The mathematical beauty of physics: simplicity, consistency, and unity

A British physicist Paul Dirac says of mathematical beauty as following:

> It is quite clear that beauty does depend on one's culture and upbringing for certain kinds of beauty, pictures, literature, poetry and so on...But mathematical beauty is of a rather different kind. I should say perhaps it is of a completely different kind and transcends these personal factors. It is the same in all countries and at all periods of time.

The law of physics posseses the mathematical beauty Dirac talked about. So, what is the mathematical beauty of physics? I propose that the law of physics is mathematically beautiful because it has three qualities: simplicity, consistency, and unity.

First, simplicity. As examples, I will give you Coulomb's law, Newton's law of universal gravitation and Maxwell's equations. Coulomb's law states that the Coulomb force (i.e., an electric force) is inversely proportional to the square of the distance between the two charges. For example, if the distance between two charges is doubled, the force is quartered. If the distance is tripled, the force becomes one ninth its previous value. Mathematically, we can express this as

$$\text{Coulomb force} \propto \frac{1}{r^2} \tag{1}$$

where r denotes the distance between the two charges. In other words, we say Coulomb's law obeys "inverse-square" property.

If we repeat now Coulomb's experiments in the 18th century, will we exactly get that the Coulomb force is inversely proportional to the square of the distance? To answer this question properly, you need to understand that there is an error whenever you measure something. For example, if the smallest scale in your ruler is 1 millimeter, at best you can measure something with that ruler within an error of 0.5 millimeter. For example, if something is 156.3248 millimeters long, at best you will measure it to be 156 millimeters i.e., an error of 0.3248 millimeters. As there are always errors in the measurement, there should be an error in Coulomb's law. For example, in 1971, it was found that

$$\text{Coulomb force} \propto \frac{1}{r^{2+q}} \tag{2}$$

where $q = (2.7 \pm 3.1) \times 10^{-16}$. If you do not know this kind of notation, it means

$$\frac{1}{r^{2.00000000000000058}} < \text{Coluomb force} < \frac{1}{r^{1.99999999999999996}} \tag{3}$$

In other words, the Coulomb force obeys inverse-square property to very high accuracy. Actually, there is another force that obeys inverse-square property to a very high accuracy: Newton's universal gravitation.

Newton's law of universal gravitation says that the gravitational force between two objects is proportional to the product of the masses of each object and its dependence on the distance satisfies inverse square property. This gravitational force applies to any objects with mass. By proposing this law, Newton succeeded in explaining diverse and seemingly-unrelated phenomena such as an apple falling to the ground, the planets orbiting around the sun, the moon orbiting around the earth, and the ebb and flow of tides caused by the moon and so on.

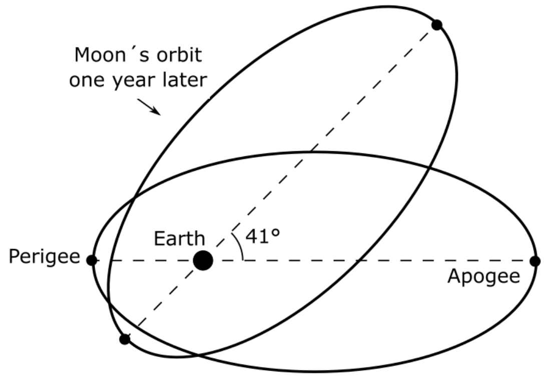

Figure 1: The figure is not in scale. In reality, the distance between the Earth and the moon changes only slightly. At the perigee, the distance to the moon is about 360,000 km. At the apogee, the distance to the moon is about 400,000 km.

However, consider the case there are three objects. For example, the Sun, the earth and the moon. The moon is attracted by the earth due to gravitation, but it is also attracted by the Sun. So, the total force acting on the moon is the sum of the gravitational force from the Sun and the gravitational force from the earth. (Of course, other planets attract the moon as well, but their forces are negligible in considering our problem.) Similarly, the total force acting on the Earth is the sum of the gravitational force from the Sun and the gravitational force from the moon. So, analyzing the motions of the earth and the moon is quite complicated; their motion is intricately related. The theory explaining the motion of the moon is called "lunar theory."

At the crudest approximation, the orbit of the moon relative to the earth is an ellipse (i.e., a squeezed circle). This is the case if the Earth is the only one that attracts the moon. We will show this in later articles. However, it is not an exact ellipse, as there is gravitational acceleration due to the Sun. Therefore, the direction of the line connecting perigee (the moon's position when it's closest to the earth) and apogee (the moon's position when it's farthest from the earth) rotates at the rate of about 41 degrees per a year. See Fig. 1. This is known as the rotation of Moon's apsides. if moon's orbit were an exact ellipse, it would mean that the dotted line doesn't rotate at all.

Newton tried to calculate these 41 degrees himself, but only got half of its value. After more mathematical technics had been developed after Newton's death, some of the best mathematicians such as Leonhard Euler, Alexis Clairaut, and Jean le Rond d'Alembert attacked this problem again in 1740s but they also got half of the correct value as Newton did. [1] Therefore, Clairaut considered the possibility that gravitational force didnt follow inverse square law property. He considered adding a k/r^3 term as follows:

$$\text{Gravitational force} \propto \frac{1}{r^2} + \frac{k}{r^3} \tag{4}$$

Of course, k should be very small so that the gravitational force doesn't deviate from the inverse square law much. Otherwise, the phenomena that Newton had already explained using his inverse square law wouldn't be still explained within observational errors. On the other hand, k should be big enough so that it can explain the 41 degrees. So did Clairaut find the value of k that explains this.

The news of Clairaut's work reached other mathematicians. For example, in 1748, a review of Clairaut's work was published in Journal de Trévoux. However, retracting his earlier position, Clairaut announced in 1749 that he succeeded in explaining the Moons apsides problem from Newton's inverse square law; he found that some of the terms he had neglected in his previous calcuation was actually important. But, he kept the details of his work secret before the official publication.

Upon hearing the news that Clairaut solved the problem, Euler sent him a letter, asking for the method. Clairaut didn't tell him the method, but just replied that he would publish his computation soon. Impatient, Euler suggested the Imperial Academy of Sciences in Saint Petersburg to propose the explanation of Moon's absides problem by Newton's gravitational law for the academic prize of 1750. As soon as his proposal was accepted, Euler wrote to Clairaut encouraing him to participate; as Euler was one of the judges, he would be able to look at their methods before anyone else. Clairaut participated in the contest and Euler soon obtained Clairaut's paper in February 1751. Then he spent the whole of March studying it. Inspired by Clairaut's paper, Euler did his research on his own on the apsidal motion. Clairaut indeed won the prize. In October, Euler wrote to the secretary of the Saint Petersburg Academy whether his own work could be published along with Clairaut's. However, he learned that the Saint Petersburg Academy could publish his work in Berlin. Possibly wanting to be the first one to publish the important result, he wrote the secretary of the Saint Petersburg Academy to give his manuscript back, but the Berlin publication failed, so he had to send the manuscript back to Saint Petersburg again. Finally, it was published in Saint Petersburg in 1753, several months after the publication of Clairaut's paper.

So much diversion. Summarizing, the Coulomb force and Newton's law of universal gravitation obey inverse-square property to a very high accuracy. Is this a coincidence? Or is there a more fundamental reason why it should be so?

Actually, it turns out that Coulomb's law and Newton's law of universal gravitation can be re-expressed by an equation now called "Gauss's law." An interesting fact is that Gauss's law can only express forces that obey the inverse square property. If the Coulomb force or Newton's law of universal gravitation were proportional to $1/r^{1.999999}$, or $1/r^{2.000001}$, or $1/r^3$, or $1/r^2 + k/r^3$, or $1/r$ then there would be no way to modify Gauss's law to account for it. That the Coulomb force and the gravitational force obey the simple inverse square property, and that they can be re-expressed by Gausss law shows the beauty of physics.

In 1861, Maxwell published "On Physical Lines of Forces" in which he established Maxwell's equations, which were collections of laws that govern everything of the electromagnetic fields. In modern notations, they are

$$\begin{aligned}
\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} &= \frac{\rho}{\epsilon_0} \\
\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} &= 0 \\
\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} &= -\frac{\partial B_z}{\partial t} \\
\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} &= -\frac{\partial B_x}{\partial t} \\
\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} &= -\frac{\partial B_y}{\partial t} \\
\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} &= \mu_0(J_z + \epsilon_0\frac{\partial E_z}{\partial t}) \\
\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} &= \mu_0(J_x + \epsilon_0\frac{\partial E_x}{\partial t}) \\
\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x} &= \mu_0(J_y + \epsilon_0\frac{\partial E_y}{\partial t})
\end{aligned} \tag{5}$$

In 1884, Heaviside reduced them to the following four simple equations.

$$\begin{aligned}
\nabla \cdot \vec{E} &= \frac{\rho}{\epsilon} \\
\nabla \cdot \vec{B} &= 0 \\
\nabla \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t}
\end{aligned} \tag{6}$$

$$\nabla \times \vec{B} = \mu_0(J + \epsilon_0 \frac{\partial \vec{E}}{\partial t})$$

The first equation (6) is Gauss's law.

In 1908, Minkowski reduced it to two. In modern notations, they are:

$$dF = 0, \quad d^*F = j \tag{7}$$

Of course, it goes without saying that Maxwell knew nothing about Oliver Heaviside's re-expression of Maxwell's equations or Minkowski's re-expression of Maxwell's equations. When I first learned Maxwell's equations, I learned a version of Heaviside's expression of Maxwell equations. Then, I didn't know much about the inevitability of nature. I thought God just happened to write Maxwell's equations in this way. However, when I learned that Maxwell's equations can be written so simply in freshman math class homework, I was mesmerized and realized that God didn't just put random terms in his equations. He had to because he didn't have much choices. This is the simplicity of physical laws. The fact that this simplification has been possible demonstrates that there is "something," a mathematical reality.

In physics, you often encounter such simplicities. It is just amazing how after a long night of calculations every ugly term cancels out and a complicated string of symbols boils down to an equation a single line long. This simplicity appears like a miracle from God. Maybe God truly is behind it. I think this must be what Albert Einstein meant in saying, "When the solution is simple, God is answering."

Second, the consistency of physicsthe ability to arrive at the same conclusion using different methodsseems to me to be another one of God's miracles. For example, Heisenberg's quantum mechanics and Schrödinger's quantum mechanics are equivalent, even though they may seem very different. (Schrödinger proved this equivalency after months of work, even though it turned out to be one-line derivation.) The actual calculations involved in each are very different, but amazingly, they give the same results.

Similarly, Feynman's quantum electrodynamics looks quite different from those of Tomonaga and Schwinger, but each set of calculations gives the same answer.[1] Yet another example: Joseph Polchinski writes in the preface of "String Theory" that "the critical dimension of the bosonic string is calculated in seven different ways in the text and exercises" in his book.

As Richard Feynman, Nobel Laureate in Physics, says, "Every theoretical physicist who is any good knows six or seven different theoretical representations for exactly the same physics."[2] I find this degree of consistency to be particularly beautiful.

Let me finally explain about a third beautiful quality of physics: unity. Albert Einstein once said:

[1] Freeman Dyson proved this equivalence.

> Creating a new theory is not like destroying an old barn and erecting a skyscraper in its place. It is rather like climbing a mountain, gaining new and wider views, discovering unexpected connections between our starting points and its rich environment. But the point from which we started out still exists and can be seen, although it appears smaller and forms a tiny part of our broad view gained by the mastery of the obstacles on our adventurous way up.

Every new theory in physics must be able to explain new phenomena in addition to describing old phenomena that an old theory has already explained adequately. This requirement constrains the development of new theories in a very strong way.

Physics is not a patchwork endeavor in which you can look at the results of new experiments independent from all others, coming up with separate formulas and theories to explain each of them. Albert Einstein once said, “A theory can be proved by experiment, but no path leads from experiment to the birth of a theory.”[3] You first have to make a theory consistent with the old and verified theories, calculate what kind of experimental results the new theory predicts, and then compare the predicted results with the experimentally attained ones. (In our later essay, “How are theories and laws in physics created?” we will talk more about this.) This is what I mean by the beauty of “unity.”

Carlo Rovelli and Francesca Vidotto wrote in their book [4]:

> Contradiction between empirically successful theories is not a curse: it is a terrific opportunity. Several of the major jumps ahead in physics have been the result of efforts to resolve precisely such contradictions. Newton discovered universal gravitation by combining Galileo’s parabolas with Kepler’s ellipses. Einstein discovered special relativity to solve the “irreconcilable” contradiction between mechanics and electrodynamics. Ten years later, he discovered that spacetime is curved in an effort to reconcile Newtonian gravitation with special relativity. Notice that these and other major steps in science have been achieved without virtually any new empirical data. Copernicus for instance constructed the heliocentric model and was able to compute the distances of the planets from the Sun using only the data in the book of Ptolemy.

‘

Indeed, it’s amazing how pure logic can lead to new theories. Regarding general relativity and Yang-Mills theory, which are fundamental to contemporary particle physics, Fundamental Physics Prize Laureate Nima Arkani-hamed said:

> So if you just hand a bunch of theorists the laws of relativity and quantum mechanics they are confident, that if you lock them up in a room, you don't let them look at what the world looks like outside, and just ask what could the world look like, this is what they will come up with.

Although general relativity was discovered in a different fashion, Yang-Mills theory actually was developed without any experimental input.

Johann Wolfgang Goethe said, "Beauty is a manifestation of secret natural laws, which otherwise would have been hidden from us forever." Certainly, if there were no beauty in physics laws, the Yang-Mills theory would have never been discovered, especially since it could not be deduced from experiments.

Let me cite two quotes. Freeman Dyson said:

> On being asked what he meant by the beauty of a mathematical theory of physics, Dirac replied that if the questioner was a mathematician then he did not need to be told, but were he not a mathematician then nothing would be able to convince him of it.

Paul Dirac said [5]:

> The mathematician plays a game in which he himself invents the rules while the physicist plays a game in which the rules are provided by Nature, but as time goes on it becomes increasingly evident that the rules which the mathematician finds interesting are the same as those which Nature has chosen.

The more you study physics, the more its simplicity, consistency, and unity will convince you of its truth. You will be sure that aliens, had they had a civilization, would have deduced Newton's laws and Einstein's theory of relativity just as we did. I am not sure the same could be said about the facts of biology or psychology. That's why I love physics the most out of all the sciences.

Finally, let me respond to a criticism of a physicist that beauty cannot be a good criterion to guide theoretical physics. She pointed out that before Kepler's time, astronomers thought that the orbit of planets must be circles, or the combination of circles, called "epicycles," because they thought that circles have beauty. Now, we know that the orbits of planets are not actually circles, but ellipses as Kepler first discovered. In other words, the orbits of planets are not of the shape which astronomers regarded as beautiful. She pointed out that scientists today would laugh at such a totally out-dated idea of ancient astronomers.

Apart from the fact that she considered an example in the era before the real physics took off by Sir Isaac Newton, let me explain what I think is wrong with this criticism. As far as I now regard the beauty of physics, it is not about phenomena, but more about its underlying principles. An orbit of a planet is a phenomenon, while the inverse square law of Newton's universal gravitation is the underlying principle that caused it. The orbits of planets are ellipses, not because ellipses are more beautiful than circles, but because the underlying principle that results in such orbits, namely, the inverse square law is beautiful. I won't repeat here why the inverse square law is beautiful, as I have already mentioned. Anyway, to put it differently, between the two choices, i.e., circular orbit and elliptical orbit, God chose the latter, because its underlying principle is more beautiful than the underlying principle that causes the circular orbits (if it exists at all).

Now, let me cite another example to support my point. We have seen that the two forces in our nature follow the inverse square law. Then, how about the other forces? What possible choices can God have, and which one does he choose? Let me provide him with two choices.

$$\text{(a) Force} \propto \frac{1}{r} \qquad \text{(b) Force} \propto \frac{e^{-\mu r}(1+\mu r)}{r^2} \tag{8}$$

For an untrained, naive person, such as the astronomers before Kepler, will think that (a) may seem more beautiful, because it is apparently simpler than (b). However, God chose (b). Why? Remember that I said that there was no way to modify Gauss's law to make the force looks like (a)? On the other hand, it is possible to modify Gauss's law to make the force follow (b). Gauss's law is given by

$$\nabla^2 \phi = \frac{\rho}{\epsilon_0} \tag{9}$$

and the force (b) can be derived from the following "modified" Gauss's law. (I put quotation marks here because this is my own expression. Nobody calls it modified Gauss's law.)

$$(\nabla^2 - \mu^2)\phi = \frac{\rho}{\epsilon_0} \tag{10}$$

Anyhow, the Japanese physicist Hideki Yukawa received the Nobel Prize in 1949 for this "modified" Gauss's law. In my later essay "How are theories and laws in physics created?" I will come up with yet another example of this kind.

References

[1] The history of lunar theory in this essay is from "The 18th-century battle over lunar motion," Physics Today, Volume 63, Issue 1, page 27 (2010) `https://physicstoday.scitation.org/doi/10.1063/1.3293410`

[2] Feynman, Richard. (2001). The Character of Physical Law (24th ed., p. 168). Cambridge, Massachusetts: The M.I.T. Press.

[3] The Sunday Times (18 Jul 1976).

[4] Rovelli, Carlo, and Francesca Vidotto. (2015). *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory* (p. 5). Cambridge University Press.

[5] Paul Adrien Maurice Dirac. *The Relation between Mathematics and Physics.* Lecture delivered on presentation of the JAMES SCOTT prize, February 6, 1939. Published in: Proceedings of the Royal Society (Edinburgh) Vol. 59, 1938-39, Part II pp. 122-129