UNIVERSAL

A JOURNEY THROUGH THE COSMOS

BRIAN COX &
JEFF FORSHAW

'Inspirational' Buzz Aldrin

CONTENTS

ABOUT THE BOOK

We dare to imagine a time before the Big Bang, when the entire Universe was compressed into a space smaller than an atom. And now, as Brian Cox and Jeff Forshaw show, we can do more than imagine: we can understand. Over the centuries, the human urge to discover has unlocked an incredible amount of knowledge. What it reveals to us is breathtaking.

UNIVERSAL takes us on an epic journey of scientific exploration and, in doing so, reveals how we can all understand some of the most fundamental questions about our Earth, Sun and solar system and the star-filled galaxies beyond. Some of these questions – How big is our solar system? How fast is space expanding? – can be answered from your back garden; the answers to others – How big is the Universe? What is it made of? – draw on the astonishing information now being gathered by teams of astronomers operating at the frontiers of the known Universe.

At the heart of all these questions – from the earliest attempts to quantify gravity, to our efforts to understand what dark matter is and what really happened at the birth of our Universe – is the scientific process. Science reveals a deeper beauty, connects us to each other, to our world, and to our Universe and, by understanding the groundbreaking work of others, reaches out into the unknown. What's more, as UNIVERSAL shows us, if we dare to imagine, we can all do it.

BRIAN COX &
JEFF FORSHAW

UNIVERSAL: A GUIDE TO THE COSMOS

ALLEN LANE
an imprint of
PENGUIN BOOKS

For Brian’s dad, David

1. THE STORY OF THE UNIVERSE

We dare to imagine a time when the entire observable Universe was compressed into a region of space smaller than an atom. And we can do more than just imagine. We can compute. We can compute how hundreds of billions of galaxies emerged from a single subatomic-sized patch of space dwarfed by a mote of dust, and there is precise agreement between those computations and our observations of the cosmos. It seems that human beings can know about the origins of the Universe.

Cosmology is surely the most audacious branch of science. The idea that the Milky Way, our home galaxy of 400 billion stars, was once compressed into a region so vanishingly small is outlandish enough. That the entire visible congregation of billions of galaxies once occupied such a subatomic-sized patch sounds like insanity. But to many cosmologists this claim isn't even mildly controversial.

This is not a book about knowledge handed down from on high. More than anything, it is about how we – all of us – can gain an understanding of the Universe by doing science. You might think that it's impossible for the average person to explore the Universe in much detail: don't we need access to Hubble Space Telescopes and Large Hadron Colliders? The answer is no, not always. Some fundamental questions about our Earth, our Sun, our solar system, and even the Universe beyond, are answerable from your back garden. How old are they? How big are they? How much do they weigh? We will answer these questions by doing science. We will observe, measure and think. One of the great joys of science is to understand something for the first time – to really understand, which is very different from, and far more satisfying than, knowing the facts. We will make our own measurements of the motion of Neptune, follow in the footsteps of the pioneering cosmologist Edwin Hubble in discovering that our Universe is expanding, and make an apparently trivial observation standing on a beach in south Wales.

As the book unfolds, our gaze will inevitably turn outwards towards the star-filled galaxies. To understand them, we will rely on observations and measurements that we cannot make ourselves. But we can imagine being a part of the teams of astronomers who can. How far away are the stars and galaxies? How big is the Universe? What is it made of? What was it like in the distant past? The answers to these questions will generate a cascade of new ideas, and, before the book is finished, we will be equipped to enquire about the origins of the Universe. Science is an enchanting journey of exploration. It is an exciting, rewarding process and one that leaves scientists with a feeling of being better connected to the world around them. It leaves a sense of awe and humility too; a feeling that the world is beautiful beyond imagination and that we are very privileged to be here to witness it.

Before we begin our journey, however, we will allow ourselves a glimpse of the destination. What follows next is the story of how our Universe evolved from a subatomic patch of space into the oceans of galaxies we see today. Perhaps, by the end of the book, you will judge that it might just be true.

Consider the Universe before the Big Bang. By 'Big Bang' we mean a time 13.8 billion years ago when all the material that makes up the observable Universe came into being in the form of a hot, dense plasma of elementary particles. Before this time, the Universe was very different. It was relatively cold and devoid of particles, and space itself was expanding very rapidly, which means that any particles it may have contained were moving away from each other at high speeds. The average distance between particles was doubling every 10^{-37} seconds. This is a staggering, almost incomprehensible, rate of expansion: two particles one centimetre apart at one instant were separated by 10 billion metres only 4×10^{-36} seconds later; more than twenty times the distance from the Earth to the Moon. We do not know for how long the Universe expanded like this, but it continued for at least 10^{-35} seconds. This pre-Big Bang phase of rapid expansion is known to cosmologists as the epoch of inflation.

Let us focus on a tiny speck of space a billion times smaller than a proton, the atomic nucleus of a hydrogen atom. At first glance, there is nothing particularly special about this tiny patch. It is one small part of a much larger, inflating Universe, and it looks much the same as all the other patches that surround it. The only reason this particular patch deserves our attention is that it is destined, over 13.8 billion years, to grow into our observable Universe: the region of space containing all the galaxies and quasars and black holes and stars and planets and nebulae visible from Earth today. The Universe is far bigger than the observable Universe, but we can't see it all because light can only travel a finite distance in 13.8 billion years.

Before the Big Bang, the Universe was filled with something called the 'inflaton' field; a material thing, like a still ocean filling space. The gravitational effect of the energy stored in the inflaton field caused the Universe's exponential expansion, and this is the origin of its name: it is the field responsible for inflating the Universe. On the whole, the inflaton field remained undisturbed as the Universe expanded, but it was not perfectly uniform. It had tiny ripples in it, as required by the laws of quantum physics.

By the time our observable Universe was the size of a melon, the period of inflation was drawing to a close as the energy driving it drained away. This energy was not lost, however; it was converted into a sea of elementary particles. In an instant, a cold, empty Universe became a hot, dense one. This is how inflation ended and the Big Bang began, delivering a Universe filled with the particles that were destined to evolve into galaxies, stars, planets and people.

We do not currently know which particles were present at the moment of the Big Bang, but we do know that the heaviest particles soon decayed to produce the lighter ones we know today: electrons, quarks, gluons, photons, neutrinos and dark matter.[1] We can also be confident about the particles that populated the Universe when it was around a trillionth of a second old, because we are able to re-create these conditions on Earth, at the Large Hadron Collider.[2] This is the time when empty space became filled with the Higgs field, which caused some of the elementary particles to acquire mass.[3] The weak nuclear force, responsible for the reactions that allow the stars to shine, became distinct from the electromagnetic force at this time.

A millionth of a second after the Big Bang, when the hot plasma had cooled to 10 trillion degrees celsius, the quarks and gluons formed into protons and neutrons, the building blocks of atomic nuclei. Although this primordial Universe consisted of an almost uniform soup of particles, there were slight variations in the density of the soup – an imprint of the quantum-induced ripples in the inflaton field. These variations were the seeds from which the galaxies would later grow.

One minute after the Big Bang, at around a billion degrees, the Universe was cool enough for some of the protons and neutrons to cluster together in pairs to form deuterium nuclei. Most of these then went on to partner with additional protons and neutrons to form helium and, in tiny amounts, lithium. This is the epoch of nucleosynthesis.

For the next 100,000 years or so, little happened as the Universe continued to expand and cool. Towards the end of this time, however, the dark matter gradually began to clump around the seeds sown by the ripples in the inflaton field. Regions of the Universe where there was a slight excess of dark matter grew denser, as their gravity pulled in yet more matter from the surroundings. This is the start of the gravitational clumping of matter that will eventually lead to the formation of galaxies. Meanwhile, photons, electrons and the atomic nuclei bounced and zig-zagged around, hitting each other so frequently that they formed something resembling a fluid. After 380,000 years, when the observable Universe was a thousand times smaller than it is today, temperatures dropped to those found on the surface of an average sun-like star, cool enough for electrons to be captured in orbit around the electrically charged hydrogen and helium nuclei. Suddenly, across the Universe, the first atoms formed and the Universe underwent a rapid transition from a hot plasma of electrically charged particles to a hot gas of electrically neutral particles. This had dramatic consequences, because photons interact far less with electrically neutral atoms. The Universe became transparent, which means the photons stopped zig-zagging around and started to head off in straight lines. The majority of these photons continued onwards, travelling in straight lines for the next 13.8 billion years. Some of them are just arriving at our Earth today in the form of microwaves. These ancient photons are messengers from the earliest times, and they carry a treasure trove of information that cosmologists have learnt to decode.

As the Universe continued to expand, its denser regions, composed mainly of dark matter, became ever denser under the action of gravity. Hydrogen and helium atoms clustered around the dark matter, and swirling atomic clouds grew until the densest regions collapsed inwards, increasing the pressure and temperature at their core to such an extent that they became nuclear furnaces; the fusion of hydrogen into helium was initiated, and stars formed across the Universe. A hundred million years after the Big Bang, the cosmic dark ages came to an end and the Universe was flooded with starlight. The most massive stars had brief lives and, as they ran out of hydrogen fuel, they began to fuse heavier elements in an ultimately futile battle with gravity: carbon, oxygen, nitrogen, iron – the elements of life – were made this way. When the fuel finally ran out, these stars scattered the newly minted heavy elements across space as they ended their lives as bright planetary nebulae or exploding supernovae. In a final flourish, the violent shock of each exploding supernova synthesized the heaviest elements, including gold and silver. New stars formed from the debris of the old, and congregated in their hundreds of billions in the first galaxies. The galaxies, numbered in hundreds of billions, were moulded into the giant filamentary webs that criss-cross the Universe by the gravitational pull of the dominant dark matter.

4.6 billion years ago in the Milky Way galaxy, a gas cloud enriched in stellar debris collapsed to form our Sun. Shortly afterwards, the Earth formed from the remains of the cloud. Then, 4 billion years ago, in a great ocean created from hydrogen formed in the first minute of the Universe's life and oxygen forged in long-dead stars, the geochemistry of the young Earth became biochemistry: life began. In 1687 Isaac Newton published the *Principia Mathematica*. We've obviously skipped a bit of biology.

This is the broad outline of the story of the evolution of the Universe, from before the Big Bang to Isaac Newton. It seems that collections of atoms on a cooling cinder, in possession of a precious thing called science for barely an instant, have found a way to glimpse the fires of creation. The rest of this book is the story of how we did it.

2. HOW OLD ARE THINGS?

The Earth is 4.55 billion years old, give or take 50 million years. This is a figure consistent with independent measurements of the age of the Universe, which place the Big Bang 13.8 billion years ago. It is also consistent with physical biological evidence and our understanding of evolution by natural selection, which suggest that the first living things appeared on Earth around 3.8 billion years ago. The life cycles of stars fit into this timeline too. The age of our Sun is estimated at 4.6 billion years, and similar stars are predicted to live for around 10 billion years before they die. More massive stars have much shorter lifetimes. There must have been time for at least some stars to live and die before the Earth formed, because the Earth is made out of heavy chemical elements like iron, carbon and oxygen: elements that are made inside stars. Leaping forward in time, the basalt columns of the Giant's Causeway in Ireland were formed 60 million years ago, around the time the dinosaurs became extinct. The oldest living tree is a bristlecone pine that lives in the White Mountains in California. It is – as of 2016 – 5066 years old.

All these dates are determined using very different kinds of science, but, remarkably – impressively – they fit together without contradiction. There is nothing special about this particular list; we chose this eclectic bunch simply because they reflect a variety of different 'old' things, and we could have chosen a different list. This raises the question: how do we know how old something is? Age is not a trivial thing to determine, especially for very old things, because it must be inferred indirectly. We can't sit around and watch while the Universe evolves from the hot plasma of its birth. We can't even point to direct evidence for the age of the oldest tree; nobody was around to write about it and record the date when it was a tiny sapling. But we don't need to have been present: knowledge can be acquired indirectly if we do a little detective work to collect evidence and then apply simple logic to draw conclusions. This book is all about taking a scientific approach to securing knowledge of the world around us. This approach is incremental – a framework of knowledge grows over time as we understand more about the Universe – and it sits in stark opposition to haphazard thinking: you don't build a computer by trial and error and you are prone to mistakes if you don't entertain the likelihood that you may be wrong. We trust our lives to scientific knowledge, in hospitals and aeroplanes, and exactly the same type of thinking can be used to great effect elsewhere in our lives. In this book, we will show how far it is possible to travel in understanding the Universe by taking simple, reasoned steps coupled with careful observations. In this chapter, we are going to begin by exploring the science that allows us to measure the age of things with such confidence and precision.

Let's begin with the age of the Earth. A very obvious way to start is to look at what we can see: to ask whether there are any features on the Earth's surface that might give us a clue to its age. To take a careful look at Nature, in other words, and see what we can work out from simple observation. For example, we know that river valleys are cut by flowing water, and that coastlines are subject to erosion. These are features that change with time; therefore, observing them carefully and understanding the physical processes that formed them should allow us to estimate their ages. On larger scales still, could the familiar shapes of the continents and oceans also tell us something about the way they have evolved, and how long it has taken them to do so?

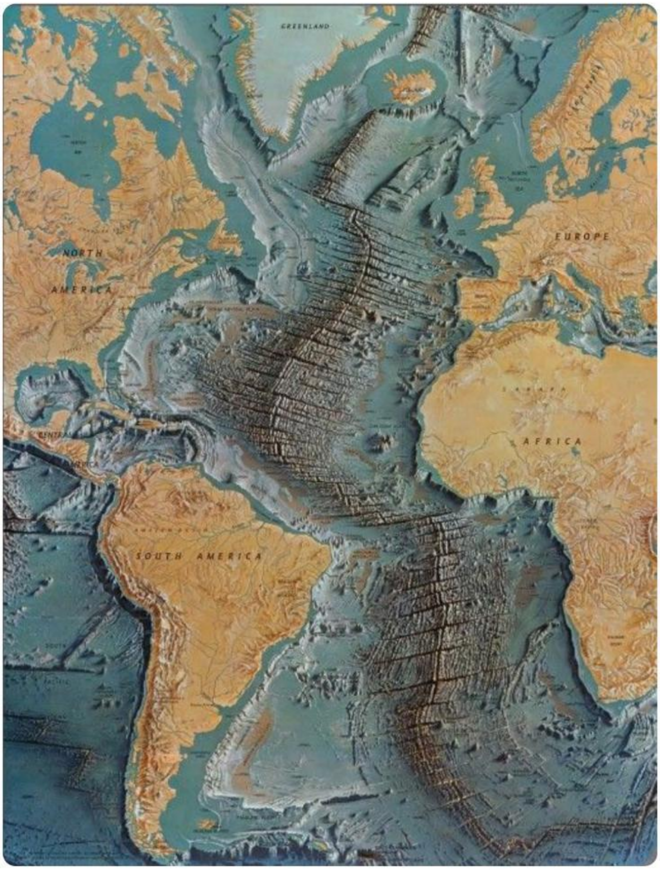

Figure 2.1
The Mid-Atlantic Ridge.

Figure 2.1 is a map of the Atlantic Ocean and the landmasses that surround it. South America and Africa in particular look as if they fit together. Let's suppose this fit is no accident and make a proposal: the continents were snuggled together at some time in the past, and have been gradually moving apart ever since. If this theory is correct, then we can make a rough estimate of the age of the Atlantic Ocean. Of course, this isn't a new idea – Alfred Wegener's idea of a global-supercontinent that broke up over time as a result of continental drift is over 100 years old. The point here, and throughout this book, is that we can uncover the science for ourselves – we want to follow in the footsteps of the great scientists, to appreciate how irresistible progress comes from simple thoughts. As a first step, we need to confirm that the broad outline of our hypothesis (that South America and Africa were once joined and have been moving apart ever since) is plausible by checking whether the Atlantic is still growing today. If it is, we can measure the current rate of separation of the continents, and – if we make the further assumption that this rate has stayed constant since the time that the continents began to separate – we will be able to make an estimate of the age of the Atlantic. There are a lot of assumptions here, but let's get on with it and see what we find.

If we were very committed experimentalists, we could measure the movements of the continents ourselves. We could pack a couple of GPS receivers into a rucksack, fly to the eastern coast of Brazil, fix one of the receivers to the ground, fly back across the Atlantic to northwest Africa – a distance of around 4000 km – and set up the second GPS receiver. Over the next few years, we could monitor how the receivers move relative to each other. We don't need to do this, because geologists have already been making such measurements for many years. Quite wonderfully, apart from using GPS receivers, the distance between North America and Europe has also been measured using a pair of radio telescopes (one in Europe and one in the USA) each focused on a distant quasar. Quasars are active galactic nuclei that most probably originate as matter accretes onto super-massive black holes in the centres of galaxies, and they are among the brightest objects and therefore the most distant we can see. Because they are so far away, they serve as excellent fixed points on the sky, which is important for triangulating the distance between Europe and the USA. We describe the measurement in a little more detail in Box 1. Do you remember those school science experiments where you had to begin with the heading ‘Apparatus: two large radio telescopes and a grid system comprising active galactic nuclei over a billion light years from Earth’?

Figure 2.3 shows a summary of the results measuring the present-day rates at which the various tectonic plates are moving. It shows that the Atlantic, between northern Brazil and northwest Africa, is currently expanding at a rate of 2.5 cm per year, which is the speed at which fingernails grow.

BOX 1. MEASURING CONTINENTAL DRIFT

The distance between two radio telescopes on the Earth's surface can be determined using a technique known as Very Long Baseline Interferometry. The two telescopes look at the same distant object in the sky, and from the difference in arrival time of light signals – determined using very precise clocks accurate to 1 second in 1 million years – the distance between the telescopes can be determined to millimetre accuracy. Quasars are so bright that they are visible at distances of many billions of light years, and being so far away guarantees that they appear still during the time of the measurement. Over twenty years, telescopes in Westford, Massachusetts, and Wettzell, Germany, have been used to determine the rate at which the Atlantic is opening between Europe and the United States. The data are shown in Figure 2.2, which shows a rate of spreading in this region of 1.7 cm/year. Satellite laser ranging, which involves bouncing laser light off satellites, and GPS measurements are also used along the length of the North and South Atlantic, and give consistent results.

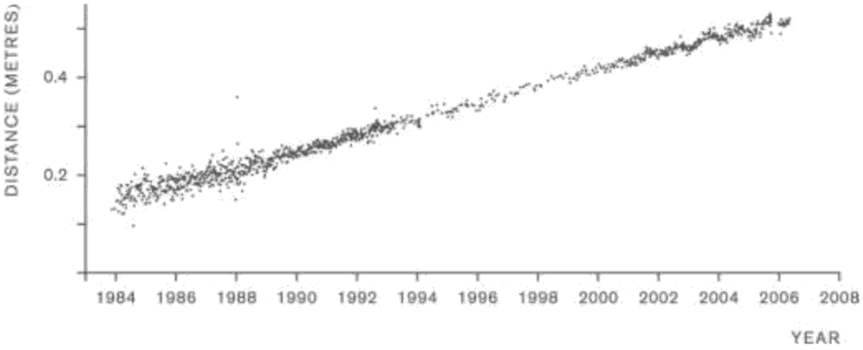

Figure 2.2 The steady rate at which Germany and the USA have been receding from each other in the recent past, as measured by a pair of radio telescopes trained on distant astronomical objects.


5.4
1.8
NORTH AMERICAN PLATE
JUAN DE FUCA PLATE
2.3
5.5
5.6
2.3
CARIBBEAN PLATE
COCOS PLATE
2.5
PACIFIC PLATE
9.2
3.0
17.2
10.1
SOUTH AMERICAN PLATE
NAZCA PLATE
6.0
4.1
3.3
7.7

Figure 2.3
How the continents are moving around. The numbers and arrows indicate the rate and direction of movement, in centimetres per year.

Working on the assumption that the continents have always been moving apart at this rate, we can now estimate the age of the Atlantic Ocean: 4000 km × 40 years/metre = 160 million years. If this figure is a good estimate, then we now also have a minimum age for the Earth – because obviously it can't be younger than the Atlantic Ocean.

We've just done what could be described as a 'back of the envelope' calculation. Obviously, we'd like to know if our number is anywhere near correct; after all, we did make a bold assertion and a very bold assumption. We asserted that the continents were once part of a single landmass and assumed that they have been moving apart at a steady rate ever since. Let's examine these assumptions more closely and try to judge how reasonable they are.

Look back at the map in Figure 2.1. It also shows the topology of the Atlantic Ocean's floor. The great range of underwater mountains running down the centre is called the Mid-Atlantic Ridge. This ridge clearly mirrors the shape of the continents on either side; it's also bang in between the two continents, in the middle of the Atlantic, and is currently spewing out material from the Earth's interior: lava that solidifies and forms a crust. This suggests a mechanism that could explain why the continents are continuing to move apart today: new ocean crust is being formed along the Mid-Atlantic Ridge.

All of which seems to indicate that our assertion is in good shape. We could, of course, have been fooled by a series of coincidences: (i) that the coastlines appear to fit together and match the shape of the Mid-Atlantic Ridge; (ii) that the Mid-Atlantic Ridge lies midway between the continents; (iii) that the lava erupting from the Mid-Atlantic Ridge has nothing to do with the currently observed widening of the ocean. But although we can be pretty confident that these are not simply coincidences, nothing we have established so far implies that the separation of these two continents has been proceeding at the same rate for over a hundred million years, and we must admit that, at this stage, this assumption is a blind guess.

evolution in the flora and fauna of the oceans (whose ages themselves are fixed using radiometric methods). The shipboard scientists involved analysed the cores and found an age–distance relationship from the Mid-Atlantic Ridge that is remarkably consistent with the assumption that the sea floor has been spreading at a constant rate. They found sediments sitting directly above the sea-floor with ages ranging from 10 million years for samples 200 km from the ridge, all the way to 70 million years for samples taken 1300 km from the ridge, corresponding to a sea-floor spreading rate of close to 2 cm/year.

So let's bring in some serious science. For decades, geoscientists have meticulously examined ocean floors across the globe and determined the age of the rocks on the seabed. This is a difficult task, and requires some beautiful science that we will discuss in a moment (see also Box 2). For now, let us just present the data, which is shown in Figure 2.4. There is a very clear pattern in the Atlantic: the youngest rocks lie along the Mid-Atlantic Ridge; the oldest are to be found bordering the continents. This fits very nicely with our proposal that the Atlantic was formed by sea-floor spreading from the Mid-Atlantic Ridge; if we are right, the rocks on the seabed should indeed get progressively older the further we travel from the ridge. Notice also that there are no sharp transitions where the rocks suddenly get much older, nor are there any extended regions where the rocks are all the same age. This is what we would expect if the rate at which new rock is being formed along the Mid-Atlantic Ridge has remained roughly constant during the time that the continents have been moving apart. The final observation we can make is to look at the age of the rocks lining the ocean floors along the edges of the continents. These are dated to be around 180 million years old – in broad agreement with our back-of-the-envelope calculation.

We haven't yet described how we go about dating rocks directly. But we can say that our suggestion that the Atlantic Ocean was created by the continents drifting apart due to geological activity along the Mid-Atlantic Ridge is consistent with the measured age of the rocks on the ocean floor.

Logical consistency and the accumulation of evidence are very important features of the way modern science works. Consider, for example, what would happen to our previous logic if the Atlantic were significantly younger than 160 million years. For the sake of argument, let's go with Bishop James Ussher, and say it is around 10,000 years old. This rather casual level of precision is doing the good bishop a disservice, because he was very specific. He asserted that the world was created on the evening of 22 October 4004 BC. The bishop performed his calculations in the late seventeenth century, using historical records and the Bible. We, on the other hand, are operating on the back of an envelope, which means we are content to work with round numbers.

If we want to accommodate an Atlantic that is 10,000 years old, but still accept that the two continents were both close together at some point, then the rate at which the continents moved apart would have had to have been much faster than the currently observed 2.5 cm per year. Instead, we would require an expansion rate of the order of 400 metres per year for most of the 10,000-year period.

The problem with an expansion rate of 400 metres per year is that the rocks along the Atlantic shores are measured to be 180,000 years old, a date that is in good agreement with the 2.5 cm per year spreading rate. If we insisted on a 10,000-year-old Earth, then it must follow that the rock ages are wrong by precisely the same factor as the spreading rate estimate is wrong. This would be quite a coincidence.

With Bishop Ussher's dating still in mind, a second possibility might be that the continents were never in fact close together, but instead they were originally created 4000 km apart, 10,000 years ago. In that scenario, the fact that the observed drift rate of 2.5 cm per year just happens to be consistent with the age inferred from dating the rocks must be regarded as a meaningless coincidence, not least because we would also need to reject as wrong the methods used to date the rocks. In addition, we'd also have to suppose that the two continents and the Mid-Atlantic Ridge all fit together quite by accident. It is clear then that the case for a young Earth requires we reject the most obvious interpretation of the facts and appeal instead to coincidence and error. We have only been studying the case of the Atlantic Ocean so far and we will meet some more examples of very old things in due course. It is up to you to judge the extent to which the evidence is convincing.

The reason that it is so difficult to make an argument against the Atlantic Ocean being around 160 million years old is that independent measurements, relying on completely different science, combine to provide a consistent picture of what happened. It is easy

image
not
available