comment

The physics of infinity

David Hilbert famously argued that infinity cannot exist in physical reality. The consequence of this statement — still under debate today — has far-reaching implications.

George F. R. Ellis, Krzysztof A. Meissner and Hermann Nicolai

It is a striking feature of modern physics and cosmology that the idea of the infinite intrudes all over the place. Well-known examples concern the possible occurrence of an infinite number of galaxies in our Universe, or multiverse models that posit the possible simultaneous existence of an infinite number of universes different from our own[1,2].

The issue of infinity also arises in connection with the infinitely small, and the search for a high-energy completion of known physics. There, one enquires about the nature and the internal consistency of physical theories at the smallest distance, where current methods often lead to apparent divergences of physical quantities. The issue of the infinite is very relevant to ongoing searches for the ultimate unification of physics into a fundamental that would encompass all phenomena, known and unknown, at arbitrarily small and at arbitrarily large distances, with no limitation to its applicability. The question then arises whether such a theory would necessarily imply the existence of actual infinity in nature.

In practice, the supposed existence of actual infinity in nature is questionable. It seems that because we have a symbol (∞) to represent infinity, many physicists believe its appearance in a theory is no big deal: it is part of the natural order. But this is not the case. Infinity is not just a very big number: it is not a number at all; rather, it is bigger than every number. It is an entity that is by its very nature unattainable, no matter what happens or how long it takes. Thus, it should not occur in the real physical universe, nor in any hypothetical multiverse. As stated eloquently by David Hilbert[3], while the concept of infinity is needed for various mathematical purposes, "The infinite is nowhere to be found in reality, no matter what experiences, observations, and knowledge are appealed to."

We can make a dual statement about zero or nothing. There is a duality between zero and infinity, expressed in the elementary identity $1/0 = \infty$. If one side of the duality does not occur in nature, also the other side ought not to. Thus, properly interpreted, zero cannot exist in physical reality either. For example, we can interpret '0' as existence in spacetime of absolutely nothing — which we know not to be the case following the uncertainty principle of quantum mechanics and the outcome of quantum field theory, which tells us that a vacuum is very far from nothing.

Hilbert's statement is thus in stark contrast to recent suggestions that infinity may actually exist in physical reality, and other unresolved issues in modern physics exemplified by the divergences of quantum field theory and the existence of spacetime singularities in general relativity.

A consistency condition

Different kinds of infinities are explained by Hilbert[3]. However, as far as potential infinities in physics are concerned, there are only two essentially different kinds: $(\infty)_{\rm VL}$, the infinity used as a placeholder for a very large number; and $(\infty)_{\rm ESS}$, essential infinity, where the paradoxical nature of infinity (such as Hilbert's hotel[3]) come into play.

Up to now, all the uses of infinity in physics have de facto been of the first rather than the second kind. Physicists have found it convenient to use the concept as a mathematical idealization where infinity occurs as a limit of large numbers, even though in physical reality there is no infinity of anything. An example is the Fourier series: while mathematically an infinite number of terms are needed to give an exact representation of the function representing an arbitrary waveform, this does not in fact mean that there are infinite physical frequencies occurring in the real physical system. In the case of a vibrating string, that would be because of the atomic nature of matter and quantum theory. But it does no harm to use an infinity of terms in a calculation if the error is smaller than the precision of any conceivable experimental test. The interesting applied mathematics question is how many terms do you need to include to get a good enough result for describing the real system (Fig. 1). How many harmonics must you include in a synthesizer to make it sound like a real violin?

In most cases, infinity is thus really just used as a placeholder $(\infty)_{\rm VL}$ for a very large number, and in such a way that the physics in question does not depend very sensitively on the actual value of this large number, provided it is taken large enough. By contrast, the paradoxical nature of essential infinity $(\infty)_{\rm ESS}$ is represented by the relations:

$$\begin{aligned}(\infty)_{\rm ESS} + a &= (\infty)_{\rm ESS}\\ b \times (\infty)_{\rm ESS} &= (\infty)_{\rm ESS}\end{aligned} \qquad (1)$$

for all numbers a and b. No finite number can satisfy both these relations (in particular these relations are not satisfied by $(\infty)_{\rm VL}$) so that they can be taken together as an operational characterization of essential infinity $(\infty)_{\rm ESS}$. It is precisely because infinity satisfies the relations (1) that it cannot occur in physical reality; in essence, it fails to obey conservation laws. By contrast, its dual zero is characterized by the relations:

$$\begin{aligned}(0)_{\rm ESS} + a &= a\\ b \times (0)_{\rm ESS} &= (0)_{\rm ESS}\end{aligned} \qquad (2)$$

for all numbers a and b. The first of these characterizes zero as unphysical: if it were to exist in some material sense, it would have no physical effect. It also does not exist in the physical world. Thus, dual to the very large number $(\infty)_{\rm VL}$, we can propose a very small number $(0)_{\rm VS}$ as an effective zero for physics purposes in different contexts. An essentially equivalent idea has recently been proposed by Gisin[8].

Therefore, in established physics neither $(\infty)_{\rm ESS}$ nor $(0)_{\rm ESS}$ appear to play a role. Our main message then is the idea that there cannot exist an essential infinity in physics or cosmology, even at the most fundamental level: this should be taken as a criterion guiding theory choice. This leads us to suggest as a consistency criterion on any fundamental theory of physics:

$$(\infty)_{\rm VL} \cong (\infty)_{\rm ESS} \leftrightarrow (0)_{\rm ESS} \cong (0)_{\rm VS} \qquad (3)$$

which is to say that in a fundamental theory there should not at a practical level exist a clash between these two notions of the infinitely large and the infinitely small.

In calculus, Leibniz (and Cauchy) had originally defined the derivative of a function

NATURE PHYSICS | VOL 14 | AUGUST 2018 | 770–772 | www.nature.com/naturephysics

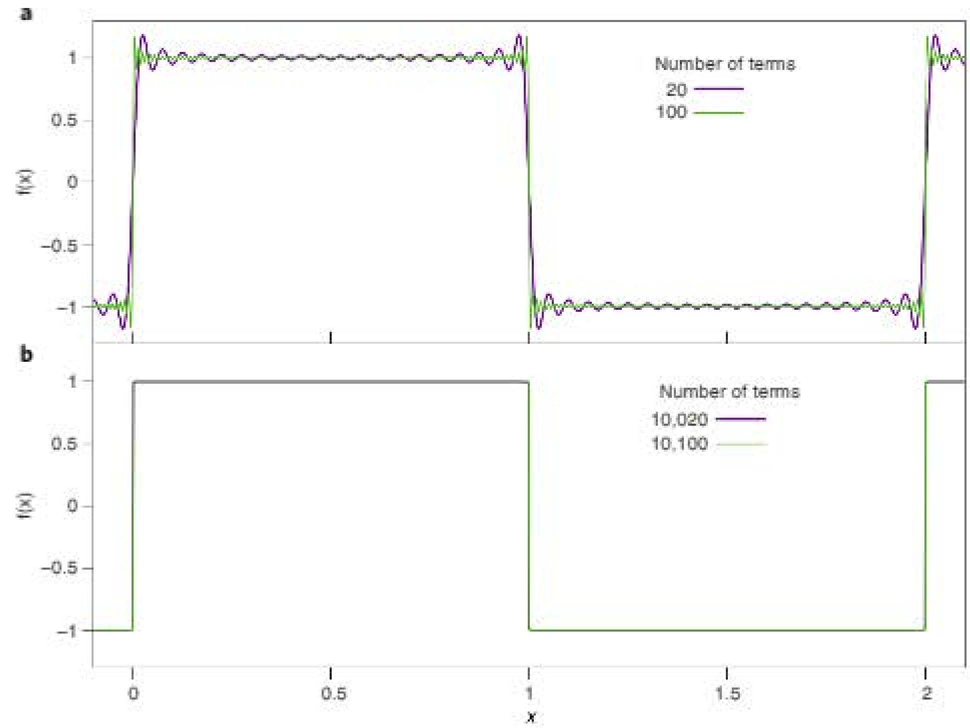

Fig. 1 | Fourier series of a square wave. a, When only few terms are considered in the Fourier series of a square wave, the accuracy of the representation of the original function depends sensitively on the number of terms included. **b**, When the number of terms is already large enough to reproduce the real function to a precision higher than what could be experimentally detected, including more terms would not change the outcome in a detectable way. Credit: courtesy of David Abergel

as the quotient of two infinitesimals — infinitely small quantities of which an infinite number would be required to fill any finite distance. Beyond this intuition, neither Leibniz nor Cauchy were able to endow the term infinitesimal with a precise meaning. Indeed, the mathematician Cantor denounced them in the nineteenth century as 'cholera bacilli' of mathematics.

The unavoidable paradoxes inherent in this way of thinking were only finally resolved when Weierstrass introduced what is now sometimes referred to as the (ε,δ)-prescription, a limiting procedure whereby one considers sequences of finite quotients that approximate the result with ever diminishing error but are never considered to actually reach zero. The notion of a derivative was thereby well-defined, such that our consistency criterion simply amounts to demanding differentiablity of a function. This discovery converted calculus to analysis in one of the great transitions in mathematics. For use of these concepts in physics and engineering, we just need small enough values of ε and δ to give the result to the desired accuracy.

The transition from calculus to analysis in mathematics corresponds to the transition in physics from use of the mathematical concept of infinity (∞) to reference to a finite but large number N, which in practice takes the place of infinity. We maintain that in each physical case where ∞ is used in a discussion, greater insight is attained by considering what large number N will suffice instead, because real physics is embodied in that number.

The very small

All successful theories of physics so far, and quantum field theory in particular, are based on the assumption that spacetime is a continuum. The apparent contradictions and inconsistencies that we encounter when we try to extrapolate the known theories to arbitrarily large energies and arbitrarily small distances are widely taken as an indication that the assumption of a spacetime continuum ultimately cannot be viable. For this reason, most approaches to quantum gravity posit that spacetime is in some sense quantized, and that there is a smallest length scale $(0)_{\mathrm{VS}}$ (usually assumed to be the Planck length $\approx 10^{-33}$ m), below which the very notion of spacetime as a continuum no longer makes sense. However, in attempts to reconcile quantum theory with Einstein's general theory of relativity (see ref. [4], for example) no consensus has emerged so far on how such a minimal length might actually be realized.

In quantum field theory, elementary particles are assumed to be point-like, leading to divergences in perturbative calculations. But do these in fact exist physically, or are the dynamical quantities in reality variables with a finite number of components $(\infty)_{\mathrm{VL}}$ in a finite volume, existing in a spacetime with a discrete structure? To deal with the infinities, an elaborate mathematical framework known as renormalization theory has been developed over the past 70 years to eliminate the high-energy (UV) infinities from perturbatively calculated physical quantities. However, it seems clear that this cannot be the last word. In fact, and in spite of the amazing quantitative successes of the renormalization programme, it appears that the standard model of particle physics cannot be extrapolated to arbitrarily large energies, and is therefore in need of a UV completion. This is even more true for Einstein's theory, which is not renormalizable. Within quantum field theory, therefore, the only hope to realize our postulate (3) would be to find a theory where all infinities cancel miraculously.

Current approaches to quantum gravity try to tackle the problem of the infinite in very different ways. String theory, widely regarded as the most promising candidate for a UV completion of physics does so by dissolving point-like constituents into extended objects (strings), and invokes a subtle mathematical property, modular invariance, for its UV finiteness. While there are indications that classical spacetime might in some sense emerge from a more fundamental (and still unknown) non-perturbative formulation of the theory, string theory in its current form cannot explain what precisely happens to space and time at the Planck scale. By contrast, loop quantum gravity posits from the outset a discrete structure that pulverizes space (and time) into an uncountable infinity of spin networks (or spin foams), a feature crucially reflected in the quantization procedure. It thereby assigns essential infinity $(\infty)_{\mathrm{ESS}}$ a central place in its very formulation. This feature has so far prevented this approach from recovering spacetime as an effective continuum and the Einstein equations in any approximate sense. There are other ideas (such as asymptotic safety) but we are still far from converging towards agreement on the right solution.

The very large

Rather different kinds of issues arise as regards claims that space is infinitely large, implying that there is no limit to spatial distances to different regions of the Universe. In this case, infinity cannot validly be used in the sense $(\infty)_{VL}$: a universe model with a claimed infinite number of galaxies is a gross misrepresentation of a physical universe where there are only a finite number of galaxies — the point being that, as stated above and represented by (1), infinity is nothing like a very large number; it is of a completely different character. With this claim, (3) would fail again. Is the number of galaxies in the Universe infinite? This is not the case if the Universe has finite spatial sections, which is necessarily true if its spatial sections are positively curved, but is possible also for spatial sections that are flat or have negative curvature. Hilbert's argument would lead us to suppose that the Universe must in fact be spatially finite, with a finite amount of matter in it, a finite number of galaxies and a finite number of living beings.

One might of course object that, even if it were true that the Universe is infinite, we could never prove this is so. If we ignore the essential problem posed by the existence of visual horizons and suppose we could in fact see an infinity of galaxies, we could not in any case count them: no matter how many we had counted, we would have made no progress in showing that there really was an infinity of entities present. There would still remain an infinity to be counted. That is the nature of $(\infty)_{ESS}$. Such untestable claims must be treated with scepticism: it is another example of claiming-to-be-scientific results that cannot be tested[7].

The most extravagant claims concern the existence of an infinite number of universes in a multiverse. Is there in fact an infinite number of expanding universe domains rather like the one in which we live, but varying as regards aspects such as the value of the cosmological constant or the fine-structure constant? Because of the existence of visual horizons as just mentioned, we cannot prove this to be the case observationally. Such a claim depends either on philosophical reasoning related to anthropic issues (how to explain why the Universe allows life to exist), or on claims that known physics related to the inflationary universe idea necessarily implies their existence. However, the physics that is meant to imply this result is not established or well-tested. Probability calculations that are supposed to legitimate these ideas run into major problems precisely because with an $(\infty)_{ESS}$ number of universes there is no way to arrive at an unambiguous probability measure.

Claims of existence of an infinite number of our own doppelgängers[1,2] are unprovable, whether they are based in claims that the one Universe we can see is infinite, or that a multiverse exists. Like all other uses of infinity, this is not testable science: it relies on $(\infty)_{ESS}$ and its paradoxical properties. It does not satisfy our proposed criterion (3) of viability. ❐

George F. R. Ellis[1]*, Krzysztof A. Meissner[2] and Hermann Nicolai[3]
[1]*Mathematics Department, University of Cape Town, Rondebosch, Cape Town, South Africa.* [2]*Faculty of Physics, University of Warsaw, Warsaw, Poland.* [3]*Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Potsdam, Germany.*
e-mail: george.ellis@uct.ac.za

Published online: 23 July 2018
https://doi.org/10.1038/s41567-018-0238-1

References

1. Tegmark, M. in *Science and Ultimate Reality: From Quantum to Cosmos* (eds Barrow, J. D., Davies, P. C. W. & Harper, C. L. Jr) Ch. 21 (Cambridge Univ. Press, Cambridge, 2004).
2. Vilenkin, A. *Many Worlds in One: The Search for Other Universes* (Hill & Wang, New York, NY, 2007).
3. Hilbert, D. in *David Hilbert's Lectures on the Foundations of Arithmetics and Logic 1917–1933* (eds Ewald, W. & Sieg, W.) 730 (Springer, Heidelberg, 2013).
4. Nicolai, H. *CERN Courier* https://cerncourier.com/gravitys-quantum-side/ (2017).
5. Hawking, S. W. & Penrose, R. *Proc. R. Soc. Lond. A* **314**, 529–548 (1970).
6. Wheeler, J. A. in *Battelle Rencontres* (eds Wheeler, J. A. & DeWitt, C. M.) (W. A. Benjamin, New York, NY, 1972).
7. Ellis, G. & Silk, J. *Nature* **516**, 321–323 (2014).
8. Gisin, N. Preprint at https://arxiv.org/abs/1803.06824 (2018).