Exploding the Big Bang

It was thought that science could tell us about the origins of the Universe. Today that great endeavour is in serious doubt

by Daniel Linford

Daniel Linford is an instructor in the Department of Philosophy and Religious Studies at Old Dominion University in Virginia, US.

Edited by Cameron Allan McKean

In the 1930s, a Belgian priest and physicist named Georges Lemaître transformed our understanding of the Universe when he envisioned its birth as a cosmic explosion. According to Lemaître, the beginning of time began with 'bright but very rapid fireworks'. His theory suggested that we lived in the fading afterglow – a slowly unfolding world of smoke and ashes. Lemaître's 'fireworks theory of evolution' painted a vivid picture, but it also presented scientists with a near-impossible puzzle: could we find evidence of the beginning of time if that slow unfolding was somehow tracked backward? Would we discover a record of the Universe's birth somewhere in the present?

Before Lemaître, the question of the Universe's birth was confined to metaphysicians and theologians. Jewish, Christian and Muslim scholars believed in divine creation, while atheist thinkers typically argued for an eternal past. The consequences of finding evidence for the beginning of time would have been enormous. If science was able to reveal when time began, the Abrahamic religions could take comfort in the confirmation of an important doctrine: the divine creation of the Universe. Alternatively, if science found that time *never* began, some conceptions of God could be ruled out. Empirical evidence, however, played no role in these philosophical and theological debates about the world's origins. In fact, no one, not even scientists, believed that the dawn of time could have left a trace in the present.

The 20th century changed everything. Lemaître's hypothesis, initially met with scepticism, suggested that the Universe had a fiery origin – one that might be discoverable. Today, many of us still believe this story. The Universe, according to popular books, television documentaries and the theme song to at least one sitcom, started with a Big Bang, marking the origins of physical matter and time itself.

The question of our Universe's birth seems settled. And yet, despite how the Big Bang is portrayed in popular culture, many physicists and philosophers of physics have long doubted whether science can truly tell us that time began. In recent decades, powerful results developed by scientifically minded philosophers appear to show that science may *never* show us that time began. The beginning of time, once imagined as igniting in a sudden burst of fireworks, is no longer an indisputable scientific fact.

When ancient, medieval and early modern thinkers debated whether the world began, they based their ideas on philosophical arguments and religious texts. 'In the beginning,' reads the first line of Genesis, 'God created the heavens and the earth.' For some theologians, however, a God who created everything (besides God himself) could also have created a *beginningless* world. Other scholars, particularly those who followed the 6th-century Christian theologian and philosopher John Philoponus, were not convinced, and argued against the idea of an eternal universe. Later, in the 13th century, Thomas Aquinas claimed that God created and sustains the world, but that we can't know about the beginning of time from the world that he created. Instead, Aquinas believed that the only way we can know about the world's beginning is through the divine word of God: the Biblical account of creation in the Book of Genesis.

Though these ideas were debated, most thinkers prior to the 20th century seemed to accept that no single moment could be reliably identified as the start of the Universe. If we could get outside our timeline, we could see whether it had a beginning, but from inside the timeline, nothing could distinguish any moment from another.

Before the 1900s, many thinkers even doubted that our world developed over time. Aristotle, the physicist Robert Hooke, the geologist Charles Lyell and others maintained that, though Earth's surface is reshaped by cyclic processes, our planet did not progressively develop. And even those who did suspect our world developed over time doubted that this development might be relevant to the world's beginning.

Time looked very different to people who lived before the 20th century. One way of understanding this difference is to consider the distinction between 'moments' and 'contents' of time. Today, we understand that a moment of time is distinct from its content: a moment is a point in time, like 8:24 am, while the content is what happens at that point, like commuting to work or eating breakfast. Imagine that everything that has happened to you today occurred an hour later. From a pre-20th-century perspective, this shift wouldn't affect your experience because your experiences are *part* of the content of each moment.

Suppose that I enter a restaurant, look at a clock, see that it reads 2:47 pm, and then order a drink. To those who lived before the 20th century, even if these events took place an hour later, everything would happen in the same way: the time would still read 2:47 pm on the clock, I would still order the drink, and still form the same sequence of memories. From this perspective, the exact timing of events is irrelevant to their contents. This led to the conclusion that time can stretch backward indefinitely for, if no moment is fundamentally different from any other and the moments themselves make no difference to the contents, there's no way to mark a beginning.

The work of physicists and cosmologists during the 20th century dramatically shifted how we view the contents and moments of time. With the development of Albert Einstein's theory of general relativity, moments and their contents became *intertwined*, allowing records of past moments to persist in the present. The contents of these moments also began to provide clues about the structure of time itself. Einstein's theory seemed to suggest that scientists might, at last, find evidence that time had a beginning.

To understand why general relativity and other scientific discoveries suggested a possible beginning to time in the 20th century, it is necessary to explore the foundational question that helped precipitate those discoveries: what is light?

In the 1860s, the Scottish physicist James Clerk Maxwell began developing an answer. In the decades before Einstein's breakthroughs, Maxwell developed equations to describe *fields* of electric and magnetic forces distributed throughout space. We can see these invisible fields in action when we rub someone's hair with a balloon or sprinkle iron filings around a magnet. To Maxwell's delight, the electric and magnetic fields described by his equations were two aspects of one unified *electromagnetic* field. And light, it turned out, was a wave in that field. For the first time, Maxwell's equations made it possible to calculate the speed of light in terms of magnetic and electric constants. But the implications of this calculation were not straightforward.

An object's speed is always measured relative to something else. For example, to find the speed of a passing car, you measure how long it takes for the car to move a certain distance on a ruler that you hold steady. Since you don't move relative to your own ruler, you always measure yourself as being at rest. Now, if you speed up until you match the speed of the other car, it will appear stationary relative to you because it no longer moves along your ruler. These principles concerning motion were part of the well-established mechanical worldview that had held sway among scientists for hundreds of years. However, Maxwell had discovered that no amount of acceleration will allow you to match the speed of light. Incredibly, light remains the same regardless of anyone's motion.

This presented physicists of the late 19th century with a paradox: though there are no absolute speeds independent of anyone's motion, the speed of light is absolute and appears to be unchanging, regardless of any observer's motion. This suggested that scientific explanations of the Universe were wrong. Something needed to be revised – either the established mechanical principles, Maxwell's new electromagnetic theory, or both.

Around the turn of the century, physicists such as Hendrik Lorentz, George Francis FitzGerald and Oliver Heaviside struggled to align the older mechanical principles with Maxwell's electromagnetic physics. In 1905, however, Einstein proposed a bold alternative, which we now call the special theory of relativity, or special relativity for short. Einstein suggested keeping the new electromagnetic physics and jettisoning the mechanical principles. This proposal yielded mind-boggling consequences for our understanding of time.

Since the exact timing of events depends on an observer's relative motion, no two events are objectively simultaneous. And since an object's length depends upon the simultaneous measurement of its front and back, an object's *length* is also relative. The same goes for the duration between two events: Einstein showed that if individuals synchronise their clocks at one location, take separate journeys, and then reunite, they will find their clocks are no longer synchronised.

In the three years after Einstein proposed special relativity, the German physicist and mathematician Hermann Minkowski began to realise that the theory did more than simply reveal the interdependence of space and time. Instead, Minkowski showed that Einstein had mathematically woven time and space into a previously unimaginable four-dimensional object: *spacetime*. With this new understanding, the pieces were falling into place for an entirely new view of the Universe's birth.

Though we perceive the world as three-dimensional, Minkowski showed that special relativity makes more sense when the world is understood as four-dimensional. Different people can have differing perspectives of the same object, like a house, which can be unified into one three-dimensional description of height, length and width. Similarly, in four-dimensional spacetime, observers experience varying perceptions of simultaneity, length and duration that can be integrated into a unified structure. Four-dimensional spacetime consolidates all reference frames – based on the measurements of rulers and clocks – into a single, unified structure that is independent of any single frame of reference. This is the profound insight Minkowski garnered from Einstein's special relativity. But though special relativity deepened our understanding of the Universe, it could not address the beginning of time on its own. A new theory of gravity was needed.

In 1907, the German physicist Johannes Stark invited Einstein to write a review of ongoing research into special relativity for a scientific journal he was editing. While writing the review, Einstein realised that Newtonian gravity and special relativity were incompatible. According to Newtonian gravity, objects exert forces on each other instantaneously, but special relativity dictates that nothing can affect anything else instantaneously. Einstein resolved this conflict over the next decade by building an entirely new theory of gravity, which is now called the general theory of relativity, or general relativity for short. Surprisingly, this theory had profound implications for the beginning of time. Through general relativity, moments and their contents become fully intertwined.

According to Einstein's new theory, spacetime affects matter, and matter affects spacetime. Just as an otherwise invisible magnetic field can be revealed by sprinkling iron filings around a magnet, the structure of spacetime can be revealed by observing how matter moves through spacetime. This insight suggested that physics might, at last, tell us something about the beginning of time.

Einstein arrived at this idea using two thought experiments. The first is known as the 'rotating disc', in which he considered the mathematical paradoxes of a rotating circle. We can reimagine this experiment in a more accessible way by considering a person encountering a merry-go-round rotating close to the speed of light. This person measures the merry-go-round's diameter and circumference by placing rulers around its lip. To their astonishment, more rulers can be placed than expected. This is because objects travelling close to the speed of light foreshorten along their direction of motion – the rulers foreshorten relative to the observing person, too. However, the diameter remains unchanged since it is perpendicular to the direction of motion. Something profound has happened. The rules of high-school geometry no longer apply. The merry-go-round appears to bend space. An object with a fixed speed and direction also has a fixed velocity. While the merry-go-round rotates with a fixed speed, the points along the circumference are constantly changing direction. Hence, the merry-go-round suggests a relationship between a changing velocity (acceleration) and the geometry of curved spaces.

In the second thought experiment, a person is standing inside a windowless elevator. While at rest on Earth, a gravitational force holds their feet to the floor. But they would feel the same force if the elevator were accelerating in just the right way – in deep space, for example. That means that local observations cannot distinguish gravitation from acceleration. Moreover, this person would feel weightless if they and the elevator were falling together on Earth. Astronauts orbiting Earth are subject to nearly the same gravitational force as we are but appear weightless because they, and their spacecraft, are falling towards Earth at the same rate. Gravitation, then, is related to acceleration, and, as Einstein showed, acceleration is related to spacetime curvature. What was less clear, however, was how gravity and spacetime curvature were related.

When viewed in the right way, such as through a fishbowl, some flat surfaces, like a tabletop, can *appear* curved. Similarly, a rapidly spinning merry-go-round can seem distorted, and a falling elevator can make a person feel weightless. Despite these appearances, mathematical procedures can distinguish truly curved spaces from apparently curved ones, and real gravitational fields from apparent ones. What Einstein's work showed was that these seemingly separate procedures turn out to be identical: apparent spacetime curvature is apparent gravity, and real curvature is real gravity.

We make these mistakes because of how we interact with the structure of spacetime. Consider the work of mapmakers. They can treat Earth as flat when mapping a city, like Chicago, but to map the entire planet, Earth's curvature must be accounted for. Imagine pasting a series of flat maps on a globe: the way each flat map connects with the others reveals the globe's curvature. Similarly, we can imagine a tiny flat map at each point of spacetime. The connections among these maps indicate spacetime's curvature.

In the absence of any forces, an object will move in a straight line at a constant speed. However, because spacetime is curved, the definition of 'straight' is not as simple as it seems. Just as someone travelling from Chicago to Paris must follow a curved path around Earth, objects in spacetime must follow curved paths to be as straight as possible within the curved spacetime around them. By observing how objects move, we can infer this curvature.

This means that matter and spacetime are intertwined. And, given this intertwining, moments of time can be distinguished by their contents: each moment, then, is unique due to its special configuration of matter and energy. And by tracking changes in the configuration of matter and energy – by tracking changes in the curvature of spacetime – perhaps some moment could be distinguished as a moment of creation? The Universe, then, might subsequently include a record of its own birth.

Einstein completed general relativity in 1916, ushering in an entirely new way of thinking about time. By the 1920s, the beginning of time stopped being a question reserved only for theologians or philosophers. The origin of the Universe now appeared to be a question with *scientific* answers.

The mathematical physics that answered this cosmological question came from general relativity's core: the Einstein field equations. These 10 equations relate the curvature of spacetime to the distribution of matter throughout spacetime. Solutions to the field equations represent possible versions of the Universe because they correspond to the many shapes that spacetime can have. If Einstein's theory is correct, our Universe should match one of these solutions.

Soon after Einstein's field equations were proposed, four physicists – Alexander Friedmann, Georges Lemaître, Howard Robertson and Arthur Walker – identified a family of solutions. The FLRW spacetimes, as they're known, describe the evolution of possible universes by assuming that each is spatially homogeneous (the same at every point) and isotropic (the same in all directions).

When some FLRW models have been extrapolated far enough backwards, the curvature of spacetime approaches infinity. According to general relativity, spacetime cannot be extended further. Thus, some FLRW spacetimes appear to expand from an initial cataclysm, beyond which – by physical law –

spacetime could not exist.

Evidence for the FLRW models, and the cataclysm, began to accumulate. In the 1920s, Edwin Hubble observed that distant galaxies are receding from us, suggesting that the observable Universe is expanding – a key feature of FLRW models. Further confirmation came in the 1940s, when the physicist George Gamow and his collaborators showed that the Universe could be explained by combining FLRW models with nuclear physics. In a 1949 radio broadcast for the BBC, the English astronomer Fred Hoyle jokingly referred to the expanding Universe as the Big Bang. The name stuck.

One alternative to this idea had already been proposed by the late 1940s. Called the steady state theory, it held that the Universe never began. However, such alternatives were largely dismissed when, in 1964, Arno Penzias and Robert Wilson discovered cosmic microwave background radiation – a kind of afterglow of the Big Bang. With Penzias and Wilson's discovery, no serious doubt remained that the observable Universe originated from a hot, dense state and has been expanding ever since.

As the 20th century progressed, questions began to emerge about the Big Bang. Was it truly the Universe's origin? The *observable* Universe may once have expanded from a hot, dense state, but that doesn't necessarily mean the entire Universe did so, or that there was nothing before the hot, dense state.

The FLRW models also came under criticism. Each of them assumed that the Universe is spatially homogeneous and isotropic. Scientists wanted to know if the catastrophe showing up in some FLRW models was a byproduct of such unrealistic assumptions. Because the Einstein field equations are so difficult to solve in anything but the simplest cases, scientists turned to Newton's theory of gravity for guidance. In some Newtonian models – which involve FLRW-like equations – there's also a past cataclysm where the gravitational field becomes undefined. But unlike in the FLRW models, Newtonian theory can be extended *past* the cataclysm.

In other Newtonian models, the cataclysm disappears altogether. In the 1950s, the physicists Otto Heckmann and Engelbert Schücking showed that the cataclysm disappears if the matter filling the Universe isn't assumed to be the same in all directions (isotropic) and, instead, changes depending on where you look. If the cataclysm disappears in non-isotropic Newtonian models, would it also disappear in more realistic general relativistic models?

In the 1960s and ’70s, physicists and mathematicians – such as Robert Geroch, Roger Penrose, Stephen Hawking and George Ellis – began studying the global properties of spacetime. Global properties are characteristics that apply to an entire space. For example, a sphere has the global property that any two initially parallel lines will meet up. Consider two global properties of spacetime: first, since a beginning must come before anything else, spacetime must have a clear direction from past to future; and, second, the entire Universe – all of spacetime – must have a boundary because without a boundary, we could always trace it back further, never encountering the beginning. Surprisingly, however, there are theoretical models (ie, solutions to the Einstein field equations) with *neither* feature. These models represent possible universes in which time does not have a past-to-future direction and spacetime has no boundary. In one example, time loops back on itself, so the Universe’s history is finite but without a beginning.

In the 1970s, landmark research by Hawking and Penrose showed that, unlike in the Newtonian models, the past cataclysm does not require the Universe to be the same everywhere and in all directions. According to what many physicists considered to be quite general and plausible assumptions, the past cataclysm appeared inescapable. However, within a decade, scientists learned that one of these assumptions about the Universe’s contents can be violated in quantum physics. The idea of the cataclysm was again up for debate, until another result appeared in 2003. Without relying on Hawking and Penrose’s assumption, three physicists – Arvind Borde, Alan Guth and Alexander Vilenkin – showed that any path along which spacetime is expanding (on average) cannot extend infinitely into the past. This means the Universe couldn’t have been expanding forever.

According to these results, no region of the Universe could have been expanding forever, but perhaps it was doing something else before it began to expand? Recently, the mathematics of Borde, Guth and Vilenkin has been challenged by Joseph Lesnefsky, Damien Easson and Paul Davies. In their view, once we do the mathematics properly, we can see that the Universe could have been expanding forever.

In recent decades, more physicists have started to think that the 'cataclysm' will be replaced with something else in a future theory. And even more radical arguments are now emerging that question our established ideas about the Big Bang – ideas that have been missed in popular accounts of the Universe. These arguments address spacetime's global structure, and they strongly suggest that no theorem and no amount of data will *ever* allow us to know whether spacetime originated in some past cataclysm.

Consider how information about spacetime is gathered. Since we can perceive light only from the past, we can receive information only from the past. As I sit at my desk, I see papers, books and a flower in a vase, but the reflected light reaching my eyes from each object is slightly delayed, taking a few nanoseconds to travel to me. I imagine I am surrounded by densely nested concentric spheres, each representing different past moments as the light travelled toward me. This collection of densely nested concentric spheres is called the past light cone.

It earned this name for how it appears when represented through diagrams. We can't draw all four dimensions of spacetime, so physicists represent spacetime with only three: two dimensions of space and one of time. With only two dimensions of space, concentric spheres become circles. And since we're representing light over time, the circles stack to form a cone, with the observer at the tip. The cone represents the region from which I can receive information – my past light cone. Every point in spacetime has its own past light cone, and together, these cones encompass all possible observations any observer can ever make.

This creates problems for physicists who hope to determine the global structure of spacetime. Can an observer determine the overall properties of spacetime only from data available within their own past light cone? The question hinges on whether there is a *single* point from which all of spacetime can be seen.

In 1977, the philosopher David Malament argued that, without an all-seeing point, no observer could fully determine the global structure of their spacetime. Only from an all-seeing point could enough information be gathered to definitively know whether the Universe has a wide variety of global properties, including an origin.

In 2009, the philosopher J B Manchak demonstrated that Malament was right. Building on Malament's proposal, Manchak showed that it is impossible to determine the overall structure of any spacetime without an all-seeing point. From any specific point within a spacetime, observers can never be certain of the global nature of their spacetime. Furthermore, all observations fit multiple possibilities – the data you have gathered from your specific past light cone can be explained by several different, even mutually exclusive, models of spacetime. In fact, all of the past light cones from all points in one spacetime (with one set of global features) can have qualitatively indistinguishable counterparts in another spacetime (with entirely different global features). Let's refer to this result as the Malament-Manchak theorem. It suggests that spacetime's global features remain unknowable.

Are there any good objections to this claim? One possibility is that our observations may be consistent with many different spacetimes. It's not uncommon for scientists to find that their observations are consistent with many different hypotheses. For example, based on all our previous observations, bits of copper conduct electricity. This observation is consistent with the hypothesis that all copper conducts electricity but is also consistent with the hypothesis that some unobserved bits of copper do *not* conduct electricity. Even though our observations are consistent with both hypotheses, we can say that all copper conducts electricity because we can confidently project from observed bits of copper to unobserved bits of copper. The philosopher Nelson Goodman calls such patterns 'lawlike'. Through them, we can project from known cases to unknown cases.

So, should we expect the unobserved parts of the Universe to behave like the parts we have observed, and could that help us infer our Universe's global properties? To make such a projection, we need a lawlike pattern. However, most lawlike patterns are defined solely by local properties. Manchak has shown that no lawlike pattern based solely on local properties would help us to determine our spacetime's global characteristics.

What about lawlike patterns that are not written in terms of local properties? The only known non-local lawlike patterns involve quantum entanglement – strange correlations in measured properties between widely separated particles. To determine whether two particles are entangled, we need to bring measured results together at a single point. However, this can't happen faster than light, which means we can't directly measure instantaneous changes taking place between the particles: we have no way of really knowing whether a particle in a terrestrial laboratory is entangled with one on the other side of the Universe. Quantum entanglement cannot help us discover spacetime's global properties either. The problem remains: as the Malament-Manchek theorem suggests, spacetime's global features remain unknowable.

This theorem has been well received by philosophers of physics during the past decade. It is often cited, but seldom rejected. Most philosophers of physics now think the matter has been settled: no amount of data can sufficiently determine spacetime's global properties. It is likely that there are no theorems strong enough to determine whether our Universe began in a past cataclysm. The Malament-Manchak theorem shows that we can't know how time began – or even *if* it began.

Once confined to metaphysics and theology, the question of whether the Universe began once seemed within the reach of science. Einstein's work transformed our understanding of space and time, binding both to matter and suggesting that spacetime itself could hold clues about its own origins. This breakthrough challenged beliefs that a 'beginning' was empirically inaccessible and led physicists to seek traces of the Universe's birth. This triumph has proven to be bittersweet.

The Malament-Manchak theorem presents us with a sobering limit: our observations, no matter how extensive, may never be sufficient to determine spacetime's global structure. Mathematically, the possible shapes and properties of the Universe remain too numerous – many versions fit equally well with the data available from our past light cones. Though the Big Bang has been popularly hailed as the origin of our

Universe, many physicists and philosophers remain unconvinced.

In the end, whether time had a beginning is a cosmological riddle. Despite dramatic scientific developments, no theorem or observation seems powerful enough to tell us whether the Universe emerged from ‘bright but very rapid fireworks’ or has always existed. Science has brought us closer to understanding the cosmos, yet it also reminds us of the limits of our knowledge. The beginning of time may remain, in the end, a mystery that we will never conclusively answer.

aeon.co 9 December 2024