The speed of light is one of Nature’s fundamental constants. It is pivotal to our understanding of space and time and is generally believed to restrict the speed at which information can be sent. But what exactly does it mean to have a maximum speed and why can’t it be exceeded, asks Paul Secular

As cleverly as physicists might try to break it, Nature seems to impose a limit on the speed at which information can be transmitted. This ultimate speed limit is one of the fundamental laws of Nature, governing the structure of our universe and ensuring that causes always precede effects.

Labelled in the scientific literature as $c$, it is more commonly known as the speed of light (strictly, the speed of light in vacuo). However, the name is slightly misleading because $c$ is not just the speed at which light travels in a vacuum; it is more universal than that. For example, Einstein’s general theory of relativity predicts that it is also the speed at which the effects of gravity propagate through space [1].

Whilst supersonic (i.e. faster than sound) air travel has been possible for many years now, we will see that Einstein’s theory of relativity also explains why we will never be able to fly faster than $c$.

Comparison with some everyday speeds [2] [3] [4] [5] [6] [7] [8] [9] shows that anything travelling at $c$ must be phenomenally fast by our usual standards. So fast in fact that, for most people’s intents and purposes, its speed can be treated as effectively infinite. When we read our watches, we do not take into account the time it takes light to travel from the face of the watch to our eyes. As far as we are concerned this process might as well be instantaneous.

The laws of mechanics as formulated by Galileo and Newton have no problem with the idea of unlimited speed. In fact, though they may not have realised it, their laws describe a classical world in which $c$ is infinite. But $c$ is not infinite, and classical mechanics is only an approximation that holds good for experiments with small speeds.

The existence of a finite speed limit has far-reaching consequences and led Einstein to reformulate the notions of speed, distance and even time itself. Although these concepts may
seem rather straightforward and intuitive at first glance, further inspection shows that they are anything but.

**A wave without a medium**

The first experimental evidence for light’s finite speed was provided by Cassini and Rømer in 1676 [10]. Almost 200 years later, in 1865, Maxwell published his theory of electromagnetism and proposed (correctly) that light is an electromagnetic wave. His theory suggested that light should therefore always travel at the same speed in a vacuum, no matter what the velocity of its source [11]. This proved puzzling to physicists, as they assumed that any wave propagation required some kind of medium; yet every experiment devised to detect this hypothetical medium failed.

“light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body.” [12]

While other physicists worried about these null results, Einstein made the radical suggestion that, if a medium could not be detected, then it simply did not exist. He instead took the constancy of the speed of light in vacuum as a basic assumed physical fact (i.e. a postulate), and then proceeded to investigate what logical implications followed [12].

In 1905, Einstein published his results in what was to become one of the most influential physics papers of all time. The special theory of relativity, as the work quickly became known, is a cornerstone of physics. It forms an indispensable part of all other modern theories and has been so accurately tested that it’s validity is now accepted without question.

A detailed account of Einstein’s theory lies beyond the scope of this paper, but for a beautifully accessible introduction see Bondi (1965).

**The “go and come” speed of light**

Relativity theory made physicists reassess everything they thought they knew about space and time; no longer could these concepts be viewed as the absolute, universal entities that form the backdrop of Newtonian mechanics. Using novel arguments and very little mathematics, Einstein showed how the constancy of the speed of light, along with Galileo’s relativity principle, implied that measurements of lengths and time intervals must in general differ from person to person, depending on their relative velocities [12]. Only when people are motionless with respect to one another will their measurements be identical.

These key findings can be summarised as follows:

- Moving clocks always run slower than clocks at rest (time dilation).
- The length of a moving object will decrease along its direction of motion (length contraction).

In view of these startling results, let us now ask what is meant by the speed of light. In general, the average speed of something is the ratio of the total distance travelled to the time taken:

$$\text{average speed} = \frac{\text{distance travelled}}{\text{time taken}}$$
Imagine shining a pulse of light in a straight line from one point in space to another (through a vacuum), say from point A to point B. Then:

\[
\text{average speed} = \frac{\vec{AB}}{t_B - t_A}
\]

To work out the average speed of the light pulse between these two points, we need to know three things:

1) The time at which the pulse was emitted from point A \( (t_A) \).
2) The time at which the pulse was received at point B \( (t_B) \).
3) The straight line distance between A and B \( (\vec{AB}) \).

Knowledge of the time at point A and the time at point B will clearly require 2 clocks (one at each point). The problem is: how can we ensure the two clocks are synchronised? Let us start with both clocks at point A and set them to tell exactly the same time. We now move one of them to point B. As discussed above, relativity theory tells us that the time on a moving clock will not agree with the time on a stationary clock. Therefore the two clocks are no longer ticking at the same rate. If we now bring the clock back to point A again, it will be found that it is no longer synchronised. This effect has actually been experimentally verified with incredibly accurate atomic clocks [13]. The amount by which the clocks differ will in practice be incredibly small, unless the clock was moved at a speed close to c, but to measure the speed of a light pulse we need incredibly accurate clocks; any difference in clock readings, no matter how small, will give us an incorrect value for the speed of light.

Einstein recognised that synchronising two clocks at different points in space requires some convention [12]. That is to say, there is no such thing as absolute simultaneity; it is another relative concept. In other words, we must define what we mean by simultaneous. This has upset many philosophers and physicists who think that there must somehow exist a correct, physically real form of simultaneity [14] [15]. But there isn’t. Simultaneity is not of itself a physical property that can somehow be measured. Instead it is a description of events that can be said to follow from some convention [16] [17] [18].

Having so shaken the foundations of Newton’s absolute space and time, we must ask if it is still meaningful to talk about a speed of light. The answer to this question is both yes and no.

To assign a physical reality to some quantity requires a way in which it can be measured. It turns out that the speed of light as it travels from one point to another—the one-way speed of light—can never be measured. Over the years, a number of physicists have unsuccessfully tried to devise ways of doing this [19]. With the deeper understanding of time and space that relativity brings to the table however, it becomes a logical impossibility [18].

What can safely be defined and measured is the average speed of light as it travels from point A to point B and then back to point A again. This is the two-way or “go and come” speed of light [18]. It requires just one clock (at point A) and so circumvents the problem of clock synchronisation. If we place a mirror at point B we can shine a light pulse from point A and use our single clock to time how long it takes for it to be reflected back to us. The average speed is then twice the distance from point A to point B.
divided by the total time taken.
Mathematically, this can be written as:

\[
\overrightarrow{\text{average speed}} = \frac{\overrightarrow{AB} + \overrightarrow{BA}}{t_2 - t_1} = \frac{2 \overrightarrow{AB}}{t_2 - t_1}
\]

Countless experiments have confirmed that this average speed for light in a vacuum always equals \( c \).

Although the one-way speed of light cannot be measured, it can be given a conventional definition. This is equivalent to choosing a method for synchronising our clocks.

**Einstein synchronisation**

The standard way to do this is to define the one-way velocity of light to be equal to the two-way velocity of light. Those of a mathematical bent might like to show that, in the previous example, this would mean that when the pulse of light arrives at point B, the time on clock B should be set to:

\[
t_B = \frac{t_1 + t_2}{2} = t_1 + \frac{\overrightarrow{AB}}{c}
\]

This convention is known as Einstein synchronisation. There are many different conventions one could choose, but Einstein’s is by far the simplest mathematically because of its symmetry. It might also be said to be more satisfactory from a physical point of view.

Despite being a matter of convention, there would be little sense in defining light as travelling at different speeds in different directions, when in all physical respects those directions are otherwise assumed to be identical [20].

“the speed of light in vacuum is exactly 299 792 458 metres per second” [9]

By 1983, the two-way speed of light had been measured so accurately, thanks to advances in technology, that its value became known to better precision than the standard international unit of length known as the metre. The International Bureau of Weights and Measures therefore decided to define the value of \( c \) as being equal to exactly 299,792,458 metres per second, and hence redefine the metre as “the length of the path travelled by light in vacuum during a time interval of 1/299 792 458 of a second” [9]. This shows how successful Einstein’s light postulate and synchronisation scheme have become. We now accept them as given and use them to define our measurements of length.

**Backwards in time?**

From Einstein’s special theory of relativity there follow two essential consequences which explain why \( c \) is a speed limit that neither matter nor information can surpass.

Firstly, Einstein found that in the framework of relativity theory, the classical expression for an object’s energy no longer holds true. Whereas Newtonian mechanics states that the speed of an object increases as the square root of its energy, the relativistic formula derived by Einstein has a more complicated form. The upshot is that, whereas classically we would expect an object’s speed to increase without limit as it is provided with more energy, Einstein showed that no matter how much energy an object has, its speed can never reach
c [21]. In fact his equation shows that if an object is capable of travelling at a speed of exactly $c$, then it must necessarily have zero mass.

Secondly, and even more surprisingly, special relativity predicts that if some type of signal were to exist that could travel faster than $c$, then it would be possible to use it to send information backwards in time. This is a consequence of the fact that time is relative and depends on velocity. So we see that, in a sense, our ultimate speed limit protects causality. If, somehow, someone could travel faster than the speed of light, they could potentially go backwards in time and prevent themselves from being born. Whilst situations very similar to this occur in science fiction movies such as “Back To The Future”, it is clearly nonsense as it leads to a logical paradox [21]. Moreover, no signs of backwards time-travel have ever been observed.

**Beyond special relativity**

Despite the fact that Einstein’s special theory of relativity cannot explain phenomena such as gravitation or the structure of matter, it lives on as an integral part of physics’ two deepest theories: general relativity (Einstein’s theory of gravity) and relativistic quantum mechanics. We have seen how special relativity forces us to redefine notions of time and space and how it predicts a maximum speed at which matter or information can travel. Although time is now known to be a relative concept, causality is enforced by adherence to this speed limit. Concepts of velocity, space, time, matter and information become even more complicated in general relativity and quantum mechanics, with certain speeds actually exceeding $c$.

A full discussion of these theories is beyond the scope of this paper but it will be instructive to briefly consider some of the mind-boggling issues they bring up.

“according to the general theory of relativity, the law of the constancy of the velocity of light in vacuo [...] cannot claim any unlimited validity” [22]

In general relativity, the structure of space and time changes throughout the universe depending on the distribution of matter. This makes it difficult to define such things as distance, time and velocity in a global sense. However, on the small (local) scale, space and time behave approximately as described by special relativity. Therefore any local measurement of the speed of light will always give the same constant value — $c$. However, from the point of view of an observer located at one point in space, time—and hence the speed of light—may seem to be different at other distant locations, since they are affected by gravity [23] [24] [25] [26] [27].

Mould explains this apparently variable speed of light by saying: “It is the nonlocal or global determinations of light velocity that normally violate Einstein’s special postulate.” [11].

“It may surprise you that there is an amplitude [probability] for a photon to go at speeds faster or slower than the conventional speed, $c$.” [28].

Quantum mechanics is full of faster-than-light effects, for example wavefunction collapse and tunnelling, yet in most cases it has been shown that these cannot be used to send a signal faster than $c$ and hence cannot disrupt causality [29].
Feynman’s outstandingly successful theory of quantum electrodynamics explains the electromagnetic force by assuming the existence of particles of light called “virtual photons” which can travel faster than $c$, and even go backwards in time [28]. However, these force carrying particles can be considered more of an abstract mathematical tool than real entities, as they can never be observed (hence the name “virtual”), and so do not lead to any paradoxes.

In conclusion, when considering quantum phenomena, the question of why $c$ should be the maximum speed at which information can be transmitted becomes a very subtle one. It proves difficult to provide a definitive answer, but physicists have tackled the problem in various different ways [30].

**Further reading**

Bondi, H. *Relativity and Common Sense* (1965)
Coleman, J A. *Relativity for the Layman* (1954)
Feynman, R P. *QED: The Strange Theory of Light and Matter* (1985)
Perkowitz, S. *Slow Light: Invisibility, Teleportation and Other Mysteries of Light* (2011)
Sullivan, J W N. *Three Men Discuss Relativity* (1925)

**References**


