International Journal of Quantum Foundations 7 (2021) 13-20

Original Paper

Derivation of the free wave equations in quantum mechanics

Shan Gao

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

Received: 9 October 2020 / Accepted: 16 December 2020 / Published: 26 January 2021

Abstract: We present a simple and rigorous derivation of the free wave equations such as the Klein-Gordon equation based on spacetime translation invariance and relativistic invariance. The new analysis may help underdstand the physical origin and significance of the laws of motion in quantum mechanics.

Keywords: Quantum mechanics; wave equations; Klein-Gordon equation; spacetime translation invariance; relativistic invariance

1. Introduction

Quantum mechanics textbooks often provide a heuristic derivation of the free wave equations such as the free Schrödinger equation (see, e.g. [1-3]). It begins with the assumption that the state of a free microscopic particle has the form of a plane wave $e^{i(kx-\omega t)}$. When combining with the de Broglie relations for momentum and energy $p = \hbar k$ and $E = \hbar\omega$, this state becomes $e^{i(px-Et)/\hbar}$. Then it uses the energy-momentum relation to obtain the free wave equations, e.g. using the non-relativistic energy-momentum relation $E = p^2/2m$ to obtain the free Schrödinger equation.

In this paper, we will show that the heuristic derivation of the free wave equations can be made more rigorous by resorting to spacetime translation invariance and relativistic invariance. Spacetime translation gives the definitions of momentum and energy, and spacetime translation invariance entails that the state of a free quantum system with definite momentum and energy assumes the plane wave form $e^{i(px-Et)}$. Moreover, the relativistic transformations of the generators of space translation and time translation further determine the relativistic energy-momentum relation, whose nonrelativistic approximation is $E = p^2/2m$. Although the requirements of these invariances are already well known, an explicit and complete derivation of the free wave equations using them seems still missing in the literature. The new analysis may help underdstand the physical origin and significance of the laws of motion in quantum mechanics.

2. Spacetime translation invariance

It is a fundamental postulate in physics that the laws of motion that govern the time evolution of an isolated system satisfies spacetime translation invariance. This is due to the homogeneity of space and time. The homogeneity of space ensures that the same experiment performed at two different places gives the same result, and the homogeneity in time ensures that the same experiment repeated at two different times gives the same result. In this section, we will analyze how the requirement of spacetime translation invariance restricts the possible forms of the laws of motion. For the sake of simplicity, we will mainly analyze one-dimensional motion.

The physical state of an isolated system is assumed to be represented by a general analytic function with respect to both x and t, $\psi(x,t)$.[1] 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.[2] $T(a)$ can be further expressed as

$$T(a) = e^{-ia\hat{p}}, \tag{2}$$

---

[1] It is arguable that $\psi(x,t)$ is the most general scalar representation of the spatial state of a system. As we will see later, however, the equation that governs the time evolution of the state will restrict the possible forms of $\psi(x,t)$.

[2] There are in general two different pictures of translation: active transformation and passive transformation. The active transformation corresponds to displacing the studied system, and the passive transformation corresponds to moving the coordinate system. Physically, the equivalence of the active and passive pictures is due to the fact that moving the system one way is equivalent to moving the coordinate system the other way by an equal amount. Here we will analyze spacetime translations in terms of active transformations.

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

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

Similarly, a time translation operator can be defined as

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

And it can also be expressed as $U(t) = e^{-it\hat{E}}$, where

$$\hat{E} = i\frac{\partial}{\partial t} \tag{5}$$

is the generator of time translation. In order to know the laws of motion, we need to find the concrete manifestation of $\hat{E}$ for a physical system, which means that we need to find the evolution equation of state:

$$i\frac{\partial\psi(x,t)}{\partial t} = H\psi(x,t), \tag{6}$$

where H is a to-be-determined operator that depends on the properties of the studied system, and it is also called the generator of time translation.[4] In the following analysis, we assume H is a linear operator independent of the evolved state, namely the evolution is linear, which is an important presupposition in our derivation of the free wave equations.

Let us now see 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 [4], 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

$$\psi(x, t_1 + \delta t) = [I - i\delta t H(t_1)]\psi_0. \tag{7}$$

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

---

3 In order to differentiate the momentum and energy eigenvalues from the momentum and energy operators, we add a hat to the momentum and energy operators as usual. But we omit the hat for all other operators in this book. In addition, for convenience of later discussion we introduce the imaginary unit i in the expression. This does not influence the validity of the following derivation.

4 Similarly we also introduce the imaginary unit i in the equation for convenience of later discussion.

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

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

$$\psi(x, t_2 + \delta t) - \psi(x, t_1 + \delta t) = i\delta t[H(t_1) - H(t_2)]\psi_0 = 0. \tag{9}$$

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 $dH/dt = 0$. It can be seen that this result relies on the linearity of time evolution. If H depends on the state, then obviously we cannot obtain $dH/dt = 0$ because the state is time-dependent. But we still have $H(t_1, \psi_0) = H(t_2, \psi_0)$, which means that the state-dependent H also satisfies time translational invariance.

Secondly, space translational invariance requires $[T(a), U(t)] = 0$, which further leads to $[\hat{p}, \hat{E}] = 0$ and $[\hat{p}, H] = 0$. This can be demonstrated as follows (see also [4], 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 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. Again, it can be seen that the linearity of time evolution is an important presupposition of this result. If $U(t)$ depends on the state, then the space translational invariance will only lead to $U(t, T\psi)T(a)\psi(x, 0) = T(a)U(t, \psi)\psi(x, 0)$, from which we cannot obtain $[T(a), U(t)] = 0$.

When $dH/dt = 0$, the solutions of the evolution equation Eq.(6) assume the basic form

$$\psi(x, t) = \varphi_E(x)e^{-iEt}, \tag{10}$$

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:

$$H\varphi_E(x) = E\varphi_E(x). \tag{11}$$

Moreover, the commutative relation $[\hat{p}, H] = 0$ further implies that $\hat{p}$ and H have common eigenfunctions. Since the eigenfunction of $\hat{p} = -i\frac{\partial}{\partial x}$ is e^{ipx} (except a normalization factor), where p is the eigenvalue, the basic solutions of the evolution equation Eq.(6) 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, $\hat{p}$ and $\hat{E}$, the generators of space translation and time translation, are also called momentum operator and energy operator, respectively, and H is called the Hamiltonian of the system. 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 electron) with definite momentum p and energy E.

3. The energy-momentum relation

The energy-momentum relation can be further determined by considering the relativistic transformations of the generators of space translation and time translation. The operator $\hat{P}_\mu = (\hat{E}/c, -\hat{p}) = i(\frac{1}{c}\frac{\partial}{\partial t}, \nabla)$ is a four-vector operator. In order that its eigenvalue equation holds in all inertial frames, its eigenvalues must transform as a four-vector too. In other words, every eigenvalue of the four-vector operator $\hat{P}_\mu$, $(E/c, -p)$, is also a four-vector. Since the dot product of two four-vectors is Lorentz invariant (a Lorentz scalar), we can form a Lorentz scalar $p^2 - E^2/c^2$ with the four-vector $(E/c, -p)$. Then the energy-momentum relationm is:

$$E^2 = p^2c^2 + E_0^2, \tag{12}$$

where p and E are the momentum and energy of a microscopic particle, respectively, and E_0 is the energy of the particle when its momentum is zero, 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

$$E^2 = p^2c^2 + m^2c^4. \tag{13}$$

In the nonrelativistic domain, this energy-momentum relation reduces to $E = p^2/2m$.

4. Derivation of the free wave equations

Since the operators $\hat{E}$ and $\hat{p}$ 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 $\hat{E}$ and $\hat{p}$. In the relativistic domain, the operator relation is $\hat{E}^2 = \hat{p}^2c^2 + m^2c^4$ for an isolated system. Then we can obtain the Klein-Gordon equation:

$$\frac{\partial^2\psi(x,t)}{\partial x^2} - \frac{\partial^2\psi(x,t)}{c^2\partial t^2} - m^2c^2\psi(x,t) = 0. \tag{14}$$

In the nonrelativistic domain, the operator relation is $\hat{E} = \hat{p}^2/2m$ for an isolated system. Then we can obtain the free Schrödinger equation:

$$i\frac{\partial\psi(x,t)}{\partial t} = -\frac{1}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}. \tag{15}$$

Here it needs to be justified that the only parameter m in Eq. (15) assumes real values; otherwise the existence of the imaginative unit i in the equation will be an illusion and the equation will be distinct from the free Schrödinger equation. Since velocity assumes real values, this is equivalent to proving that momentum or the eigenvalue of the generator of space translation assumes real values, namely that the generator of space translation itself is Hermitian. This is indeed the case. Since the space translation operator $T(a)$ preserves the norm of the state: $\int_{-\infty}^{\infty}\psi^*(x,t)\psi(x,t)dx = \int_{-\infty}^{\infty}\psi^*(x-a,t)\psi(x-a,t)dx$, $T(a)$ is unitary, satisfying $T^{\dagger}(a)T(a) = I$. Thus the generator of space translation, $\hat{p}$, which is defined by $T(a) = e^{-ia\hat{p}}$, is Hermitian.

In addition, it is worth noting that the reduced Planck constant $\hbar$ with dimension of action is missing in the above free wave equations. However, this is in fact 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 [5] 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 essentially the free wave equations in quantum mechanics.

5. Further discussion

We have derived the free wave equations in quantum mechanics based on an analysis of spacetime translation invariance and relativistic invariance. The new analysis may not only make these equations more logical and understandable, but also help understand the origin of the complex wave function.

As noted before, the free Schrödinger equation is usually derived in quantum mechanics textbooks by analogy and correspondence with classical physics. There are at least two mysteries in this heuristic 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 also puzzled bythe inevitable appearance of the imaginary unit "*i*" in the equation [6]. Next, one doesn't know why there are the de Broglie relations for momentum and energy and why the nonrelativistic energy-momentum relation is $E = p^2/2m$.

According to the analysis given in the previous sections, the key to answer these questions 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 $\hat{p}$ 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 $\hat{p} = -i\partial/\partial x$, and its eigenfunctions are e^{ipx} , where p is the corresponding real eigenvalue. Similarly, the energy operator $\hat{E}$ is defined as the generator of time translation, and its universal form is $\hat{E} = i\partial/\partial t$. But the concrete manifestation of this operator for a physical system, denoted by H and called the Hamiltonian of the system, is determined by the concrete situation.

Fortunately, for an isolated system, the form of H, which determines the evolution equation of state, can be fixed for linear 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 $[\hat{p}, H] = 0$. This means that $\hat{p}$ 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 electron) 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 domain it is $E = p^2/2m$. Then we can obtain the Hamiltonian of an isolated system, $H = \hat{p}^2/2m$, and the free Schrödinger equation, Eq.(15).

Finally, we emphasize again that the linearity of time evolution is an important presupposition in the above derivation of the free Schrödinger equation. It is only for linear evolution that spacetime translation invariance of laws of motion can help determine the precise form of the equation of motion for isolated systems. It might be possible that the free evolution equation also contains nonlinear evolution terms. However, although nonlinear time evolution may also satisfy spacetime translation invariance, the invariance requirement cannot help determine the precise form of the nonlinear evolution equation. Nonlinear time evolution, if it exists, must have an additional physical origin.

References

[1] Schiff, L. (1968). Quantum Mechanics. New York: McGraw-Hill.

[2] Landau, L. and Lifshitz, E. (1977). Quantum Mechanics. Oxford: Pergamon Press.

[3] Greiner, W. (1994). Quantum Mechanics: An Introduction. New York: Springer.

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

[5] Duff, M. J., Okun, L. B., and Veneziano, G. (2002). Trialogue on the number of fundamental constants. Journal of High Energy Physics, 03, 023.

[6] Karam, R. Schrödinger's original struggles with a complex wave function. American Journal of Physics 88, 433 (2020).