
A Physical Explanation on Why Our Space Is Three Dimensional

Hua Ma

The College of Science, Air Force University of Engineering, Xi'an, People's Republic of China

Email address:

mahuar@xjtu.edu.cn

To cite this article:

Hua Ma. A Physical Explanation on Why Our Space Is Three Dimensional. *American Journal of Modern Physics*.

Vol. 6, No. 6, 2017, pp. 122-126. doi: 10.11648/j.ajmp.20170606.12

Received: August 12, 2017; **Accepted:** August 28, 2017; **Published:** September 21, 2017

Abstract: It is a basic, ancient and mysterious issue: why our space is three dimensional? This issue is related to philosophy, mathematics, physics and even religion, and thus aroused great research interests. The author makes an in-depth analysis of the problem, and finally comes to a conclusion: For any vector space with symmetry, orthogonality, homogeneity and completeness, the space dimension must be three on condition that: the energy obeys the law of conservation, the dynamics law is governed by the covariance principle, and thus the cross-product must can be defined in the space. Our space just meets and requires the above constraints, so its dimension is three.

Keywords: Space-time, Dimension, Energy Conservation Law, Covariance Principle, Field Tensor, Cross-product

1. Introduction

In mathematics, space-time is a manifold, namely, it is a topological space similar to the Euclidean space near each point. In physics, space-time is mathematical models that combine space and time into a coherent continuum. From history to the present, there are two points of view in space-time: the classic space-time view and the relativistic space-time view. Classic space-time view regards time as a universal quantity of measurement, which is separated from space. However, Albert Einstein's theory of relativity holds that the velocity of light through space has a definite value, and that invariance is independent with the movement of the light source. As a corollary of this theory, the relativistic space-time view holds that, time and space are non-separable, and time is unified in space model by the relativistic mechanics [1, 2].

The dimensionality of a mathematical space is defined as the minimum number of coordinates needs to specify the range of any coordinates. Whenever in classical mechanics or relativistic mechanics, the dimensions of space and time are different categories, which is usually called $(N+1)$ dimensional space-time divided into one-dimensional time and N -dimensional space. Therefore, the so-called multidimensional space-time refers to the space-time which is composed of one-dimensional time and multidimensional

space. Conceptually, dimension is not limited to physical objects, so high dimensional spaces often occur in mathematics and science. For instance, they can be parameter spaces or configuration spaces which are abstract spaces independent of the physical space we live in [3, 4].

Our world is a four-dimensional space-time, of which the three dimensional space is classified as Euclidean space in classical mechanics or as Riemann manifold in relativistic mechanics. The string theory proposed the concept of ten dimensional space in which we live, however the spatial coordinates of the ten dimensions are unequal, and thus the space concept of string theory is different with our space. Our three dimensional space is a vector space with symmetry, orthogonality, homogeneity and completeness, supporting classical mechanics and relativistic mechanics to achieve axiomatic system. So, the fact that our space is three dimensional is above suspicion. Why our space is three dimensional? Why it is not two-, four-, or other dimensional? The questions lasted for centuries and were related to philosophy, mathematics, physics, and even religion, arousing many thoughts and discussions especially in recent years [5-8]. Here, on the basis of two basic principles of physics: the energy conservation law and the covariance principle, and in a deductive way, the spatial properties of physical fields are studied, providing a physical explanation on why our space is three dimensional.

2. Derivations Based on Physical Principles

2.1. The Dual Patters of Physical Fields Ruled by the Energy Conservation Law

The energy conservation law states that the total energy of an isolated system remains constant—it is said to be conserved over time. Energy can neither be created nor destroyed; rather, it transforms from one form to another. In technical terms, conservation of energy can be rigorously proven by the Noether’s theorem, which is the result of a space-time symmetry. So, in our study of temporal and spatial properties, the law of conservation of energy is an important inspiration [9-11].

Starting with the discussion of harmonic wave fields: suppose that there is a physical field which propagates in space in the form of harmonic waves:

$$\vec{B}_1 = \vec{B}_0 \cos(\vec{k} \cdot \vec{r} - \omega t) \tag{1}$$

In the propagation process, if there is no energy loss such as damping and radiation, the energy conservation must be kept for an isolated system. According to Eq. (1), the instantaneous energy of field \vec{B}_1 is

$$E_1 = |\vec{B}_0|^2 \cos^2(\vec{k} \cdot \vec{r} - \omega t) \tag{2}$$

In fact, the total energy is $|\vec{B}_0|^2$, so another part of the instantaneous energy is

$$E_2 = |\vec{B}_0|^2 \sin^2(\vec{k} \cdot \vec{r} - \omega t) \tag{3}$$

carried by another field patter \vec{B}_2 , for which following condition must be satisfied:

$$|\vec{B}_2|^2 = |\vec{B}_0|^2 \sin^2(\vec{k} \cdot \vec{r} - \omega t) \tag{4}$$

Comparing Eqs. (1) and (4), the temporal variation between patters \vec{B}_1 and \vec{B}_2 is obvious:

$$\vec{B}_2 \propto \vec{B}_0 \sin(\vec{k} \cdot \vec{r} - \omega t) = \omega^{-1} \frac{\partial}{\partial t} \vec{B}_1 \tag{5}$$

Above deductions is based on the hypothesis of harmonic waves, whereas the conclusion can be generalized for that any field waves can be decomposed into harmonics. For instance, this proposal is confirmed in electromagnetics: electric field and magnetic field are two patters of electromagnetic field. Therefore, in the course of electromagnetic wave propagation, electric field and magnetic field change alternately, ensuring the conservation of electromagnetic energy. Now, extend this proposal to a general conclusion, namely: Any physical fields, as long as which can be transmitted with wave form in space, and obey the laws of conservation of energy, must have two

patters. The specific properties of the two patters can’t be deduced only from the laws of conservation of energy, however their physical natures and mechanical mechanisms are identical. This conclusion suggests that when study and describe any fields with energy conservation constraints in our real space, two patters of the field must be taken into account, even though one of them has not been found.

2.2. The Tensor Form of Physical Fields Ruled by the Covariance Principle

Any physical fields must be represented as a tensor form to coordinate the covariance principle. When a physical law is expressed by an equation, the equation is covariant if the form of the equation is constant in different coordinate systems. The covariance principle of physics emphasizes the use of finite physical quantities to describe the laws of physics, so that their measurements can be significantly associated with different reference systems. Historically, many equations of dynamics were not covariant when they were established firstly, such as Newton's second law and Schrodinger’s equation. However, these equations can be covariant after covariance treatments, and their covariance forms can reveal the laws of physics at a deeper level [12-14]. Using this principle to field physics, the dynamical equations with covariance form requires that the field quantities must be described by a tensor, and thus has the form of

$$F_{\mu\nu} \quad (\mu, \nu = 0, 1, \dots, N)$$

in a (N+1) dimensional space-time. Although the tensor description of physical fields is a mathematical means, it stems from the requirement of covariance principle. So, it embodies the physical properties of field quantities, and here can reasonably regard it as a basic principle of Physics.

In order to discuss the properties of field tensors, adopt relativistic space-time covariant form, i.e. (N+1) dimensional Minkowski space-time $\{ict, x_1, x_2, \dots, x_N\}$, where $(\mu = 0, 1, \dots, N)$ and $x_0 = ict$. Obviously, this space-time is composed of one-dimensional time and N dimensional space. Firstly, by spatial inverse transformation $\vec{r} \rightarrow -\vec{r}$, to investigate the polar or axial properties of the tensor components. Namely, keeping time direction invariant, inverse the space base vectors [14-17]:

$$\{\hat{e}_0, \hat{e}_1, \hat{e}_2, \dots, \hat{e}_N\} \rightarrow \{\hat{e}_0, -\hat{e}_1, -\hat{e}_2, \dots, -\hat{e}_N\} \tag{6}$$

The coordinate transformation between the origin coordinate system $\{x_\mu\}$ and the new coordinate system $\{x'_\mu\}$ is:

$$x_\mu = \beta^\mu_\nu x'_\nu \quad (\mu, \nu = 0, 1, \dots, N) \tag{7}$$

where

$$\beta^i_j = -\delta^i_j \quad (i, j = 1, 2, \dots, N) \tag{8}$$

and

$$\beta_\mu^0 = \delta_\mu^0, \beta_0^\mu = \delta_0^\mu (\mu = 0, 1, \dots, N) \quad (9)$$

The field tensors in the origin and the new coordinate systems are $F_{\mu\nu}$ and $F'_{\mu\nu}$, respectively, and the transformation between them is

$$F'_{ij} = \beta_i^k \beta_j^m F_{km} = (-\delta_i^k)(-\delta_j^m)F_{km} = F_{ij} \quad (10)$$

$$F'_{0j} = \beta_0^k \beta_j^m F_{km} = (\delta_0^k)(-\delta_j^m)F_{km} = -F_{0j} \quad (11)$$

$$F'_{i0} = \beta_i^k \beta_0^m F_{km} = (-\delta_i^k)(\delta_0^m)F_{km} = -F_{i0} \quad (12)$$

$$F'_{00} = \beta_0^k \beta_0^m F_{km} = (\delta_0^k)(\delta_0^m)F_{km} = F_{00} \quad (13)$$

where $(i, j = 1, 2, \dots, N; k, m = 0, 1, \dots, N)$.

Eq. (10) indicates that F'_{ij} is axial vectors, so it can be described by introducing a N dimensional vector potential $\{A_i\}$:

$$F_{ij} = \partial_i A_j - \partial_j A_i \quad (14)$$

Eq. (13) indicates that F_{00} may be an axial vector, so it must be 0 if defined according to Eq. (14). Otherwise, if F_{00} is a polar vector, it also must be 0, for that Eq. (13) and $F'_{00} = -F_{00}$ need to be satisfied simultaneously. These two possibilities bring to a certain result that:

$$F'_{00} = F_{00} = 0 \quad (15)$$

Eqs. (11) and (12) indicate that F_{0j} and F_{i0} are polar vectors, so they may arise from the field source, or from time-varying excitation of another field patter F_{ij} . For the field source, a scalar potential φ can be introduced to describe the field F_{0j} :

$$F_{0j} = \partial_j(\lambda\varphi) \quad (16)$$

where λ is the proportionality coefficient relating to the units and the attributes of space-time.

For the part of time-varying excitation of another field patter F_{ij} , Eq. (5) provides an inspiration: taking the field \vec{B}_1 as an axial vector, its temporal variation is equivalent to that of its vector potential $\{A_i\}$ due to the separation of coordinates of time and space. So, by introducing vector potential $\{A_i\}$ of \vec{B}_1 and choosing the proper unit system, Eq. (5) can be adjusted as:

$$(B_2)_i = \eta \frac{\partial}{\partial t} A_i \quad (17)$$

where η is the proportionality coefficient relating to the units and the attributes of space-time. Thus the time-varying excitation from F_{ij} contributes to F_{0j} , and can be expressed

as follows according to Eq. (17):

$$F_{0j} = \eta \frac{\partial}{\partial t} A_j \quad (18)$$

Merging Eqs. (16) and (18), the polar vector F_{0j} is

$$F_{0j} = \partial_j(\lambda\varphi) + \eta \frac{\partial}{\partial t} A_j \quad (19)$$

Considering $x_0 = ict$, and in order to obtain covariant form of Eq. (19), $\eta \partial A_j / \partial t$ should be in the form of $\eta(ic) \partial A_j / \partial(ict) = -\partial A_j / \partial(ict)$, therefore $\eta = i/c$. In the same way, λ can be assigned as i/c , and Eq. (19) comes to be:

$$F_{0j} = \partial_j \left(\frac{i}{c} \varphi \right) - \frac{\partial}{\partial(ict)} A_j \quad (20)$$

Now, introducing $(N+1)$ dimensional potential $\{\frac{i}{c} \varphi, A_1, A_2, \dots, A_N\}$ and denoting $\frac{i}{c} \varphi$ as A_0 , Eqs. (14) and (20) can be expressed as follows uniformly in a covariant form:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (\mu, \nu = 0, 1, \dots, N) \quad (21)$$

Eq. (21) expresses not only F_{ij} and F_{0j} , but also F_{i0} , so it is a general formula that describe field tensors. From the form of Eq. (21), the characteristics of the field tensor can be obtained:

Firstly, the main diagonal elements of the tensor $F_{\mu\nu}$ is 0, since that $F_{\mu\mu} = 0$ when $\mu = \nu$. Secondly, the tensor $F_{\mu\nu}$ is antisymmetric, since that

$$F_{\nu\mu} = \partial_\nu A_\mu - \partial_\mu A_\nu = -(\partial_\mu A_\nu - \partial_\nu A_\mu) = -F_{\mu\nu} \quad (22)$$

2.3. The Dimension of Our Space

For a $(N+1)$ dimensional space-time, the corresponding field tensor $F_{\mu\nu}$ is $(N+1)$ dimensional and 2 order, thus it has $(N+1)^2$ elements. Among these elements, there are $(N+1)$ principal diagonal elements which are 0, and the other $(N+1)^2 - (N+1)$ elements are antisymmetric, so there are just $((N+1)^2 - (N+1))/2$ elements that are independent of each other. According to previous analysis, tensor $F_{\mu\nu}$ contains two patters of field quantities, each patter has N components in the N dimensional space, so there are $2N$ field components in total forming $((N+1)^2 - (N+1))/2$ elements of tensor $F_{\mu\nu}$. As a result, the following dimension equation must be satisfied:

$$\frac{(N+1)^2 - (N+1)}{2} = 2N \quad (23)$$

Solving the Eq. (23), its solution is exactly 3, indicating that the space must be three dimensional in our space-time.

Based on the same reasoning, the same conclusion can be

drawn from another point of view. Eq. (10) indicates that F_{ij} is axial vectors, and is defined only in the space. For an axial vector, its accommodation space must be a cross-product vector space. Considering the antisymmetric properties of axial vector tensors, F_{ij} has N^2 elements, including N 0-valued principal diagonal elements, and other (N^2-N) antisymmetric elements. In these (N^2-N) antisymmetric elements, each element appears twice as the antisymmetric form, so there are $(N^2-N)/2$ elements to be independent of each other. According to the completeness of cross-product vector space, the number of independent vector elements is equal to the dimension of the vector space, so the dimension equation is

$$\frac{N^2 - N}{2} = N \quad (24)$$

and its solution is 3, which is consistent with the conclusion of Eq. (23).

3. Discussions

In the above derivation, three points to be noted: (1) The space discussed here is our living space, not including the intrinsic spaces of quantum mechanics, the function spaces of mathematics, and so on. (2) On the dimension problems, the distinction between classical mechanics and relativistic mechanics was not considered, for that the conclusions from these two mechanics are consistent. (3) The order of the derivation is: firstly starting from the energy conservation law, concluded that any physical field must have two patters, and the time variation relation between the two patters; then, using this conclusion and the covariance principle of physics, derived the expressions and structural features of the field tensor, and consequently draw the conclusion that our space must be three dimensional.

So far, it has been demonstrated that why our spatial dimension is three. In combination with the above derivations, our conclusions stem from two mechanism premises that have been proved by experiments and physical theories. The first premise is the law of conservation of energy. Starting from it, arrive at the conclusion that the physical field must have two patters, and predict the time variation relationship between the two patters. On this premise, perhaps from another point of view, the law of conservation of energy is the embodiment of space-time symmetry, so regarding it as the basic premise is more meaningful to reveal space-time attributes. The second premise is the covariance principle of the law of physics. On the basis of it, predict the tensor representation of the physical field which consists of axial vectors and polar vectors, and has covariant representations with antisymmetric forms. Combining corollary to this two premises, draw the conclusion that our space must be three dimensional, and from another point of view, also draw the same conclusion that our space is a vector space in which cross-product can be defined, and thus the dimension of our space must be three.

4. Conclusions

It is a universally acknowledged fact that our space is three dimensional, but there is no reasonable physical explanation erenow. Based on the accepted physical principles, this paper gives the theoretical explanation of this question. The conclusion can be summarized as: For any vector space with symmetry, orthogonality, homogeneity and completeness, the space dimension must be three on condition that: the energy obeys the law of conservation, the dynamics law is governed by covariance principle, and thus the cross-product must be defined in it. The significance of this work is that: firstly, have answered the question that why our space is three dimensional; secondly, draw a conclusion that the physical field must have two patters, which has guiding significance to the study of other physical field; and thirdly, have proved the self-consistency between the covariance principle, space-time symmetries and tensor forms, opening up a road to perfecting the axiomatic system of modern physics.

Acknowledgements

The author is grateful to the support from the National Natural Science Foundation of China (Grants Nos. 61331005), and the Innovation Team of Shaanxi province (Grant No.2014KCT-05).

References

- [1] Rynasiewicz, Robert, "Newton's Views on Space, Time, and Motion", *Stanford Encyclopedia of Philosophy* (Stanford University).
- [2] Schutz, Bernard, *Gravity from the Ground Up: An Introductory Guide to Gravity and General Relativity* (Cambridge University Press, 2004).
- [3] WS Massey, "Cross products of vectors in higher dimensional Euclidean spaces", *The American Mathematical Monthly*. 90 (10): 697–701. (1983).
- [4] Rolfsen, Dale, *Knots and Links* (Berkeley, California: Publish or Perish, 1976).
- [5] N Mankoc Borstnik and H B Nielsen, Why Nature has made a choice of one time and three space coordinates? *Journal of Physics A: Mathematical and General*, 2002, Vol. 35, no 49.
- [6] Y Itin, FW Hehl, Is the Lorentz signature of the metric of spacetime electromagnetic in origin? *Annals of Physics*, 2004, 312(1): 60-83.
- [7] D Kothawala, Minimal Length and Small Scale Structure of Spacetime, *Physical Review D*, 2013, 88(10): 130-130.
- [8] T Padmanabhan, S Chakraborty and D Kothawala, Spacetime with zero point length is two-dimensional at the Planck scale, *General Relativity & Gravitation*, 2016, 48(5): 1-8.
- [9] David J Griffiths, *Introduction to electrodynamics* (Prentice Hall, 3rd ed., pp. 559–562, 1999).
- [10] Landau LD and Lifshitz EM, *Mechanics* (Pergamon Press, 3rd ed., pp. 2–4, 1976).

- [11] Griffiths, David J., *Introduction to Quantum Mechanics* (Prentice Hall, 2nd ed., 2004).
- [12] Arnold, V. I., *Mathematical Methods of Classical Mechanics* (Springer, 1989).
- [13] Ashby, Neil, Relativity and the Global Positioning System, *Physics Today*, 55 (5): 41–47(2002).
- [14] E. J. Post, *Formal Structure of Electromagnetics: General Covariance and Electromagnetics* (Dover publications).
- [15] M. R. Spiegel, S. Lipschutz, D. Spellman, *Vector Analysis* (USA: McGraw Hill., Schaum's Outlines, 2nd ed., 2009).
- [16] John Lee, *Introduction to smooth manifolds* (Springer, p. 173, 2000).
- [17] Danielson, Donald A, *Vectors and Tensors in Engineering and Physics* (Westview (Perseus), 2nd ed., 2003).