Do the Laws of Nature Have Necessity? If Yes, Where Does It Come From?

Shan Gao
Research Center for Philosophy of Science and Technology,
Shanxi University, Taiyuan 030006, P. R. China
E-mail: gaoshan2017@sxu.edu.cn.

August 1, 2024

Abstract

In this paper, I present a new analysis of the origin of laws of nature. It is argued that the quantum mechanical equations of motion for free particles in Minkowski spacetime such as the Klein-Gordon equation and its nonrelativistic limit (i.e. the free Schrödinger equation) are determined by the properties of such spacetime. This result strongly supports necessitarianism and may also help us understand where the physical necessity of the laws of nature comes from.

1 Introduction

The nature of the laws of nature has been a hot topic of debate in philosophy of science. Recent years have witnessed a growing interest in this long-standing question (Ott and Patton, 2018; Carroll, 2020; Adlam, 2022; Ott, 2022; Meacham, 2023; Chen, 2024). Exactly what is a law of nature? And where do the laws of nature come from? There are two competing views in metaphysics: the regularity view and necessitarianism.[1] The former says that laws of nature are only statements of the regularities in the world or descriptions of the way the world is, and physical necessity and the explanation it provides are an illusion. By contrast, the latter claims that the laws of nature must have physical necessity, which is needed to explain why the world is as it is, e.g. the occurring event had to happen given the laws of nature and antecedent conditions. However, it is still unclear where

[1]Note that the regularity view is widely called the Humean account of laws of nature. But this is a misnomer, since Hume himself was not Humean but a necessitarian as regards laws of nature (Swartz, 2024).

the necessity of the laws of nature comes from. According to primitivism about laws, the laws of nature are metaphysically fundamental, and its necessity is just its nature and it has no further origin (Maudlin, 2007). In this paper, I will present a new analysis of the origin of laws of nature. My approach is to investigate certain concrete laws of physics to see whether they have physical necessity. This will avoid pure philosophical speculation. In particular, I will show that the quantum mechanical equations of motion for free particles in Minkowski spacetime such as the Klein-Gordon equation and its nonrelativistic limit (i.e. the free Schrödinger equation) can be derived from the properties of such spacetime. This result, as I will argue, strongly supports necessitarianism and may also help us understand where the physical necessity of the laws of nature comes from.

2 A derivation of quantum wave equations for free particles in Minkowski spacetime

In Minkowski spacetime, the laws of motion for isolated systems satisfy spacetime translation invariance and relativisitic invariance. In the following, I will argue that the quantum wave equations for free particles in Minkowski spacetime such as the Klein-Gordon equation and its nonrelativistic limit (i.e. the free Schrödinger equation) can be derived from these two invariance requirements when assuming linearity of time evolution (see also Gao, 2017). For simplicity, I consider only free spinless particles, whose states are represented by a scalar function with respect to both x and t, $\psi(x,t)$.[2]

A space translation operator can be defined as

$$T(a)\psi(x,t) = \psi(x-a,t). \tag{1}$$

It means translating rigidly the state of the system, $\psi(x,t)$, by an infinitesimal amount a in the positive x direction. $T(a)$ can be further expressed as

$$T(a) = e^{-iaP}, \tag{2}$$

where P is the generator of space translation.[3] By expanding $\psi(x-a,t)$ in order of a, we can further get

$$P = -i\frac{\partial}{\partial x}. \tag{3}$$

[2] For particles with spin, their states will be represented by a more complex funtion such as a vector function.

[3] I introduce the imaginary unit i in the expression so that it is consistent with the standard definition.

Similarly, a time translation operator can be defined as

$$U(t)\psi(x,0)=\psi(x,t). \tag{4}$$

Let the evolution equation of state be the following form:

$$i\frac{\partial\psi(x,t)}{\partial t}=H\psi(x,t), \tag{5}$$

where H is a to-be-determined operator that depends on the properties of the studied system. Then the time translation operator $U(t)$ can be expressed as $U(t)=e^{-itH}$, and H is the generator of time translation.[4] In the following analysis, I assume H is a linear operator independent of the evolved state, namely the evolution is linear, which is a key feature of the quantum wave equations.

Let's now analyze the implications of spacetime translation invariance for the laws of motion. First of all, time translational invariance requires that H have no time dependence, namely $dH/dt=0$. This can be demonstrated as follows (see also Shankar, 1994, p.295). Suppose an isolated system is in state ψ_0 at time t_1 and evolves for an infinitesimal time δt. The state of the system at time $t_1+\delta t$, to first order in δt, will be

$$\psi(x,t_1+\delta t)=[I-i\delta tH(t_1)]\psi_0. \tag{6}$$

If the evolution is repeated at time t_2, beginning with the same initial state, the state at $t_2+\delta t$ will be

$$\psi(x,t_2+\delta t)=[I-i\delta tH(t_2)]\psi_0. \tag{7}$$

Time translational invariance requires the outcome state should be the same:

$$\psi(x,t_2+\delta t)-\psi(x,t_1+\delta t)=i\delta t[H(t_1)-H(t_2)]\psi_0=0. \tag{8}$$

Since the initial state ψ_0 is arbitrary, it follows that $H(t_1)=H(t_2)$. Moreover, since t_1 and t_2 are also arbitrary, it follows that H is time-independent, namely $dH/dt=0$.

Secondly, space translational invariance requires $[T(a),U(t)]=0$, which further leads to $[P,H]=0$. This can be demonstrated as follows (see also Shankar, 1994, p.293). Suppose at $t=0$ two observers A and B prepare identical isolated systems at $x=0$ and $x=a$, respectively. Let $\psi(x,0)$ be the state of the system prepared by A. Then $T(a)\psi(x,0)$ is the state of the system prepared by B, which is obtained by translating (without distortion) the state $\psi(x,0)$ by an amount a to the right. The two systems look identical to the observers who prepared them. After time t, the states evolve into $U(t)\psi(x,0)$ and $U(t)T(a)\psi(x,0)$. Since the time evolution of each identical

[4] Similarly I also introduce the imaginary unit i in the equation of state for convenience of later analysis.

system at different places should appear the same to the local observers, the above two systems, which differed only by a spatial translation at $t = 0$, should differ only by the same spatial translation at future times. Thus the state $U(t)T(a)\psi(x,0)$ should be the translated version of A's system at time t, namely we have $U(t)T(a)\psi(x,0) = T(a)U(t)\psi(x,0)$. This relation holds true for any initial state $\psi(x,0)$, and thus we have $[T(a), U(t)] = 0$, which says that space translation operator and time translation operator are commutative.

When $dH/dt = 0$, the solutions of the evolution equation Eq.(5) assume the basic form

$$\psi(x,t) = \varphi_E(x)e^{-iEt}, \tag{9}$$

and their linear superpositions, where E is an eigenvalue of H, and $\varphi_E(x)$ is an eigenfunction of H and satisfies the time-independent equation:

$$H\varphi_E(x) = E\varphi_E(x). \tag{10}$$

Moreover, the commutative relation $[P, H] = 0$ further implies that P and H have common eigenfunctions. Since the eigenfunction of $P = -i\frac{\partial}{\partial x}$ is e^{ipx} (except a normalization factor), where p is the eigenvalue, the basic solutions of the evolution equation Eq.(5) for an isolated system assume the form $e^{i(px-Et)}$, which represents the state of an isolated system with definite properties p and E. In quantum mechanics, P and H, the generators of space translation and time translation, are called momentum operator and energy operator (or the Hamiltonian of the system), respectively. Correspondingly, $e^{i(px-Et)}$ is the eigenstate of both momentum and energy, and p and E are the corresponding momentum and energy eigenvalues, respectively. Then the state $e^{i(px-Et)}$ describes an isolated system (e.g. a free particle) with definite momentum p and energy E in Minkowski spacetime. Note that since the Hamiltonian and the generator of space translation are both Hermitian, the eigenvalues of energy and momentum are both real, and thus the state $e^{i(px-Et)}$ is a complex function.

The energy-momentum relation can be further determined by considering the relativistic structure of Minkowski spacetime. The energy-momentum operator $P_\mu = i(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla)$ is a four-vector operator. The operator $P_\mu P^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$, also called the d'Alembert operator, is a Lorentz scalar operator. This means that for the eigenvalues of energy and momentum there is the following relation:

$$E^2 - p^2c^2 = E_0^2, \tag{11}$$

where E_0^2 is a Lorentz scalar. It can be seen that E_0 is the energy of the particle when its momentum is zero, usually called the rest energy of the particle. By defining $m = E_0/c^2$ as the (rest) mass of the particle, we can further obtain the familiar energy-momentum relation

$$E^2 = p^2c^2 + m^2c^4. \tag{12}$$

Since the operators H and P have common eigenfunctions for an isolated system, the relation between their eigenvalues E and p or the energy-momentum relation implies the corresponding operator relation between H and P, which is $H = \sqrt{P^2c^2 + m^2c^4}$ or $H^2 = P^2c^2 + m^2c^4$ for an isolated system. Then we can obtain the Klein-Gordon equation:[5]

$$[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} + m^2c^2]\psi(x,t) = 0. \tag{13}$$

In the nonrelativistic limit (where $p \ll mc$ and the rest energy is absorbed into the Hamiltonian), the operator relation becomes $H = P^2/2m$ for an isolated system. Then we can obtain the free Schrödinger equation:

$$i\frac{\partial\psi(x,t)}{\partial t} = -\frac{1}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}. \tag{14}$$

Note that the eigenvalues of energy and momentum are both real, and thus the mass parameter m in the above equation also assumes real values due to the energy-momentum relation $E = p^2/2m$. This ensures that the appearance of the imaginative unit i in the equation is not apparent and the equation is indeed the free Schrödinger equation.

In addition, it is worth noting that the reduced Planck constant $\hbar$ with dimension of action is missing in the above Klein-Gordon equation and free Schrödinger equation. However, this is not a problem. The reason is that the dimension of $\hbar$ can be absorbed in the dimension of m. For example, we can stipulate the dimensional relations as $p = 1/L$, $E = 1/T$ and $m = T/L^2$, where L and T represent the dimensions of space and time, respectively (see Duff, Okun and Veneziano, 2002 for a more detailed analysis). Moreover, the value of $\hbar$ can be set to the unit of number 1 in principle. Thus the above equations are just the Klein-Gordon equation and free Schrödinger equation in quantum mechanics.

A final point. It is usually thought that by Noether's theorem symmetries yield conservation laws, e.g. spacetime translation invariance of laws of motion implies the laws of conservation of energy and momentum, which says that the total energy and total momentum of an isolated system are always constant. However, since the proof of Noether's theorem assumes the laws of motion (in both the classical and quantum cases, e.g. it assumes the Euler-Lagrange equation in the classical case), one cannot really derive conservation laws from symmetries without knowing the laws of motion. By

[5] When the state of a particle is represented by a more complex vector function, we can also obtain the Dirac equation and other wave equations for particles with spin.

contrast, the above derivation of the wave equations in quantum mechanics also provides a genuine derivation of the laws of conservation of energy and momentum for isolated systems.

3 Further discussions about the derivation

I have derived the quantum mechanical equations of motion for free spinless particles in Minkowski spacetime, namely the Klein-Gordon equation and the free Schrödinger equation (in the nonrelativistic limit). The derivation may help us understand the physical origin of these wave equations, which are usually derived in quantum mechanics textbooks by analogy and correspondence with classical physics. There are at least two mysteries in the textbook derivation. First of all, even if the behavior of microscopic particles likes wave and thus a wave function is needed to describe them, it is unclear why the wave function must assume a complex form. Indeed, when Schrödinger invented his equation, he was puzzled by the inevitable appearance of the imaginary unit "i" in his equation. Next, one doesn't know why there are the de Broglie relations for momentum and energy and why the energy-momentum relation is as it is.

According to the above analysis, the key to unveiling these mysteries is to analyze spacetime translation invariance of laws of motion. Spacetime translation gives the definitions of momentum and energy in quantum mechanics. The momentum operator P is defined as the generator of space translation, and it is Hermitian and its eigenvalues are real. Moreover, the form of the momentum operator is uniquely determined by its definition, which turns out to be $P = -i\partial/\partial x$, and its eigenfunctions are e^{ipx}, where p is the corresponding real eigenvalue. Similarly, the energy operator or the Hamiltonian of the system H is the generator of time translation, and its concrete form is determined by the concrete properties of the system.

Fortunately, for an isolated system, the form of H, which determines the evolution equation of state, can be fixed for linear time evolution by the requirements of spacetime translation invariance and relativistic invariance. Concretely speaking, time translational invariance requires that $dH/dt = 0$, and this implies that the solutions of the evolution equation $i\partial\psi(x,t)/\partial t = H\psi(x,t)$ are $\varphi_E(x)e^{-iEt}$ and their superpositions, where $\varphi_E(x)$ is the eigenfunction of H. Moreover, space translational invariance requires $[P, H] = 0$. This means that P and H have common eigenfunctions, and thus $\varphi_E(x) = e^{ipx}$. Therefore, $e^{i(px-Et)}$ and their superpositions are solutions of the evolution equation for an isolated system, where $e^{i(px-Et)}$ represents the state of the system with momentum p and energy E. In other words, the state of an isolated system (e.g. a free particle) with definite momentum and energy assumes the plane wave form $e^{i(px-Et)}$. Furthermore, the relation between p and E or the energy-momentum relation can be determined by considering the relativistic transformation of the generators of space translation and time translation, and in the nonrelativistic limit it is $E = p^2/2m$. Then we can obtain the Hamiltonian of an isolated system, $H = P^2/2m$, and the free Schrödinger equation, Eq.(14), in the nonrelativistic limit.

To summarize, the above derivation tells us that the quantum mechanical equations of motion for free spinless particles in Minkowski spacetime, namely the Klein-Gordon equation and the free Schrödinger equation (in the nonrelativistic limit), are determined by the structure and properties of such spacetime. Spacetime translation invariance, which is used to obtain the plane wave representation of the state of a free particle with definite momentum and energy, is a consequence of the homogeneity of spacetime. Minkowski spacetime is homogeneous. The homogeneity of spacetime ensures that the laws of motion are the same in two different places and at two different times. Moreover, the energy-momentum relation is also entailed by the structure of Minkowski spacetime, e.g. the dot product of two four-vectors is a Lorentz scalar that is Lorentz invariant.

4 Possible implications for the nature of laws of nature

Then, what are the implications of the above analysis for the nature of laws of nature? I think there are at least three possible implications.

First of all, the analysis may help solve the ontological issues about laws of nature and answer what kind of things laws are, e.g. it supports necessitarianism and disfavors the regularity view. According to the analysis, the Klein-Gordon equation, being the law of motion for free spinless particles in Minkowski spacetime, is determined by the structure and properties of the spacetime in which these particles exist and evolve. In other words, the equation of motion for these particles in Minkowski spacetime must assume the form of the Klein-Gordon equation, and thus it has physical necessity. As a result, the free Schrödinger equation, being the nonrelativistic limit of the Klein-Gordon equation, also has physical necessity. If the state of a free particle (whose speed is much less than the speed of light) is $\psi(x, t_0)$ at instant t_0, its state must be $\psi(x, t)$ at another instant t, where $\psi(x, t)$ is obtained from $\psi(x, t_0)$ by solving the free Schrödinger equation.[6] This means that for the Humean mosaic, which is a 4-dimensional Minkowski spacetime occupied by free particles, its states at two different instants have a necessary connection. If this is not the case, then Minkowski spacetime cannot have all properties that it should have by definition including homogeneity,

[6]Note that my analysis does not imply that laws *produce* the subsequent states from earlier ones (cf. Maudlin, 2007); rather, they only establish the necessary connection between the states of the world at different instants.

which is a logical contradition.

Here it is worth noting that this implication for necessitarianism does not depend on whether the law of motion for interacting particles or the final theory of quantum gravity can be derived in a similar way. The reason is that on the regularity view, the nonrelativistic law of motion for free particles whose speed is much less than the speed of light in Minkowski spacetime (without gravity) does not have physical necessity within its domain either, but this contradicts the above analysis.

In addition, by the same reasoning, it can be argued that the above analysis seems to also disfavor primitivism about laws. According to this view, the laws of nature are metaphysically fundamental, and their physical necessity has no further origin or explanation. Now if the laws of nature include the Klein-Gordon equation and the free Schrödinger equation in quantum mechanics as usually thought, then since the above derivation of these equations explains their physical necessity, primitivism about laws is disfavored. What if the laws of nature include only the final complete laws? (which seems to be a minority view) In this case, the above wave equations in quantum mechanics, being reduced forms of the final complete laws holding within a certain domain, also have physical necessity that has no further explanation (besides that they can be derived from the final complete laws) on primitivism about laws, and thus the above derivation also disfavors this view.

Second, the above analysis may further help solve the epistemological issues about laws of nature and answer how we have epistemic access to laws. An obvious epistemological issue is the so-called the epistemic gap: if laws are really objective and mind-independent as widely thought, it seems puzzling how we can have epistemic access to them, since laws are not consequences of our observations (Chen, 2024). The above analysis provides a possible way to close the epistemic gap for laws of nature. It is that we may have epistemic access to the laws holding within a certain domain by deriving them with the help of mathematics and logic. For example, we can obtain the equations of motion for free particles in Minkowski spacetime by deriving them from the properties of such spacetime.

Finally, the above analysis may also help understand the criteria of laws of nature such as objectivity, universality and simplicity etc. If the laws holding within a certain domain can be derived, then the criteria of laws will not be useful tools for us to discover or have epistemic access to the laws, but be the derived features of laws as a result of the derivation. This will settle the controversies about whether these features should be regarded as the criteria of laws.

5 Conclusions

In order to explain the world, e.g. why the world is as it is, it seems that the laws of nature must have a kind of necessity. However, it has been a deep mystery for necessitarians that where the physical necessity of laws comes from.[7] In this paper, I have argued that the quantum mechanical equations of motion for free spinless particles in Minkowski spacetime, namely the Klein-Gordon equation and the free Schrödinger equation (in the non-relativistic limit), are determined or necessiated by the properties of the spacetime in which these particles exist and evolve. This result strongly supports necessitarianism and also suggests that the laws of nature are not unanalyzable facts about the world.

How about the final law of motion for particles and spacetime? It will be a unified equation of quantum gravity, which is still unknown to us. Since there is nothing besides these particles and spacetime in the world, it seems that the law cannot be derived in a similar way as in this paper; the law for something cannot be determined by this thing itself. My conjecture is that in this case, the complete state of the universe including all particles and spacetime is uniquely determined by the boundary condition. In the end, the world should be self-explainable, and we need not resort to something outside the world to explain why the world is as it is.

References

[1] Adlam, E. (2022). Laws of nature as constraints. Foundations of Physics, 52(1), 1–41.

[2] Carroll, J. W. (2020). Laws of Nature. The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2020/entries/laws-of-nature/.

[3] Chen, E. K. (2024). Laws of Physics (Elements in the Philosophy of Physics). Cambridge: Cambridge University Press.

[4] Gao, S. (2017). The Meaning of the Wave Function: In Search of the Ontology of Quantum Mechanics. Cambridge: Cambridge University Press.

[5] Maudlin, T. (2007). The Metaphysics Within Physics. Oxford: Oxford University Press.

[7] "Twentieth-century Necessitarianism has dropped God from its picture of the world. Physical necessity has assumed God's role: the universe conforms to (the dictates of? / the secret, hidden, force of? / the inexplicable mystical power of?) physical laws. God does not 'drive' the universe; physical laws do. But how? How could such a thing be possible? The very posit lies beyond (far beyond) the ability of science to uncover." (Swartz, 2024)

[6] Meacham, C. J. G. (2023). The nomic likelihood account of laws. Ergo: An Open Access Journal of Philosophy, 9(9), 230–284.

[7] Ott, W. (2022). The Metaphysics of Laws of Nature: The Rules of the Game. Oxford: Oxford University Press.

[8] Ott, W and Patton, L (eds) (2018). Laws of Nature. Oxford: Oxford University Press.

[9] Shankar, R. (1994). Principles of Quantum Mechanics, 2nd ed. New York: Plenum.

[10] Swartz, N. (2024) Laws of Nature. The Internet Encyclopedia of Philosophy. https://iep.utm.edu/lawofnat/.