quantum-size microscopic changes in energy are not easy to measure.

What next? In the previous paragraph, we mentioned the fact that energy conservation has been experimentally confirmed. This is a natural thing to mention, since physics in general is about the physical world. So to see which physical laws, which physical principles are valid, a reasonable idea is to compare their predictions with the results of actual measurements and observations.

Thus, a natural question is: how can we check that energy is indeed conserved? In other words, how can we experimentally compare a theory in which energy is conserved and a theory in which energy is not conserved? To be able to do that in a quantum context, we need to be able to describe quantum theories in which energy is not conserved. In the next section, we will show that in a pure quantum physics, this is not possible.

4 Another Quantum-Related Problem with Energy Conservation: Looks Like We Cannot Even Test It in Quantum Context

As we have just mentioned, in order to test whether energy is conserved, we need to compare the experimental results with (at least) two models:

– a model in which energy is conserved, and
– a model in which energy is not conserved.

So, in quantum context, to make this comparison possible, we need to have a quantum theory in which energy is not conserved. In this chapter, following [19], we show that if we try to do that, then we immediately get outside the usual quantum physics: e.g., we need modifications of quantum physics in which Planck's constant is no longer a constant, but a new field who value changes from one point to another.

It is worth mentioning that the need to test the energy conservation law on quantum level is not purely theoretical:

– on a pragmatic level, serious physicists considered the possibility of micro-violations of energy conservation starting from the 1920s [7];
– as the previous section shows, on a more foundational level, the intuitive ideas of free will seem to lead to possible energy non-conservation.

Ehrenfest's equations: reminder. In quantum physics (see, e.g., [7]), states are described by elements of a Hilbert space – e.g., of the space of all square-integrable functions $\psi(x)$ – and physical quantities are described by linear operators in this space.

Historically, quantum physics started with a description by Heisenberg, in which states are fixed but operators change.

Very soon, it turned out that in most cases, an alternative representation is more computationally advantageous – a representation in which operators are fixed but states change. This representation was originally proposed by E. Schroedinger and is therefore known as the Schroedinger representation. However, in many cases, the Heisenberg's representation is still used, since it is closely related to the corresponding classical (pre-quantum) equations. One such case is Ehrenfest's description of the quantum analogue of Newton's mechanics.

In the Heisenberg representation, physical quantities like coordinates x_i and components $p_i \stackrel{\text{def}}{=} m \cdot \dfrac{dx_i}{dt}$ of the momentum vector are represented by operators. In the first approximation, the usual quantum mechanics is described by the usual Newton's equations

$$\frac{dx_i}{dt} = \frac{1}{m} \cdot p_i, \quad \frac{dp_i}{dt} = f_i, \tag{31}$$

with the only difference that instead of scalars x_i and p_i, we now consider operators. This description, as we have mentioned earlier, was first found by P. Ehrenfest (see, e.g., [7]).

The difference between the scalars and operators is that operators, in general, do not commute, i.e., in general, for two operators a and b, we have $[a, b] \stackrel{\text{def}}{=} ab - ba \neq 0$. Specifically, in the usual quantum physics, operators x_i and x_j corresponding to different coordinates commute with each other, operators p_i and p_j commute with each other, but operators x_i and p_i do not commute:

$$[x_i, x_j] = 0, \quad [p_i, p_j] = 0, \quad [p_i, x_j] = \mathrm{i} \cdot \hbar \cdot \delta_{ij}, \tag{32}$$

where the Kronecker's delta δ_{ij} is equal to 1 when $i = j$ and to 0 otherwise. For the usual energy-preserving quantum mechanics, these commuting relations get conserved as the operators x_i and p_i change in time – in accordance with Ehrenfest equations (31).

Classical (pre-quantum) physics as a limit case of Ehrenfest-type equations. As we have mentioned, the main advantage of Ehrenfest-type equations is that they provide a good relation to classical (non-quantum) physics. Namely, in the classical limit, operators x_i and p_j simply turn into the corresponding scalar quantities.

It is also known that in the first approximation, the commutator $[a, b]$ can be described in terms of the Poisson brackets (see, e.g., [7]). Namely, for arbitrary functions $a(x, p)$ and $b(x, p)$ of coordinates $x = (x_1, x_2, x_2)$ and momentum $p = (p_1, p_2, p_3)$, we have

$$[a, b] = \mathrm{i} \cdot \hbar \cdot \{a, b\} + o(\hbar), \tag{33}$$

where

$$\{a, b\} \stackrel{\text{def}}{=} \sum_k \left(\frac{\partial a}{\partial p_k} \cdot \frac{\partial b}{\partial x_k} - \frac{\partial a}{\partial x_k} \cdot \frac{\partial b}{\partial p_k} \right). \tag{34}$$

As an example, let us show what happens for the Heisenberg commutator $[a, b]$ for which $a = p_i$ and $b = x_j$. Since $a = p_i$ depends only on p_i, we have $\frac{\partial p_i}{\partial p_k} = \delta_{ik}$ and $\frac{\partial p_i}{\partial x_k} = 0$. Similarly, since $b = x_j$ depends only on x_j, we have $\frac{\partial x_j}{\partial p_k} = 0$ and $\frac{\partial x_j}{\partial x_k} = \delta_{jk}$. Thus,

$$\{p_i, x_j\} = \sum_k \delta_{ik} \cdot \delta_{jk} = \delta_{ij}.$$

When the force comes from a potential field, we have a consistent quantum description. Let us show that in the potential field with potential energy $V(x)$, when $f_i = -\frac{\partial V}{\partial x_i}$, the commutator relation $[p_i, x_j] = \mathrm{i} \cdot \hbar \cdot \delta_{ij}$ is preserved by the system's dynamics.

Indeed, let us assume that at a certain moment of time t_0, we have $[p_i, x_j] = \mathrm{i} \cdot \hbar \cdot \delta_{ij}$. Let us show that – at least in the first approximation – this relation is conserved, in the sense that $\frac{d}{dt}[p_i, x_j] = 0$.

First, we should note that since $[a, b] = ab - ba$, we have

$$\frac{d}{dt}([a, b]) = \frac{d}{dt}(ab - ba) = \frac{da}{dt}b + a\frac{db}{dt} - \frac{db}{dt}a - b\frac{da}{dt} = \left[\frac{da}{dt}, b\right] + \left[a, \frac{db}{dt}\right].$$

Thus, we have

$$\frac{d}{dt}([p_i, x_j]) = \left[\frac{dp_i}{dt}, x_j\right] + \left[p_i, \frac{dx_j}{dt}\right]$$

Due to Ehrenfest's equations, we have $\frac{dp_i}{dt} = f_i$ and $\frac{dx_i}{dt} = \frac{1}{m} \cdot p_i$, so:

$$\frac{d}{dt}([p_i, x_j]) = [f_i, x_j] + \frac{1}{m} \cdot [p_i, p_j]. \tag{35}$$

Since f_i depend only on the coordinates, and all coordinate operators commute, we have $[f_i, x_j] = 0$. Since all the components of the momentum commute, we have $[p_i, p_j] = 0$. Thus, we conclude that $\frac{d}{dt}([p_i, x_j]) = 0$.

Similarly, we can conclude that the second derivative of the Heisenberg commutator is also equal to 0. Indeed, by differentiating both sides of the equation (35), we conclude that

$$\frac{d^2}{dt^2}([p_i, x_j]) = \left[\frac{df_i}{dt}, x_j\right] + \frac{1}{m} \cdot [f_i, p_j] + \frac{1}{m} \cdot [f_i, p_j] + \frac{1}{m} \cdot [p_i, f_j]\,. \tag{36}$$

Here, since f_i depends only on coordinates, we have

$$\frac{df_i}{dt} = \sum_\ell \frac{\partial f_i}{\partial x_\ell} \cdot \frac{dx_\ell}{dt} = \frac{1}{m} \cdot \sum_\ell \frac{\partial f_i}{\partial x_\ell} \cdot p_\ell,$$

so

$$\frac{d^2}{dt^2}([p_i, x_j]) = \frac{1}{m} \cdot \left(\left[\sum_\ell \frac{\partial f_i}{\partial x_\ell} \cdot p_\ell, x_j\right] + 2[f_i, p_j] + [p_i, f_j]\right).$$

Thus, to prove that this second derivative is equal to 0, it is sufficient to prove that the expression in parentheses is equal to 0. In the first approximation, this expression is proportional to the sum S of the corresponding Poisson brackets

$$S = \left\{\sum_\ell \frac{\partial f_i}{\partial x_\ell} \cdot p_\ell, x_j\right\} + 2\{f_i, p_j\} + \{p_i, f_j\};$$

so, in the first approximation, it is sufficient to prove that the sum S is equal to 0. In the first bracket, x_j depends only on x_j, so

$$\left\{\sum_\ell \frac{\partial f_i}{\partial x_\ell} \cdot p_\ell, x_j\right\} = \sum_k \sum_\ell \frac{\partial f_i}{\partial x_\ell} \cdot \delta_{k\ell} \cdot \delta_{jk} = \frac{\partial f_i}{\partial x_j}.$$

For the second term of the sum S, since p_j only depends on the momentum, we get

$$\{f_i, p_j\} = -\sum_k \frac{\partial f_i}{\partial x_k} \cdot \delta_{jk} = -\frac{\partial f_i}{\partial x_j}.$$

Similarly,

$$\{p_i, f_j\} = \frac{\partial f_j}{\partial x_i}.$$

Thus, we have

$$S = \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}.$$

For the potential field, we have $f_i = -\dfrac{\partial V}{\partial x_i}$ and therefore,

$$\frac{\partial f_i}{\partial x_j} = -\frac{\partial^2 V}{\partial x_i \partial x_j}.$$

Hence, we have

$$S = -\frac{\partial^2 V}{\partial x_i \partial x_j} + \frac{\partial^2 V}{\partial x_i \partial x_j} = 0.$$

What happens when energy is not conserved: an example. We are interested in situations where energy is not conserved. Let us start our analysis with the simplest such situation of the friction force $f_i = -k \cdot v_i$, i.e., force of the type $f_i = -k_0 \cdot p_i$, where $k_0 \stackrel{\text{def}}{=} \frac{k}{m}$. In this case, from the formula

$$\frac{d}{dt}([p_i, x_j]) = [f_i, x_j] + \frac{1}{m} \cdot [p_i, p_j],$$

by using $[p_i, p_j] = 0$, we get

$$\frac{d}{dt}([p_i, x_j]) = -k_0 \cdot [p_i, x_j].$$

In other words, for $h \stackrel{\text{def}}{=} [p_i, x_i]$, we have a differential equation

$$\frac{dh}{dt} = -k_0 \cdot h.$$

From this equation, we conclude that $\frac{dh}{h} = -k_0 \cdot dt$ hence $\ln(h) = \text{const} - k_0 \cdot t$, and $h(t) = \text{const} \cdot \exp(-k_0 \cdot t)$. We know that $h(t_0) = \mathrm{i} \cdot \hbar$, hence

$$h(t) = h(t_0) \cdot \exp(-k_0 \cdot (t - t_0)).$$

At the initial moment t_0, we have $h(t_0) = \mathrm{i} \cdot \hbar$. So, the above equation means, in effect, that Planck's constant $\hbar \stackrel{\text{def}}{=} \frac{[p_i, x_i]}{\mathrm{i}}$ is no longer a constant – it exponentially decreases with time.

Discussion. Let us show that the same phenomenon – of Planck's constant no longer being a constant – occurs for *every* theory in which energy is not conserved.

Formulation of the main result. Let us consider the general case, when each component f_i of a force is a function of coordinates x and momentum p. We will show that if in the quantum version of this theory, Planck's constant remains a constant, i.e., we have $[p_i, x_j] = \mathrm{i} \cdot \hbar \cdot \delta_{ij}$ for all moments of time, then the field f_i is a potential field, i.e., has the form $f_i = -\frac{\partial V}{\partial x_i}$ for some function $V(x)$.

This means that if f_i is *not* a potential field, then Planck's constant is no longer a constant.

Proof. Indeed, suppose that we have $\frac{d}{dt}([p_i, x_j]) = 0$. Explicitly differentiating the left-hand side, we conclude that $[f_i, x_j] + \frac{1}{m} \cdot [p_i, p_j] = 0$. Since $[p_i, p_j] = 0$, we get $[f_i, x_j] = 0$. In the first approximation, this means that the corresponding Poisson bracket is equal to 0: $\{f_i, x_j\} = 0$. Since x_j depends only on the coordinate, we get

$$\{f_i, x_j\} = \sum_k \frac{\partial f_i}{\partial p_k} \cdot \delta_{kj} = \frac{\partial f_i}{\partial p_j} = 0.$$

The fact that all partial derivatives of f_i relative to p_j are equal to 0 means that f_i does not depend on the momentum. In other words, the force f_i depends only on the coordinates x_j.

Now, since we know that f_i depend only on the coordinates, for the second time derivative $\frac{d^2}{dt^2}([p_i, x_j])$, we can repeat arguments from the previous section and conclude that in the first approximation, this second derivative is proportional to

$$S = \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}.$$

So, from the fact that the second derivative is equal to 0, we conclude that $S = 0$, i.e., that

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$$

for all i and j. It is known that these equalities are necessary and sufficient conditions for the existence of a field V for which $f_i = -\frac{\partial V}{\partial x_i}$. Thus, we have proved that f_i is indeed a potential field.

The conclusion has been proven.

Where are we in relation to Joule's principle. In the previous section, we showed that in quantum context, if we take free will into account, then energy will *not* be preserved. In other words, we show that in this sense, energy conservation is *not* compatible with quantum physics. In this section, we describe another difficulty in trying to relate quantum physics and energy conservation: that, strictly speaking, in quantum physics, we cannot even formulate the question of experimentally testing whether energy is conserved. To be able to even formulate this question properly, we need to go beyond the usual quantum physics – to its modifications in which, e.g., Planck's constant is no longer a constant.

Based on the previous two sections, it looks like the relation between quantum physics and energy conservation is a disaster :-(

Comment. It should be mentioned that theories in which Planck's constant is no longer a constant but a new physical field $s(x)$ have actually been proposed; see, e.g., [23].

These theories only consider a *scalar* field $s(x)$ corresponding to

$$[p_i, x_j] = \mathrm{i} \cdot \hbar \cdot s(x) \cdot \delta_{ij}.$$

In general, we may need to go even further, to the commutator $[p_i, x_j]$ can be an arbitrary *tensor*.

What now? The main gist of this section is that to be able to deal with a possibility that energy may be not conserved, we need to go beyond the usual quantum physics.

In this section, we assumed that the space is Euclidean, that the time is the usual time, and we concluded that in this case, to describe a possible non-conservation, we need to go beyond quantum physics. From this, we concluded that quantum physics has a problem with energy conservation.

Now that we have written it down explicitly, it is easy to see what is the problem with this argument: the problem is that, according to relativity theory, Euclidean space-time is only an *approximation* to the actual curved one. So, let us repeat this analysis, but this time we will take relativistic effects into account.

5 Taking Relativistic Effects into Account Resolves the Puzzle

So far, out of the two main contributions of 20 century to physics – quantum mechanics and relativity – we only took into account quantum mechanics. Let us now take into account relativity as well.

Before we show how to take into account effects of *general* relativity (i.e., curved space-time), let us briefly describe how *special* relativity affects questions related to energy conservation.

Special relativity: a brief reminder. The main idea behind special relativity is the relativity principle – which goes back to Galileo. Galileo noticed that when a ship smoothly goes forward, then a ship's passenger in a cabin with no windows will not be able to tell whether the ship is moving or not – and how fast it is moving. Einstein made this principle a cornerstone of his Special Relativity Theory: that all the equations describing physical processes will remain the same if we consider them from the viewpoint of a moving observer, as long as this observer moves in the same direction with a constant speed.

In relativity theory, many physical quantities are *relative* – in the sense that their numerical values are not absolute, they change if we measure them in a moving system. For example, the object's velocity is relative:

- the velocity of a fast moving object may be huge when we measure it from the Earth, but
- from the viewpoint of an observer who travels on this objects, the object's velocity is 0.

Energy and energy conservation in special relativity: a brief reminder. Similarly to velocity, energy is also relative:

- from the viewpoint of the observer on the Earth, a fast moving object has a huge kinetic energy, but
- from the viewpoint of the observer who travels on this object, the object's kinetic energy is 0.

To describe the object's energy in all possible coordinate systems, we need to use a 4-dimensional energy-momentum vector (p_0, p_1, p_2, p_3), in which the first component p_0 is the object's energy and the other three components p_1, p_2, and p_3 are components of the object's momentum. Similarly, to describe the energy density in all possible coordinate systems, we need to use a 4-dimensional energy-stress tensor T_{ij}; the actual energy density corresponds to the component T_{00}.

In terms of the tensor T_{ij}, energy conservation can be expressed as $T_{ij,j} = 0$, where:

– as before, $a_{,i}$ means partial derivative with respect to the i-th coordinate, and
– it is implicitly assumed that we add over repeated indices, i.e., in this case, that we actually mean the formula $\sum\limits_j T^{ij}_{,j} = 0$. (It is worth mentioning that the above simplifying sum-less notation was first introduced by Einstein himself and is thus known as *Einstein's notations*.)

Now, we are ready to describe the effects of general relativity.

General relativity equations: a brief reminder. According to Einstein's General Relativity, the equations for the metric tensor field g_{ij} (that describes gravity, i.e., curved space-time) have the following form (see, e.g., [7,21,29]):

$$G_{ij} = T_{ij},$$

where T_{ij} is the stress-energy tensor,

$$G_{ij} \stackrel{\text{def}}{=} R_{ij} - \frac{1}{2} R g_{ij},$$

and R_{ij} and R are special expressions in terms of the components of the metric tensor and their first and second order derivatives.

The observable effects of the metric field g_{ij} is that a freely moving particle – which in the empty space-time will follow a straight line trajectory in which velocity does not change $\dot{v}_i = 0$, in full accordance with the law of inertia – will follow a more complex trajectory, in which the derivative $\dot{v}_i$ is proportional to a special combination Γ of the first derivatives of the metric tensor g_{ij}.

General relativity as a particular case of flat-space-time field theory. Einstein derived his equations by using general physical and geometric ideas. However, later, it turned out that the exact same equations can be obtained from the usual field theory in the usual non-curved ("flat") space-time if we consider a physical field $\gamma_{ij} = \sqrt{-g} \cdot g^{ij} - \eta^{ij}$, where:

– $\eta^{ij} = \text{diag}(1, -1, -1, -1)$ is a diagonal matrix with elements 1, -1, -1, and -1,
– g^{ij} is an "inverse" matrix to the metric tensor, i.e., a matrix for which $\sum_j g^{ij} \cdot g_{jk} = \delta_{ik}$, where δ_{ik} is the above-defined Kronecker (which is equal to 1 if $i = k$ and to 0 otherwise), and
– g is the determinant of the matrix g_{ij}.

To get Einstein's equations, we need to consider a field γ_{ij} whose source is the overall energy-stress tensor $E_{ij} = T_{ij} + t_{ij}(\{\gamma_{kl}\})$ that includes both:

– the energy of the external fields T_{ij} and
– the energy $t_{ij}(\{\gamma_{kl}\})$ of the field γ_{ij} itself:

$$\Box\gamma_{ij} = E_{ij} = T_{ij} + t_{ij}(\{\gamma_{kl}\}); \tag{37}$$

see, e.g., [10,16].

The observable effects of the metric field g_{ij} – i.e., in these terms, of the field γ_{ij} – depends only on the expression Γ – just like in electrodynamics, where the observable effect of the electromagnetic 4-potential A_i depends only on the electromagnetic field $F_{ij} = A_{i,j} - A_{j,i}$. Some changes to A_i do not affect the values F_{ij} and thus, lead to the same observable results. Because of this, we can always select the field A_i so as to satisfy an additional relation. In field theory, it turns out to be convenient to use the relation $A_{i,i} = 0$. Similarly, without changing the observable effect of the field γ_{ij}, it is always possible to select the metric tensor for which $\gamma_{ij,j} = 0$.

The geometric meaning of this choice is that we can always select coordinates in which this condition is satisfied; in general relativity, such conditions are known as *harmonic coordinates*. In these coordinates, we always have $(\Box\gamma_{ij})_{,j} = 0$ and thus, $E_{ij,j} = 0$.

In general relativity, energy depends on the observer. We may recall that Einstein's general relativity started with Einstein's observation that:

– while we usually strongly feel the law of gravity,
– a person in a freely falling elevator (Einstein's example) or on board of a spaceship orbiting the Earth (our natural example) feels absolute weightlessness.

So:

– from *our* viewpoint, there is a strong potential energy of the gravitational field – so strong that a small flower pot falling from the 10-th floor can easily kill a pedestrian;
– on the other hand, from the *astronaut's* viewpoint, there is no potential energy.

So, from the viewpoint of general relativity, gravitational energy is something depending on the coordinate system: it is the term $t_{ij}(\{\gamma_{kl}\})$.

What will happen if the energy of the external fields is not preserved. Normally, physicists consider situations in which the external fields T_{ij} are described by usual theories like electrodynamics, theories for which energy is preserved.

However no one prevents us from considering a general case, in which the energy T_{ij} of the external field is *not* preserved, i.e., for which $T_{ij,j} \neq 0$. In this case, since for (37) we always have $E_{ij,j} \neq 0$, this would mean that we will have $t_{ij,j} = -T_{ij,j}$. So, if we have a strong energy non-conservation, with a large non-zero value of $T_{ij,j}$, this will create a large value of $t_{ij,j}$ and thus, a large gravitational field t_{ij}.

The fact that $E_{ij,j} = 0$ means that the overall energy is preserved – so that non-conservation of energy of the other fields will generate a string gravitational field. The energy will not disappear – it will get transformed into the gravitational energy.

This phenomenon can be explained in the original general relativistic terms. The same phenomenon can be explained in the original Einstein's curved space-time.

Indeed, from the definition of the tensor G_{ij}, it follows that $G^{ij}_{;j} = 0$, where, for each object a (scaler or vector or tensor), $a_{;k}$ denotes *covariant derivative*, i.e., derivative of the type $a_{;k} = a_{,k} + \Gamma a$, where Γ is an expression containing g_{ij} and its first derivatives – and which is equal to 0 in non-curved (Minkowski) space-time. Due to Einstein's equation, the formula $G^{ij}_{;j} = 0$ implies that $T^{ij}_{;j} = 0$, i.e., that $T^{ij}_{j} + \Gamma T = 0$. If in the original theory, energy is not conserved, i.e., we have $T^{ij}_{,j} \neq 0$, this means that we have $\Gamma T \neq 0$, i.e., that $\Gamma \neq 0$.

The value $\Gamma = 0$ corresponds to non-curved space-time, so $\Gamma \neq 0$ means that the space-time is curved – i.e., that there is a gravitational field. The larger non-conservation of energy, the larger Γ and thus, the stronger the corresponding gravitational field. Thus, in the relativistic version of the original non-energy-conserving theory, energy does not disappear, it simply gets transformed into the gravitational energy – just like in Joule's experiments, mechanical energy and heat energy got transformed into each other.

Where are we in relation to Joule's principle: final conclusion. In the previous sections, we saw that if we only take quantum effects into account, then we will most probably have to abandon energy conservation ideas – and, what is even worse, we will not even be able to test energy conservation experimentally without the need to go beyond non-relativistic quantum physics.

As we saw in this section, if we do go beyond non-relativistic quantum physics – by considering effects of special and general relativity – then Joule's principle becomes valid again.

Not only Joule's principle becomes valid again – this principle is strengthened. It is no longer a principle that all physical theories *should* satisfy, it is a principle that all physical theories *do* satisfy – at least in the current picture of the physical world, a picture governed by quantum physics and general relativity.

Conclusion. Our main conclusion is that conservation of energy is with us always, with or without free will included. In the previous sections, we show that this conclusion is *not* valid for classical physics or for non-relativistic quantum theories, but it *is* valid at the all-inclusive level, when we take into account both quantum and relativistic effects.

Acknowledgements We would like to dedicate to Yakov Kreinovich, late Vladik's father and Olga's father-in-law, whose ideas were one of the main triggers for this research.

We are thankful to the referees for their very useful suggestions and very interesting references and ideas.

Conflict of Interest Statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Compatibilism, Stanford Encyclopedia of Philosophy, https://plato.stanford.edu, downloaded on February 8, 2020
2. F. Crick, The Astonishing Hypothesis: The Scientific Search for the Soul, Charles Scribner & Sons, New York (1993)
3. P. Davies, God and the New Physics, Simon & Schuster, New York (1983)
4. D. Dennett, Elbow Room, MIT Press, Cambridge, Massachusetts (1984)
5. D. Dennett, Freedom Evolves, Viking Penguin, New York (2003)
6. J. Earman, A Primer on Determinism, D. Reidel, Dordrecht, Holland (1986)
7. R. Feynman, R. Leighton, and M. Sands, The Feynman Lectures on Physics, Addison Wesley, Boston, Massachusetts (2005)
8. J. M. Fischer, God, Foreknowledge and Freedom, Stanford University Press, Stanford, California (1989)
9. H. Frankfurt, Alternate possibilities and moral responsibility, Journal of Philosophy, Vol. 66, pp. 829–839 (1969)
10. S. N. Gupta, Einstein's and other theories of gravitation, Review of Modern Physics, Vol. 29, pp. 334–336 (1957)
11. J. Hospers, What means this freedom?, In: S. Hook (ed.), Determinism and Freedom in the age of modern science, Collier-Macmillan, New York, pp. 126–142 (1958)
12. R. Kane, The Oxford Handbook of Free Will, Oxford University Press, New York (2002)
13. J. P. Joule, On the calorific effects of magneto-electricity, and on the mechanical value of heat, Philosophical Magazine, Vol. 3, No. 23, pp. 263–276, 347–355, 435–443 (1843)
14. J. P. Joule, On the changes of temperature produced by the rarefaction and condensation of air, Philosophical Magazine, Vol. 3, No. 26 (174), pp. 369–383 (1845)
15. J. P. Joule, Joint Scientific Papers of James Prescott Joule, Taylor and Francis, London, 1887.
16. V. Kreinovich, Gupta's derivation of Einstein equations, Soviet Academy of Sciences Doklady, Vol. 222, No. 2, pp. 319–321 (1975, in Russian); translated into English as Soviet Physics Doklady, Vol. 20, No. 5, pp. 341–342 (1975)
17. V. Kreinovich, In quantum physics, free will leads to non-conservation of energy, Journal of Uncertain Systems, Vol. 7, No. 3, pp. 176–178 (2013)
18. V. Kreinovich and O. Kosheleva, Is Lagrangian formalism adequately describing energy conservation?, Mathematical Structures and Modelling, Vol. 28, No. 2, pp. 21–27 (2013)
19. V. Kreinovich and A. Ortiz, If energy is not preserved, then Planck's constant is no longer a constant: a theorem, Mathematical Structures and Modeling, Vol. 26, pp. 57–63 (2012)
20. L. D. Landau and E. M. Lifschitz, The Classical Theory of Fields, The Classical Theory of Fields, Butterworth-Heinemann, Oxford, UK (1980)
21. C. W. Misner, K. S. Thorne, and J. A. Wheeler, Gravitation, W. H. Freeman, New York (1973)
22. R. Penrose, A. Shimony, N. Cartwright, S. Hawking, and M. Longair, The Large, the Small and the Human Mind, Cambridge University Press, Cambridge, Massachusetts (2000)

23. T. E. Phipps Jr., Time asymmetry and quantum equations of motion, Foundations of Physics, Vol. 3, pp. 435–455 (1973)
24. S. Smilansky, Free Will and Illusion, Oxford University Press, New York (2000)
25. E. Sober, Simplicity, Claredon Press, Oxford, UK (1975)
26. P. Suppes, The Transcendental character of determinism, Midwest Studies in Philosophy, Vol. 18, pp. 242–257 (1993)
27. L. Susskind and G. Hrabovsky, The Theoretical Minimum: What You Need to Know to Start Doing Physics, Basic Books, New York (2014)
28. F. J. Tipler, The Physics of Immortality: Modern Cosmology, God and the Resurrection of the Dead, Doubleday, New York (1994)
29. K. S. Thorne and R. D. Blandford, Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics, Princeton University Press, Princeton, New Jersey (2017)
30. J. Trusted, Free Will and Responsibility, Oxford University Press, New York (1984)
31. J. C. Urenda and O. Kosheleva, How to reconcile physical theories with the idea of free will: from analysis of a simple model to interval and fuzzy approaches, Proceedings of the IEEE World Congress on Computational Intelligence WCCI'2008, Hong Kong, China, June 1–6, 2008, pp. 1024–1029 (2008)
32. P. Van Inwagen, An Essay on Free Will, Clarendon Press, Oxford (1983)