# Introduction

VERY OCCASIONALLY, A SINGLE EVENT CAN expand your mind into a new realm. My astronomer colleagues variously trace their love of space to being given a telescope, spending a night under the stars, or watching the moon landings. The moment that sticks in my memory was discovering, at age seven, my father's ZX Spectrum computer. A musician by profession, Dad worked with a succession of early digital music synthesisers and had a background in electronic engineering. Home computing was the next frontier. The plasticky Spectrum, with its rubber keyboard and rainbow motif, was plugged into an old TV set in the damp basement and soon occupied me for hours a day. It could be instructed to do almost anything, or so it seemed.

The Spectrum stored games and other programs – apps or code, in today's terminology – on audio cassettes. Starting a program was an unreliable process that involved guesswork: fast-forwarding or rewinding to the right point on the tape, typing LOAD, pressing play on the tape deck, and waiting for a few minutes while bizarre sci-fi sounds blared and psychedelic colours flashed on the screen. Eventually the process would come to an abrupt end and, if you were lucky, the game would begin.

One day, somewhere on Dad's numerous cassettes, I found a program called SatOrb.[1] This gem challenged you to launch a satellite around a planet of your choice (you could select any from the solar system). It asked for an initial height and speed, then traced what happened to your hypothetical craft. As a yellow, pixely path slowly drew itself on the black screen, one could start to guess: will it crash onto the planet's surface, shoot off into space, or achieve the goal of a stable orbit? With practice you could place your vehicle on a suitable trajectory, and it would satisfyingly loop-the-loop – just like the moon, or one of the thousands of artificial satellites that circle Earth.

SatOrb helped kindle my interest in physics and computing, setting me on a course to spend much of my teens in the basement, writing computer code to create programs of my own. I had some books about space, which I enjoyed browsing, and I did glance at the night sky from time to time. But I never thought to ask for a telescope. The blocky, blurry, garish universe inside this little black box seemed to me more real than the far-off, out-there reaches of space.

What I didn't know at the time was that SatOrb is a rudimentary simulation.

Simulations attempt to mimic a real scenario inside a computer, and they are in such common use that they touch every part of our lives. The weather forecasts we all rely upon are based on simulations of Earth's atmosphere; when we drive cars or fly in aeroplanes, they were simulated and tested before being built; simulations are at the heart of computer-generated special effects for cinema and TV; computer games, architectural modelling, financial planning, and even public-health decision-making are all underpinned by simulations.

My job as a cosmologist involves simulating the entire universe on computers. The goal is to understand what is out there, where it came from, and how it relates to our lives here on Earth. Loosely, we are using the computer in place of a laboratory. Cosmologists can't perform traditional experiments like other scientists might: there is no way to control the universe at large, and even if there were, we would wait cosmic timescales – billions of years – for the results. Simulations offer us a computerised universe where space and time are under our control.

The ability to sculpt virtual worlds is what hooked me on computers, but my life today doesn't involve sitting alone in a darkened room, tapping away at a keyboard. I work with dozens of colleagues here in London and around the world. We publish our results in journals which reach hundreds of others. The whole endeavour rests on the cumulative work of thousands and makes use of powerful computers which fill entire air-conditioned rooms.

There is another difference between my work today and SatOrb: the trajectory of a spacecraft orbiting a planet can be calculated using pen and paper. Manual calculations might be tedious and error-prone, but there is nothing that SatOrb does that can't be accomplished by a determined human being, and no result that SatOrb produces is a surprise to a degree-level physicist. It certainly does not reveal any new truths about the reality in which we live. By contrast, when we try to simulate the universe as a whole we really do learn something new, because the results often defy expectations.

I am going to uncover the reasons for that over the course of the book. It is not just about the absurd physical extent of the universe, although that's certainly worth pausing to contemplate. It's hard enough to imagine Earth being almost 13,000 kilometres across, let alone comprehend the size of the sun, into which our planet would fit 1.3 million times over. The sun is just one of hundreds of billions of stars in our home galaxy, the Milky Way, which in turn is one of hundreds of billions of galaxies of various shapes, sizes and colours, all arranged into a vast pattern known as the cosmic web. Simulations reveal how these various structures, despite their inordinate scale, have all played a role in our own origins: as I will show, carbon-based life forms on a small rocky planet couldn't have arisen without these gargantuan support structures. It is boggling. I don't think there is any way truly to come to terms with it.

But the universe isn't just enormous; it is also enormously complex. Simulations are at their most valuable when they trace a kaleidoscope of billions of individual stars, black holes, gas clouds and dust specks. It can be exceptionally hard to anticipate the collective behaviour of such a large number of elements in combination. It does not follow easily from the physics of the individual components.

This stark difference between individual and collective behaviour can be appreciated by studying social insects here on Earth. Army ants, for example, swarm to locate colonies of smaller insects, which they then devour. While swarming, they perform extraordinary feats of cooperation, using their bodies to smooth out the terrain, or even to build bridges over uneven ground. And yet no one plots a route to the food, draws up blueprints for a bridge, or dictates where to fill potholes. There is no organising principle, and yet organised structures still emerge, structures that are hard to anticipate by studying an ant in isolation.

This can be counterintuitive at first, since human social organisations are so heavily based on hierarchies and plans. To human eyes, the collective behaviour of the army ants suggests that an executive within the colony formulates strategies to reach prey efficiently. But there is no such individual. There are just lone ants, following simple unchanging rules, like joining an ant-bridge if there are many individuals pushing behind and leaving the structure if no others crawl over.[2] The sophistication emerges from the sheer number of individuals following these rules.[3]

Understanding how a coherent, organised universe emerges from a melee of stars, gas, and dust is one of cosmologists' central goals. We build computer simulations based on the laws of nature – gravity, particle physics, light, radiation, and more – in order to obtain predictions that can be tested against night-sky observations. Because their arithmetic is accurate and fast, computers can repeatedly apply simple rules to millions or billions of sub-elements, and reveal for us how a fixed set of rules can give rise to new and surprising collective behaviours.

Simulations help us see the big picture, in which the universe transcends its small-scale laws. By the end of this book, you will have seen just how radical that picture is, describing an intricate cosmic ecosystem upon which our own existence is contingent.

## The craft of simulations

Setting out to capture the universe inside a computer requires a level of chutzpah. The difficulties are inherent in the goal: understanding how a multitude of tiny influences combine to determine an overall outcome is intrinsically hard. If the simulation misrepresents any one of the influences, even by a small amount, the conclusion might be very wrong indeed. The art of simulation lies in characterising the individual elements as precisely as possible, while understanding any remaining shortcomings so that the conclusions can be framed with appropriate caution.

These vagaries may come as a surprise. The universe follows a rigid, inarguable set of laws, or so we are taught in school, and it is true in principle that a virtual universe might be constructed by appealing directly to clockwork laws of physics that have been rigorously and extensively verified. This would seem to leave little room for error. The laws are a formalised collection of knowledge and expectations, written in the precise language of mathematics – perfect for translating into computer code. But all is not quite as it seems.

Consider the weather forecast. The presenters who tell you what to expect tomorrow base their expectations on simulations of Earth's atmosphere, combining all the innumerable tiny influences on wind, clouds and rain to make predictions for the future. But wind, clouds and rain don't appear directly within the laws of physics, which are instead written in relation to individual atoms or molecules. The weather emerges from the combined effects of the $10^{44}$ molecules in Earth's atmosphere, and a simulation would seemingly need to know the location and motion of each one.

That is not possible. Any computer's storage capacity is finite, and can be measured in bits, the smallest possible units of storage, corresponding to single switches that can either be on or off. By itself, a single bit is not terribly descriptive but you can store anything if given enough of them. Black-and-white images, for example, can be represented by bits on a grid: switched-on might represent a black dot, and switched-off an empty cell. Numbers, letters, colours, sounds, videos, Facebook friendships: all can be stored as series of bits, and the more bits you have, the more descriptive you can afford to be. A ZX Spectrum had almost 400,000 bits of memory; the laptop on which I am typing has 100 billion bits; some supercomputers have more than 10,000 trillion of them.

This is still nowhere near enough to enable simulations of Earth's atmosphere at the molecular level. If you wanted to store even a single bit of information for each molecule, you'd need to increase the current storage capacity[4] of the world's computing centres by a factor of $10^{21}$.

So a weather forecast cannot be constructed on the basis of atoms and molecules, and a simulation that tries to tell us about entire galaxies certainly won't be able to track these most fundamental constituents either. To fit inside computers, a description of the weather, of a galaxy, or of the entire universe has to lump together vast numbers of molecules, describing how they move en masse, push on each other, transport energy, react to light and radiation, and so on, all without explicit reference to the innumerable individuals within.

If the goal is to mimic reality inside computers, the available resources are laughably inadequate to reach it; the limitations of what can be achieved in practice are often stark. And yet over the last fifty years, as the technology has steadily improved, a growing community of astrophysicists has made cosmological simulation tractable with the aid of crafty physics shortcuts and tricks.

I am going to give you a taste of how these tricks were invented – sometimes through the hard work of lone PhD students who battled to have their ideas recognised; other times, by entire laboratories that banded together to crack tough problems; and, in some cases, as a result of national research priorities set by the highest level of governments. Some of the resulting shortcuts are well justified, while others are admittedly more like a stab in the dark. For that reason, not everything within a simulated universe can be taken at face value.

This problem is not unique to cosmology. Humanity has a creeping reliance on simulations, models and algorithms, with the dividing lines between these categories being fuzzy. I tend to think of algorithms as rules that determine an action to be taken: the way an autopilot corrects the course of an aeroplane, or a social-media site decides which posts to display, or a satnav calculates which route you should follow, for example. In cases where these decisions aren't totally straightforward, there needs to be an underlying model – a description of relevant phenomena like flight dynamics, human attention spans, or future traffic flow. And if a model involves large numbers of different elements interacting, it is best characterised as a simulation.

A good example of the fine line between algorithm, model and simulation is financial trading, where inspiration from physics played a major role in the 2008 economic crash.[5] The goal of financial modelling is to predict the future movement of stocks, starting from whatever real-world information is available. Such predictions are impossible in detail; but in the early 2000s, hedge funds fell in love with theoretical physicists, and their ability to make informed guesses at the future. Using a few simple assumptions about how individual stocks change value over time, so-called ‘quants’ built simulations of the long-term market movements which emerge.[6] Based on the resulting predictions, fund managers started placing speculative bets.

But models and simulations are not recreations of reality, and are therefore only as good as the simplified assumptions on which they rest. When the markets jitter, individual traders panic, trying to second-guess every decision. It is very hard to write rules for how stocks behave in these circumstances, and bets can turn out spectacularly wrong. Fund managers without the right circumspection, who were too assured, too blindly committed to the prophecies of the models or simulations, found their fortunes turned very quickly.

As early as the 1960s, mathematicians argued that the assumptions that underpin financial modelling underpredicted the risks of rare but catastrophic market falls.[7] Wise financiers of the 2000s hedged against just such eventualities and took the promises of modellers with a grain of salt. But others were impressed by the sheen of computer predictions, and lost their investors mind-boggling sums of money as a result.

The lesson here isn’t that simulations are useless, but that they are nuanced, and not to be taken literally. To understand a simulation requires a deep appreciation for its limitations, and those lie in the simplifications that separate virtual worlds from the impossibly complicated reality; the better we understand the imperfections, the more we appreciate what the simulation is really telling us.

In the aftermath of the 2008 financial crash, two leading quants published a Modeller’s Hippocratic Oath: ‘I will remember that I didn’t make the world, and it doesn’t satisfy my equations . . . I will not give the people who use my models false comfort about their accuracy. I will make the assumptions and oversights explicit to all who use them.'[8] It's a maxim that should be applied to cosmological simulations, too.

The financial risks of simulating the universe are small compared with the trillions of dollars staked on stock market bets. Still, cosmologists would like to understand which aspects of our simulations can be trusted and which cannot. We are trying to construct a story of creation that is sufficiently accurate to guide wise investment in new telescopes and laboratories; whatever money is devoted to foundational physics research should be spent shrewdly, maximising the chance of new discoveries.

## The cosmic laboratory

There are some fantastical elements in the simulations that I am going to introduce. A good place to start is with dark matter and dark energy: exotic substances, never encountered on Earth, invisible to even the most sensitive telescope, yet seemingly vital to making sense of cosmic history. Without them, simulations are unable to make sense of the universe.

The absurdity of hypothesising these materials raises the stakes considerably. On the one hand, it increases the onus to show the working of simulations, admit the limitations, and make the case for why we still, on balance, accept the outrageous conclusions. On the other, if one accepts the case for dark matter and dark energy, they are pointing towards entirely new realms of physics, so far untouched by laboratory experiments. There is nothing more exciting to scientists than this kind of frontier; we are driven by the hope that, one day, humanity will know and understand nature's secrets.

Simulations explore the perimeter of contemporary understanding in another respect, related to science's most basic assumption: that everything happens for a reason, through an unbroken chain of cause and effect. From the perspective of a weather forecast, wind, cloud, rain, heat and cold don't simply appear and disappear; they exist in distinct weather systems that can move thousands of miles before finally dispersing. Accurately charting the weather today is therefore crucial for predicting the weather tomorrow or in a few days' time.

Similarly, the universe doesn't just do whatever it pleases at a given moment, but follows a domino-like progression of events. The chain extends over almost 13.8 billion years, the current estimated age of time itself, but what happened at the start? What toppled the first domino? When building a simulation, we have no choice but to include some informed speculation about what set events in motion.

At least some aspects of the universe's creation are uncontroversial. There is overwhelming evidence that the universe has been expanding throughout its life, and that this expansion has been so extreme that the entirety of space was once microscopic. The expansion can easily be incorporated into simulations, but on its own is not sufficient to define a starting point for them.

Calculations since the 1980s have suggested that any description of our cosmic origins must lie in the theory of quantum mechanics, which is more usually regarded as a description of atomic and subatomic phenomena. Quantum physics has been well tested in laboratories for more than a century, but its implications are highly counterintuitive. The strangest assertion, at the heart of the theory, is that nothing can ever be completely certain. Subatomic particles don't have a precise location within an atom; they jump, seemingly at random, from one place to another.

Since the universe was once so tiny, it has been imprinted with these quantum phenomena. In the early cosmos, matter can't spread evenly because its tendency to jump randomly will create, through sheer luck, some regions with a little more and others with a little less material. According to simulations, these accidental differences act as seeds that grow into every astronomical structure – every galaxy, star and planet that we can see around us today, 13.8 billion years later.

The upshot is that the universe might easily have looked very different; there is a strong element of chance in our own existence, which to my mind is distinctly uncomfortable. Quantum mechanics in our initial conditions dooms any hope of predicting precisely what should be in the sky; simulations can only say what sorts of things, in what sorts of quantities, in what sorts of places, might be present. Yet, from such a weakened starting point, I am going to show how it's still possible to draw surprisingly strong conclusions about the universe.

Depending on your perspective, the expansion of space, the central role for invisible materials, and the influence of quantum mechanics may seem rather unlikely. What makes cosmology particularly difficult is appreciating and accepting the otherness of the cosmos. Reality out there doesn't accord with our human experience, and for good reason: our perspective is limited in scale, in speed and in circumstance. What would it be like to be microscopic or galactic in extent? How would it feel to travel alongside a light beam? What would happen if we fell into a black hole?

When dealing with all this, it's wise to prepare for some surprises. The materials that sculpt space aren't the ones we know from here on Earth. The rules of time and space that we intuitively understand cease to apply. The distances involved defy comprehension. Even looking through a telescope can be counterintuitive: the light we receive tells us not about the universe today, but about the universe in the past. Light travels fast, but still it can take billions of years to cross the vast expanses over which we are peering. Common sense, exquisitely honed on human experience, becomes irrelevant.

# The universe in a box

To understand the origins of our existence, we need to trace them back into deep space. To fathom deep space, how it nurtures new galaxies, stars and planets, and how these elements interrelate, we need simulations: mini-universes inside computers. And to build and interpret simulations, we need a meticulous appreciation of physics.

But this isn't physics like it's taught in schools and universities, where there is a menu of compartmentalised topics, a list of equations to memorise and a correct way to solve every problem. Nobody can simulate every subatomic particle and its influence on every other and so the physics in simulations is, at best, approximate. It is much more messy, much more open to debate, much more *human* than what we teach to undergraduates.

Nor is the physics in simulations very much to do with the future that theoreticians sometimes fantasise about, in which a single equation will come to describe every type of particle and force. Maybe one day we will have such an equation; maybe not. Such a final theory of physics, even if it perfectly describes the behaviour of individual microscopic elements of our universe, may have only marginal implications for the overarching narrative of creation. The simulator's quest lies elsewhere, in understanding the way that things – subatomic particles, or stars, or clouds of gas, or whatever – behave en masse. Just as watching a single isolated ant tells you little about the behaviour of the colony, so studying abstract equations that describe single particles reveals little about the universe.

Simulations enable a new type of understanding, offloading any hard arithmetic to a computer and allowing humans to focus instead on the connections and relationships which emerge. That, at any rate, is the dream. Getting there requires cosmologists to confront the hidden weaknesses of physics, where there are limits to what we know, restrictions on the computational power at our disposal, and compromises at every turn. Choosing and understanding the compromises is where the excitement and challenge is at its most intense.

The reward is a far-sighted vision of our home, the cosmos. And while there is a long way to go before that vision is complete – indeed, it may never be complete – simulations have already taught us about dark matter, dark energy, black holes, galaxies, and the way all these interplay to bring the universe to life. Towering far above their foundations in physics, simulations blend computation, science and human ingenuity in a way that has transformed what it means to be a cosmologist in the twenty-first century. This is their story.