ALSO BY LEE SMOLIN

The Trouble with Physics

Three Roads to Quantum Gravity

The Life of the Cosmos

Time Reborn

From the Crisis in Physics to the Future of the Universe

LEE SMOLIN

HOUGHTON MIFFLIN HARCOURT
BOSTON · NEW YORK · 2013

1

Falling

Before starting this or any other journey of discovery, we should heed the advice of the Greek philosopher Heraclitus, who, barely a few steps into the epic story that is science, had the wisdom to warn us that "Nature loves to hide." And indeed she does; consider that most of the forces and particles that science now considers fundamental lay hidden within the atom until the last century. Some of Heraclitus's contemporaries spoke of atoms, but without really knowing whether or not they existed. And their concept was wrong, for they imagined atoms as indivisible. It took until Einstein's papers of 1905 for science to catch up and form the consensus that matter is made of atoms. And six years later the atom itself was broken into pieces. Thus began the unraveling of the interior of atoms and the discoveries of the worlds hidden within.

The largest exception to the modesty of nature is gravity. It is the only one of the fundamental forces whose effects everyone observes with no need for special instruments. Our very first experiences of struggle and failure are against gravity. Consequently, gravity must have been among the first natural phenomena to be named by our species.

Nonetheless, key aspects of the common experience of falling remained hidden in plain sight until the dawn of science, and much remains hidden still. As we shall see in later chapters, one thing that remains hidden about gravity is its relation to time. So we start our journey toward the discovery of time with falling.

◆

"Why can't I fly, Daddy?"

We were on the top deck, looking down three floors to the back garden.

"I'll just jump off and fly down to Mommy in the garden, like those birds."

"Bird" had been his first word, uttered at the sparrows fluttering in the tree outside his nursery window. Here is the elemental conflict of parenthood: We want our children to feel free to soar beyond us, but we also fear for their safety in an uncertain world.

I told him sternly that people can't fly and he was absolutely never to try, and he burst into tears. To distract him, I took the opportunity to tell him about gravity. Gravity is what holds us down to Earth. It is why we fall, and why everything else falls.

The next word out of his mouth was, unsurprisingly, "Why?" Even a three-year-old knows that to name a phenomenon is not to explain it.

But we could play a game to see *how* things fall. Soon we were throwing all kinds of toys down into the garden, doing "speriments" to see whether they all fell the same way or not. I quickly found myself thinking of a question that transcends the powers of a three-year-old mind. When we throw an object and it falls as it moves away from us, it traces a curve in space. What sort of curve is it?

It's not surprising that this question doesn't occur to a three-year-old. It doesn't seem to have occurred to anyone for thousands of years after we regarded ourselves as highly civilized. It seems that Plato, Aristotle, and the other great philosophers of the ancient world were

content to watch things fall around them without wondering whether falling bodies travel along a specific kind of curve.

The first person to investigate the paths traced by falling bodies was the Italian Galileo Galilei, early in the 17th century. He presented his results in *Dialogue Concerning Two New Sciences,* which he wrote during his seventies, when he was under house arrest by the Inquisition. In this book, he reported that falling bodies always travel along the same sort of curve, which is a parabola.

Galileo not only discovered how objects fall but also explained his discovery. The fact that falling bodies trace parabolas is a direct consequence of another fact he was the first to observe, which is that all objects, whether thrown or dropped, fall with a constant acceleration.

Galileo's observation that all falling objects trace a parabola is one of the most wonderful discoveries in all of science. Falling is universal, and so is the kind of curve that falling bodies trace. It doesn't matter what the object is made of, how it is put together, or what its function is. Nor does it matter how many times, from what height, or with what forward speed we drop or throw the object. We can repeat the experiment over and over, and each time it's a parabola. The parabola is one of the simplest curves to describe. It is the set of points equidistant from a point and a line. So one of the most universal phenomena is also one of the simplest.

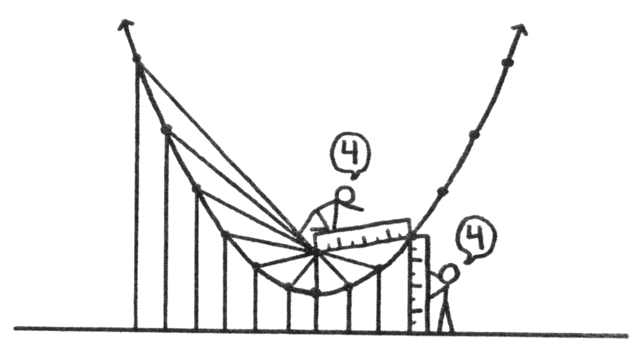

Figure 1: Definition of a parabola: the points equidistant from a point and a line.

A parabola is a concept from mathematics — an example of what we call a mathematical object — that was known to mathematicians well before Galileo's time. Galileo's observation that bodies fall along parabolas is one of the first examples we have of a law of nature — that is, a regularity in the behavior of some small subsystem of the universe. In this case, the subsystem is an object falling near the surface of a planet. This has happened a great number of times and in a great number of places since the universe began; hence there are many instances to which the law applies.

Here's a question children may ask when they're a bit older: What does it say about the world that falling objects trace such a simple curve? Why should a mathematical concept like a parabola, an invention of pure thought, have anything to do with nature? And why should such a universal phenomenon as falling have a mathematical counterpart that is one of the simplest and most beautiful curves in all of geometry?

◈

Since Galileo's discovery, physicists have profitably used mathematics in the description of physical phenomena. It may seem obvious to us now that a law must be mathematical, but for almost 2,000 years after Euclid codified his axioms of geometry no one proposed a mathematical law applying to the motion of objects on Earth. From the time of the ancient Greeks to the 17th century, educated people knew what a parabola was, but not a single one of them seems to have wondered whether the balls, arrows, and other objects they dropped, flung, or shot fell along any particular sort of curve.[1] Any one of them could have made Galileo's discovery; the tools he used were available in the Athens of Plato and the Alexandria of Hypatia. But nobody did. What changed to make Galileo think that mathematics had a role in describing something as simple as how things fall?

This question takes us into the heart of some questions easy to state but hard to answer: What is mathematics about? Why does it come into science?

Mathematical objects are constituted out of pure thought. We don't discover parabolas in the world, we invent them. A parabola or a circle or a straight line is an idea. It must be formulated and then captured in a definition. *"A circle is a set of points equidistant from a single point. . . . A parabola is a set of points equidistant from a point and a line."* Once we have the concept, we can reason directly from the definition of a curve to its properties. As we learned in high school geometry class, this reasoning can be formalized in a proof, each argument of which follows from earlier arguments by simple rules of reasoning. At no stage in this formal process of reasoning is there a role for observation or measurement.[2]

A drawing can approximate the properties demonstrated by a proof, but always imperfectly. The same is true of curves we find in the world: the curve of a cat's back when she stretches or the sweep of the cables of a suspension bridge. They will only approximately trace a mathematical curve; when we look closer, there's always some imperfection in the realization. Thus the basic paradox of mathematics: The things it studies are unreal, yet they somehow illuminate reality. But how? The relationship between reality and mathematics is far from evident, even in this simple case.

You may wonder what an exploration of mathematics has to do with an exploration of gravity. But this is a necessary digression, because mathematics is as much at the heart of the mystery of time as gravity is, and we need to sort out how mathematics relates to nature in a simple case, such as bodies falling along curves. Otherwise when we get to the present era and encounter statements like "The universe is a four-dimensional spacetime manifold," we will be rudderless. Without having navigated waters shallow enough for us to see bottom, we'll be easy prey to mystifiers who want to sell us radical metaphysical fantasies in the guise of science.

Although perfect circles and parabolas are never to be found in nature, they share one feature with natural objects: a resistance to manipulation by our fantasy and our will. The number pi — the ratio of a circle's circumference to its diameter — is an idea. But once the concept was invented, its value became an objective property, one that must be discovered by further reasoning. There have been attempts to legislate the value of pi, and they have revealed a profound misunderstanding. No amount of wishing will make the value of pi anything other than it is. The same is true for all the other properties of curves and other objects in mathematics; these objects are what they are, and we can be right or wrong about their properties but we can't change them.

Most of us get over our inability to fly. We eventually concede that we have no influence on many of the aspects of nature. But isn't it a bit unsettling that there are concepts existing only in our minds whose properties are as objective and immune to our will as things in nature? We invent the curves and numbers of mathematics, but once we have invented them we cannot alter them.

But even if curves and numbers resemble objects in the natural world in the stability of their properties and their resistance to our will, they are not the same as natural objects. They lack one basic property shared by every single thing in nature. Here in the real world, it is always some moment of time. Everything we know of in the world participates in the flow of time. Every observation we make of the world can be dated. Each of us, and everything we know of in nature, exists for an interval of time; before and after that interval, we and they do not exist.

Curves and other mathematical objects do not live in time. The value of pi does not come with a date before which it was different or undefined and after which it will change. If it's true that two parallel lines never meet in the plane as defined by Euclid, it always was and always will be true. Statements about mathematical objects like curves and numbers are true in a way that doesn't need any qualification with regard to time. Mathematical objects transcend time. But how can anything exist without existing in time?[3]

People have been arguing about these issues for millennia, and philosophers have yet to reach agreement about them. But one proposal has been on the table ever since these questions were first debated. It holds that curves, numbers, and other mathematical objects exist just as solidly as what we see in nature—except that they are not in our

world but in another realm, a realm without time. So there are *not* two kinds of things in our world, time-bound things and timeless things. There are, rather, two worlds: a world bound in time and a timeless world.

The idea that mathematical objects exist in a separate, timeless world is often associated with Plato. He taught that when mathematics speaks of a triangle, it is not any triangle in the world but an ideal triangle, which is just as real (and even more so) but exists in another realm, one outside time. The theorem that the angles of a triangle add up to 180 degrees is not precisely true of any real triangle in our physical world, but it is absolutely and precisely true of that ideal mathematical triangle existing in the mathematical world. So when we prove the theorem, we are gaining knowledge of something that exists outside time and demonstrating a truth that, likewise, is not bounded by present, past, or future.

If Plato is right, then simply by reasoning we human beings can transcend time and learn timeless truths about a timeless realm of existence. Some mathematicians claim to have deduced certain knowledge about the Platonic realm. This claim, if true, gives them a trace of divinity. How do they imagine they pulled this off? Is their claim credible?

When I want a dose of Platonism, I ask my friend Jim Brown for lunch. Both of us enjoy a good meal, during which he will patiently, and not for the first time, explain the case for belief in the timeless reality of the mathematical world. Jim is unusual among philosophers in coupling a razor-sharp mind with a sunny disposition. You sense that he's happy in life, and it makes you happy to know him. He's a good philosopher; he knows all the arguments on each side, and he has no trouble discussing those he can't refute. But I haven't found a way to challenge his confidence in the existence of a timeless realm of mathematical objects. I sometimes wonder if his belief in truths beyond the ken of humans contributes to his happiness at being human.

One question that Jim and other Platonists admit is hard for them to answer is how we human beings, who live bounded in time, in contact only with other things similarly bounded, can have definite knowledge of the timeless realm of mathematics. We get to the truths of mathematics by reasoning, but can we really be sure our reasoning is correct? Indeed, we cannot. Occasionally errors are discovered in the proofs published in textbooks, so it's likely that errors remain. You can try to get out of the difficulty by asserting that mathematical objects don't exist at all, even outside time. But what sense does it make to assert that we have reliable knowledge about a domain of nonexistent objects?

Another friend I discuss Platonism with is the English mathematical physicist Roger Penrose. He holds that the truths of the mathematical world have a reality not captured by any system of axioms. He follows the great logician Kurt Gödel in arguing that we can reason directly to truths about the mathematical realm — truths that are beyond formal axiomatic proof. Once, he said something like the following to me: "You're certainly sure that one plus one equals two. That's a fact about the mathematical world that you can grasp in your intuition and be sure of. So one-plus-one-equals-two is, by itself, evidence enough that reason can transcend time. How about two plus two equals four? You're sure of that, too! Now, how about five plus five equals ten? You have no doubts, do you? So there are a very large number of facts about the timeless realm of mathematics that you're confident you know." Penrose believes that our minds can transcend the ever changing flow of experience and reach a timeless eternal reality behind it.[4]

We discovered the phenomenon of gravity when we realized that our experience of falling is an encounter with a universal natural occurrence. In our attempts to comprehend this phenomenon, we discerned an amazing regularity: All objects fall along a simple curve the ancients invented called parabolas. Thus we can relate a universal phenomenon affecting time-bound things in the world with an invented concept that, in its perfection, suggests the possibility of truths — and of existence — outside time. If you're a Platonist, like Brown and Penrose, the discovery that bodies universally fall along parabolas is no less than the perception of a relationship between our earthly time-bound world and another, timeless world of eternal truth and beauty.

Galileo's simple discovery then takes on a transcendental or religious significance: It is the discovery of a reflection of timeless divinity acting universally in our world. The falling of a body in time in our imperfect world reveals a timeless essence of perfection at nature's heart.

This vision of transcendence to the timeless via science has drawn many into science, including myself, but now I'm sure it's wrong. The dream of transcendence has a fatal flaw at its core, related to its claim to explain the time-bound by the timeless. Because we have no physical access to the imagined timeless world, sooner or later we'll find ourselves just making stuff up (I'll present you with examples of this failing in chapters to come). There's a cheapness at the core of any claim that our universe is ultimately explained by another, more perfect world standing apart from everything we perceive. If we succumb to that claim, we render the boundary between science and mysticism porous.

Our desire for transcendence is at root a religious aspiration. The yearning to be liberated from death and from the pains and limitations of our lives is the fuel of religions and of mysticism. Does the seeking of mathematical knowledge make one a kind of priest, with special access to an extraordinary form of knowledge? Should we simply recognize mathematics for the religious activity it is? Or should we be concerned when the most rational of our thinkers, the mathematicians, speak of what they do as if it were the route to transcendence from the bounds of human life?

It is far more challenging to accept the discipline of having to explain the universe we perceive and experience only in terms of itself — to explain the real only by the real, and the time-bound only by the time-bound. But although it's more challenging, this restricted, less romantic route will ultimately be the more successful. The prize that awaits us is to understand, finally, the meaning of time on its own terms.

2

The Disappearance of Time

GALILEO WAS NOT the first to associate motion with curves. He was just the first to do it for motion on Earth. One reason it may never have occurred to anyone before Galileo that bodies fall on parabolas is that no one had perceived those parabolas directly. The paths of falling bodies were simply too fast to see.[1] But long before Galileo, people did have examples of motion slow enough to easily record. These were the motions of the sun, moon, and planets in the sky. Plato and his students had records of their positions, which the Egyptians and Babylonians had been keeping for thousands of years.

Such records amazed and delighted those who studied them, because they contained patterns — some obvious, like the annual motion of the sun, and others far from obvious, like the cycle of eighteen years and eleven days found in records of solar eclipses. These patterns were clues to the true constitution of the universe the ancients found themselves in. Over many centuries, scholars worked to decipher them, and it is by these efforts that mathematics first entered science.

But this isn't the whole answer. Galileo used no tool not available to the Greeks, so there must have been some conceptual reason for the

lack of progress on earthly motion. Did Galileo's predecessors have some blind spot about motion on Earth that Galileo lacked? What did they believe that he didn't?

Let's consider the discovery of one of the simplest and most profound patterns found by ancient astronomers. The word "planet" comes from the Greek word for wanderer, but the planets don't wander all over the sky. They all move along a great circle called the ecliptic, which is fixed with respect to the stars. The discovery of the ecliptic must have been the first step in decoding the records of planetary positions.

A circle is a mathematical object, defined by a simple rule. What does it mean if a circle is seen in the motions in the sky? Is this the visitation of a timeless phenomenon into the ephemeral, time-bound world? This might be how we would see it, but this is not how the ancients understood it. The universe, for the ancients, was split into two realms: the earthly realm, which was the arena of birth and death, of change and decay, and the heavenly realm above, which was a place of timeless perfection. For them the sky was already a transcendental realm; it was populated by divine objects that neither grew nor decayed. This was, after all, what they observed. Aristotle himself noted that "in the whole range of time past, so far as our inherited records reach, no change appears to have taken place either in the whole scheme of the outermost Heaven or in any of its proper parts."[2]

If the objects in this divine realm were to move, these movements could only be perfect and thus eternal. To the ancients, it was evident that the planets move along a circle because, being divine and perfect, they could move only on the curve that was the most perfect. But the earthly realm is not perfect, so it might have seemed bizarre to them to describe motion on Earth in terms of perfect mathematical curves.

The division of the world into an earthly realm and heavenly spheres was codified in Aristotelian physics. Everything in the earthly realm was composed of mixtures of four elements: earth, air, fire, and water. Each had a natural motion: The natural motion of earth, for example, was to seek the center of the universe. Change followed from the mixing of these four essences. Aether was a fifth element, the quintessence, which made up the heavenly realm and the objects that moved across it.

This division was the origin of the connection of elevation with transcendence. God, the heavens, perfection — these are above us, while we are trapped here below. From this perspective, the discovery that mathematical shapes are traced by motions in the sky makes sense, because both the mathematical and the heavenly are realms that transcend time and change. To know each of them is to transcend the earthly realm.

Mathematics, then, entered science as an expression of a belief in the timeless perfection of the heavens. Useful as mathematics has turned out to be, the postulation of timeless mathematical laws is never completely innocent, for it always carries a trace of the metaphysical fantasy of transcendence from our earthly world to one of perfect forms.

Long after science has moved on from the cosmos of the ancients, its basic shape influences everyday speech and metaphor. We speak of rising to the occasion. We look upward for inspiration. Whereas to fall (as in "falling in love," for instance) means to surrender to loss of control. More than that, the opposition of "ascending" and "falling" symbolizes the conflict between the corporeal and the spiritual. Heaven is above us, Hell is below. When we degrade ourselves, we sink downward into the earth. God, and everything we ultimately seek, is above us.

Music was another way the ancients experienced transcendence. Listening to music, we often experience a profound beauty that takes us "out of the moment." It's not surprising that behind the beauty of music the ancients sensed mathematical mysteries waiting to be decoded. Among the great discoveries of the school of Pythagoras was the association of musical harmonies with simple ratios of numbers. For the ancients, this was a second clue that mathematics captures the patterns in the divine. We know few personal details about Pythagoras and his followers, but we can imagine that they noticed that an affinity for mathematics often accompanies a talent for music. We would say that mathematicians and musicians share an ability to recognize, create, and manipulate abstract patterns. The ancients might have talked instead of a shared ability to perceive the divine.

Galileo was exposed to music as a child, before he was a scientist.[3] His father, Vincenzo Galilei, was a composer and an influential music theorist, who is said to have stretched violin strings across the attic of their house in Pisa so his young son could experience the relationship between harmony and ratio. Bored during a service in the Pisa cathedral, Galileo noticed that the time it took a hanging lamp to sway from side to side was independent of how wide its swing was. This independence of the period (meaning the time it takes to complete one swing or orbit) from the amplitude of a pendulum was one of his first discoveries. How did he manage it? We would use a stopwatch or a clock, but Galileo didn't have those available. We can imagine that he simply sang to himself as he watched the lamps sway over his head, since he later claimed to be able to measure time to within a tenth of a pulse-beat.

Galileo evinced a musician's showmanship as well, when he took the case for Copernicanism to the people. He wrote his ideas down in Italian instead of Latin, the language of scholars, vividly conveying them through dialogues in which imagined characters converse about science as they share a meal or a walk. For this he is praised as a democrat who disdained the hierarchy of church and university to appeal directly to the intelligence of the common person.

But as brilliant a polemicist and experimenter as he was, what's stunning about Galileo's work are the new questions he asked — thanks in part to the liberation from ancient dogma that was the legacy of the Italian Renaissance. The ancient distinction between the earthly and divine realms that had long kept people from thinking seems to have left Galileo unimpressed. Leonardo had discovered proportion and harmony in static form, but Galileo looked for mathematical harmony in everyday motions, such as those of pendulums and balls rolling down inclined planes. Before he was democratic in his strategy of communication, he was a democrat about the universe.

Galileo destroyed the divinity of the sky when he discovered that heavenly perfection was a lie. He did not invent the telescope, and he may not have been the only one who used the new invention to look at the heavens. But his unique perspective and talents led him to make a fuss about what he saw there, which was imperfection. The sun has spots. The moon is not a perfect sphere of quintessence; it has mountains, just like Earth. Saturn has a strange threefold shape. Jupiter has moons, and there are vastly more stars than those seen with the naked eye.

This decline of divinity had been anticipated a few years earlier, in 1577, when the Danish astronomer Tycho Brahe watched a comet penetrate the perfect spheres of Heaven. Tycho was the last and greatest of the naked-eye astronomers, and he and his assistants accumulated over his lifetime the best measurements of planetary motions that had ever been made. These sat in his record books undecoded until 1600, when he employed an irascible young assistant, Johannes Kepler.

The planets move along the ecliptic, but they are not seen to move consistently. They all move in the same direction, but occasionally pause and reverse themselves, moving backward for a while. This retrograde motion was a great mystery to the ancients. Its real meaning is that the Earth is a planet, too, which moves around the sun as the other planets do. The planets appear to stop and start only from Earth's perspective. Mars moves eastward in our sky when it's ahead of us and reverses direction when Earth catches up. Its retrograde motion is simply an effect of Earth's motion, but the ancients couldn't see it that way, because they were stuck with the false idea that the Earth is at rest at the center of the universe. Since Earth is still, the perceived motion of the planets must be their real motion; hence the ancient astronomers had to explain the retrograde motions as if they were caused by the planets' intrinsic motion. To do so, they imagined an awkward arrangement involving two kinds of circles, in which each planet was attached to a small circle rotating around a point that itself moved on a bigger circle around the Earth.

The epicycles, as these mini-circles were called, rotated with a period of one Earth year, because they were nothing but the shadow of Earth's motion. Other adjustments required still more circles; it took

fifty-five circles to get it all to work. By assigning the right periods to each of the big circles, the Alexandrian astronomer Ptolemy calibrated the model to a remarkable degree of accuracy. A few centuries later, Islamic astronomers fine-tuned the Ptolemaic model, and in Tycho's time it predicted the positions of the planets, the sun, and the moon to an accuracy of 1 part in 1,000—good enough to agree with most of Tycho's observations. Ptolemy's model was beautiful mathematically, and its success convinced astronomers and theologians for more than a millennium that its premises were correct. And how could they be wrong? After all, the model had been confirmed by observation.

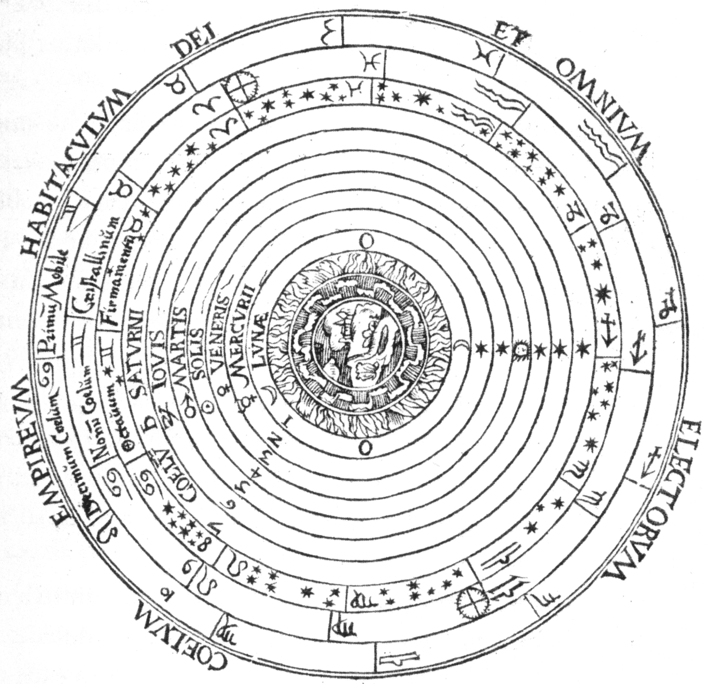

Figure 2: A schematic of the Ptolemaic vision of the universe.[4]

There's a lesson here, which is that neither mathematical beauty nor agreement with experiment can guarantee that the ideas a theory is based on bear the slightest relation to reality. Sometimes a decoding of the patterns in nature takes us in the wrong direction. Sometimes we fool ourselves badly, as individuals and as a society. Ptolemy and Aristotle were no less scientific than today's scientists. They were just unlucky, in that several false hypotheses conspired to work well together. There is no antidote for our ability to fool ourselves except to keep the process of science moving so that errors are eventually forced into the light.

It fell to Copernicus to decipher the meaning of the fact that all the epicycles have the same period and move in phase with the sun's orbit. He put Earth in its rightful place, as a planet, and the sun near the center of the universe. This simplified the model but introduced a tension the ancient cosmology couldn't survive. Why should Earth's sphere be any different from those of the heavens, if Earth is just another planet traveling through the heavens?

However, Copernicus was a reluctant revolutionary who missed other clues. A big one was that even after Earth's motion was accounted for, the planets' orbits were not precisely circles. Unable to escape the idea that motions on the sky must be compounded from circles, he solved this problem just as Ptolemy had fourteen centuries earlier. He introduced epicyles as needed to get the theory to fit the data.

The least circular orbit is that of Mars. It was Kepler's great luck — and science's too — that Tycho assigned to him the problem of deciphering the orbit of Mars, and, after working for many years after he left Tycho's service, Kepler found that Mars traces an ellipse, not a circle, in space.

This was revolutionary in ways that may not be apparent to a modern reader. In an Earth-centered cosmology, the planets don't trace a closed path of any sort, because their paths relative to Earth each combine two circular motions with different periods. It is only when the orbits are plotted with respect to the sun that they make closed paths. Only then does it become possible to ask what the shape of an orbit is. So putting the sun at the center deepens the harmony of the world.

Once the planetary orbits were understood to be ellipses, the explanatory power of Ptolemy's theory was shattered. A slew of new

questions arose: *Why* do the planets move in elliptical orbits? And what keeps them from wandering off? What compels them to move at all, rather than just sitting still in space? Kepler's answer was a wild guess that turned out to be half right: *What moves the planets around in their orbits is a force from the sun.* Imagine the sun as a rotating octopus, its arms sweeping the planets around as it turns. This was the first time anyone had suggested the sun as the source of a force that affects the planets. He just got the direction of the force wrong.

Tycho and Kepler smashed the heavenly spheres and in so doing unified the world. This unification had grave implications for the understanding of time. In the cosmology of Aristotle and Ptolemy, a timeless realm of eternal perfection surrounds the earthly realm. Growth, decay, change, all the evidence of a time-bound world is restricted to the small domain below the sphere of the moon. Above it is perfect circular motion, unchanging and eternal. Now that the sphere separating the time-bound and the timeless was smashed, there could be only one notion of time. Would this new world be time-bound throughout, with the whole universe subject to growth and decay? Or would timeless perfection be extended to all of creation, so that change, birth, and death would be seen as mere illusions? We still struggle with this question.

Kepler and Galileo did not solve the mystery of the relationship between the divine, timeless realm of mathematics and the real world we live in. They deepened it. They breached the barrier between sky and Earth, putting Earth in the sky as one of the divine planets. They found mathematical curves in the motions of bodies on Earth and the planets around the sun. But they could not heal the fundamental rift between time-bound reality and timeless mathematics.

By the middle of the 17th century, scientists and philosophers confronted a stark choice. Either the world is in essence mathematical or it lives in time. Two clues to the nature of reality hung in the air, expectant and unresolved. Kepler had discovered that the planets move along ellipses. Galileo had discovered that falling objects move along parabolas. Each was expressed by a simple mathematical curve and each was a partial decoding of the secret of motion. Separately they were profound discoveries; together they were the seeds of the Scientific Revolution, which was about to flower.

This is not unlike the present juncture in theoretical physics. We have two great discoveries, quantum theory and general relativity, whose unification we seek. Having worked on this problem for most of my life, I'm impressed by the progress we've made. At the same time, I'm certain that some simple idea lies hidden in plain sight that will be the key to its resolution. Admitting that progress can be held up as we await the invention of nothing more substantial than an idea is humbling, but it's happened before. The Scientific Revolution launched by the simple discoveries of Galileo and Kepler was long delayed because of the idea that the universe was divided into an earthly and a heavenly realm. This idea prevented the thorough application of mathematics to the lower world, while our understanding of the upper world was thwarted by the belief that there was no need to look for causes of perfect heavenly motions.

It's thrilling to think about what might have happened had this basic conceptual mistake not blinded, for more than 1,000 years, the thinking of smart people who had in their hands the data and mathematics needed to take the steps that Galileo did. A Hellenistic or Islamic astronomer could well have made some or all of Kepler's discoveries from data available 1,000 years earlier than Tycho. The idea that Earth orbits the sun did not have to wait for Copernicus; it was on the table ever since it was proposed by Aristarchus in the 3rd century BC. His heliocentric cosmology was discussed by Ptolemy and others and would have been known to such great scholars as Hypatia, a brilliant mathematician and philosopher who lived in Alexandria from about AD 360 to 415. Suppose she or one of her brighter students had discovered Galileo's law of falling bodies, or Kepler's elliptical orbits?[5] There might have been a Newton by the 6th century, and the Scientific Revolution might have started a full 1,000 years earlier.

Historians may protest that Copernicus, Galileo, and Kepler could not have made their discoveries before the Renaissance prepared the way by freeing thinkers from the dogmatism of the Dark Ages. But in Hypatia's time, the Dark Ages had not yet descended and the struggle between the exponents of Greek learning and religious fundamentalism had not yet killed the spirit of rational inquiry. History may have been quite different if someone in Roman Alexandria, or, for that matter, the great centers of learning that flourished in the Islamic world a few centuries later, had done away with the geocentric universe. However, the brightest scientists in the best conditions could not make the conceptual leap of imagining mathematical laws governing motion in the earthly sphere or dynamic forces playing a role in the heavens. It took the shattering of the spheres separating the two realms for Galileo and Kepler to make their discoveries.

on blindness due to conceptual mistake

But even they could not take the next step, which was to see the unity lying in the earthly parabola and the planetary ellipse. That took Isaac Newton.

Because they lived after the shattering of the spheres, Galileo and Kepler could have asked whether throwing something hard enough leads to orbiting and the slowing of an orbiting object leads to falling. To us, it's obvious that these are not two phenomena but one. But this was not apparent to them. Sometimes it takes a generation or so before the simplest implications of new discoveries come into focus. Half a century later, Newton understood that orbiting is a form of falling and completed the unification of the heavens and the Earth.

One clue was a mathematical unity shared by the two curves that code motion. Ellipses trace the planetary orbits and parabolas trace the paths of falling bodies on Earth. These two curves are closely related: They both can be made by intersecting a cone with a plane. Curves that can be so constructed are called conic sections; the other examples are circles and hyperbolas.

The question for the second half of the 17th century was to discover the physical unity explaining this mathematical unity. The insight that impelled Newton to embark on the Scientific Revolution was about nature, not mathematics, and it was not his alone. Several of his contemporaries had perceived the great secret: *The force that*

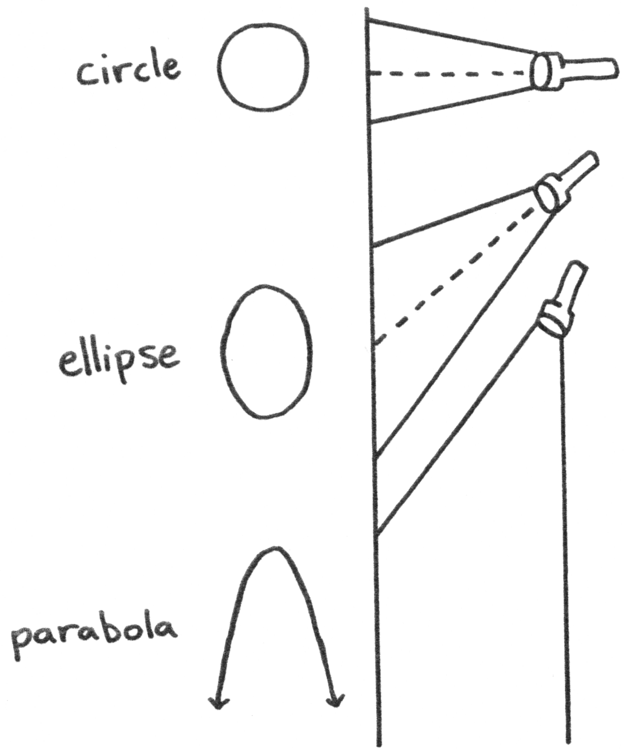

Figure 3: Conic sections illustrated with the image made by a flashlight on a wall.

impels everything on Earth to fall toward it is universal and acts also to pull the planets toward the sun and the moon toward the Earth. Gravity.

Newton, according to legend, had this epiphany while sitting in his garden noticing apples falling from a tree as he contemplated the motion of the moon. To complete the thought, he asked another crucial question: How does that force decrease with the distance between the objects? For decrease it must, otherwise we would be pulled upward to the sun rather than downward to Earth. And how does a force produce motion?

Others, such as Newton's contemporary, Robert Hooke, asked these questions, but the true accomplishment of Newton was in his answers

to them. The effort took him two decades and resulted in the theory of motion and forces we call Newtonian physics.

For our purposes, the most important thing about these questions is that they're mathematical. How a force decreases with distance can be specified by giving a simple equation. The right answer, as every first-year physics student knows, is that the force decreases proportional to the square of the distance. The astounding consequence for our conception of nature is that such a simple mathematical relation captures a universal phenomenon in nature. Nature did not have to be so amazingly simple—and, indeed, the ancients had never contemplated such a simple and universal application of mathematics to the causes of motion.

To ask how a force causes motion, you have to think of a moving object tracing a curve in space. The question then is how the curve differs depending on whether there is a force acting on it or no force. The answer is stated in the first two of Newton's laws. If there is no force, the curve along which a body moves is a straight line. If there is a force, the force acts to cause an acceleration of the body.

It's impossible to state these laws without mathematics. A straight line is an ideal mathematical concept; it lives not in our world but in the Platonic world of ideal curves. And what is acceleration? It is the rate of change of velocity, which is itself the rate of change of position. To describe this adequately, Newton needed to invent a whole new branch of mathematics: the calculus.

Once you have the necessary mathematics, it's straightforward to work out the consequences. One of the first questions Newton must have answered with his new tools[6] was what path a planet would take under the influence of a force from the sun that decreases proportionally to the square of the distance. The answer: It can be an ellipse, a parabola, or a hyperbola, depending on whether the planets travel on a closed orbit or make a one-time pass by the sun. Newton was also able to subsume Galileo's laws of falling in his law of gravitation.[7] Galileo and Kepler had thus seen different aspects of a single phenomenon, which is gravity.

There is little in the history of human thought more profound than the discovery of this hidden commonality between falling and orbiting. But beneath the enormity of Newton's accomplishment is an unintended consequence, which is that his work made our conception of nature far more mathematical than before. Aristotle and his contemporaries had described motion in terms of tendencies: Earthly objects have a tendency to seek Earth's center, air has a tendency to flee the center, and so on. Theirs was an essentially descriptive science. There is no suggestion that the paths along which objects move have any special properties, and hence they had no interest in applying mathematics to the description of motion on Earth. Mathematics, being timeless, was divine and applicable only to those divine and timeless phenomena we could see, which were only in the heavens.

When Galileo discovered that falling bodies are described by a simple mathematical curve, he captured an aspect of the divine, brought it down from the sky, and showed that it could be discovered in the motion of everyday objects on Earth. Newton demonstrated that the tremendous variety of motions on Earth and in the sky, whether impelled by gravity or by other forces, are manifestations of a hidden unity. The diverse motions are all consequences of a single *law of motion*.

By the time Newton had finished joining motion in the sky and on Earth, we lived in a single, unified world. And it was a world infused with divinity, because timeless mathematics was at the heart of everything that moved, on Earth and in the sky. If timelessness and eternity are aspects of the divine, then our world — that is, the whole history of our world — can be as eternal and divine as a mathematical curve.

3

A Game of Catch

TO ADDRESS THE ISSUES raised in the first two chapters, we need to know more about how we define motion. Nothing seems simpler: Motion is change in position over time. But what is position, and what is time?

There are two answers that physicists have given to the seemingly innocuous question of defining position. The first is the commonsense idea that the position of an object is defined relative to a landmark of some sort; the second is that there is something absolute about position in space, beyond its relation to something else. These are called the relational and the absolute notions of space.

The relational notion of position is familiar to all of us. I am now three feet from my chair. The airplane is approaching the airport from the west and is now two kilometers from the end of runway 1 at a height of 1,000 feet. These are all descriptions of relative position.

But relational position seems to leave something out. Where is the ultimate reference? You give your coordinates on Earth, but where is Earth? So many miles from the sun, in the direction of the constellation Aquarius. But where is the sun? So many thousands of light-years from the center of the Milky Way Galaxy. And so forth.

Proceeding in this way, you can give the position of everything in the universe relative to everything else. This is a lot of information, but is it enough? Is there not some absolute notion of position — of where something *really* is, behind all these relative positions?

This debate between relational and absolute notions of space runs through the whole history of physics. Roughly speaking, Newton's physics was a triumph of the absolute picture, which was overthrown by Einstein's relativity theory, which established the relational view. I believe the relational view is correct, and I hope to convince the reader of this. But I would also like to give the reader a vivid sense of why savants like Newton embraced the absolute view and what is given up when we reject it in favor of the relational view.

To appreciate how Newton thought about the problem, we have to ask not only about position but about motion. Let's leave time aside for a moment and apply what we have just discussed. If position is relative, then motion is change of relative position — i.e., change of position relative to some reference body.

All commonplace talk of motion is talk of relative motion. Galileo studied bodies that fall relative to the surface of the Earth. I throw a ball and see it move away from me. The Earth moves around the sun. These are all examples of relative motion.

A consequence of relative motion is that who or what is moving is always a matter of point of view. Earth and the sun move around each other, but which is really moving? Is the real story that the sun moves around an Earth fixed at the center of the universe? Or is it rather the sun that is fixed, and Earth that orbits? If motion is only relative, there can be no right answer to this question.

The fact that anything can be moving or fixed makes it hard to explain the causes of motion. How could something be the cause of Earth's motion around the sun if there is a different and equally valid point of view according to which Earth isn't moving at all? If motion is relative, an observer is free to adopt the point of view that all motion is defined relative to him. To resolve the impasse and be able to speak

of causes of motion, Newton proposed that there must be an absolute meaning to position. This was, for him, position with respect to what he called "absolute space." Bodies are moving or not, in an absolute sense, relative to this absolute space. Newton argued that it was the Earth and not the sun that moved absolutely.

The postulation of absolute space stops the infinite regression and gives a meaning to the location of every single thing in the universe, with no need to refer to anything else. This may be a comforting notion, but there's one problem. Where is this absolute space, and how would you measure the position of a body with respect to it?

No one has ever seen or detected absolute space. No one has ever measured a position that was not a relative position. So to the extent that the equations of physics refer to position in absolute space, they cannot be connected to experiment.

Newton knew this and it didn't bother him. He was a deeply religious thinker, and absolute space had a theological meaning for him. God saw the world in terms of absolute space, and that was enough for Newton. He would put it even more strongly: Space was one of God's senses. Things exist in space because they exist in the mind of God.

This isn't as strange as it sounds if you're a master decoder, as Newton was. He devoted years of work to searching for hidden meaning in the Scriptures, and as an alchemist he sought the hidden code for virtue and perhaps immortality. As a physicist, he uncovered universal laws that governed all motion in the universe but had previously been hidden. It was in character for him to believe that the essence of space was hidden from our senses yet seen by God.

Besides, he had a physical argument for absolute space. Even if position in absolute space could not be humanly perceived, some kinds of motion with respect to absolute space could be.

Children can't fly, but they can spin. And spin they do. Nothing matches the delight on the face of a child who has just discovered that she can make herself dizzy. Anytime she wants, over and over. Again! Newton had no children, but I like to imagine him being struck silent by the delight of his young niece, Catherine, spinning around in his study. Newton takes the wobbly, laughing child on his knee and tells her that her dizziness is a direct perception of absolute space. And absolute space is God. "What you feel when you feel dizzy is the hand of God upon you," he offers. She giggles, squirms as he starts to explain that she's dizzy not because she's rotating with respect to the furniture, or the house, or the cat, but because she spins with respect to space itself. And if space can make her dizzy, it must be something real. "Why?" she says, jumping off his lap to chase the cat out of the room. Let's leave Newton there, pondering gravity and mortality, and return to the question of how motion is defined.

When we say that something moves, we mean it is changing its position over time. This is common sense, but to be precise about it we need to be sure we know what we mean by time. Here we face the same *dilemma of the relational versus the absolute.*

Human beings perceive time as change. The time an event takes place is measured relative to other events—for example, the reading of the dial of a clock. All clock and calendar readings are relative times, just as addresses are relative positions. But Newton believed there is hidden behind change an absolute time, which God perceives.

Here's a taste of the debate that has raged over the issue of absolute time ever since. Newton's rival Gottfried Leibniz believed in God, too, but his God was not free, as Newton's was, to do as He pleased. Leibniz worshipped a supremely rational God. But if God is perfectly rational, then everything in nature must have a reason. This is Leibniz's principle of sufficient reason. One way to state it is that every question of the form "Why is the universe like this rather than like that?" must have a rational answer. There are, of course, questions for which there cannot possibly be any rational answer. Leibniz's point was that to ask a question that cannot have a rationally justified answer is to commit an error of thinking.

Leibniz illustrated his principle thus: He asked, "Why did the universe start when it did and not ten minutes later?" He replied that there cannot be any rational reason to prefer the history of the universe to one in which everything happens ten minutes later. All the relative times will be the same in both universes; only the absolute times are different. But the laws of nature speak only about relative times. Consequently, Leibniz argued, if there is no reason to prefer the universe to start at one absolute time rather than another, there can be no meaning to absolute time.

I accept Leibniz's reasoning, so whenever I refer to a time, I will mean a relative time. Indeed, although we can argue about whether there might be some transcendent sense in which absolute time exists, one thing that's certain is that we humans, living in the real world, have access only to relative times. So for the purpose of describing motions, we will consider time to be measured by clocks. For our purposes, a clock is any device that reads out a sequence of increasing numbers.

Now that we've defined both time and position, we can go on to measure motion: Motion is change of position, measured relative to some reference object, during a period of time, measured relative to the reading of a clock.

This brings us to the next, crucial step in our argument. To do science it is not enough just to make definitions and argue about concepts. You have to measure motions. This means using tools like clocks and rulers to associate positions and times with numbers.

Unlike absolute position, which is invisible, relative distances and relative times can be measured in numbers, which in turn can be recorded on a piece of paper or in a digital memory. In this way, observations of motion are turned into tables of numbers that can be studied with methods from mathematics. One such method is to make a graph of the records, thus turning the tables of numbers into pictures that can spark our understanding and imagination.

This powerful tool was developed by René Descartes and is taught to every schoolchild. It is undoubtedly something Kepler would have done as he struggled with Tycho's data on the orbit of Mars. In Figure 4 we see a graph of the orbit of the moon with respect to Earth.

In school, we learned a second way to draw a motion, which is to add an axis for time and graph the position against time. This repre-

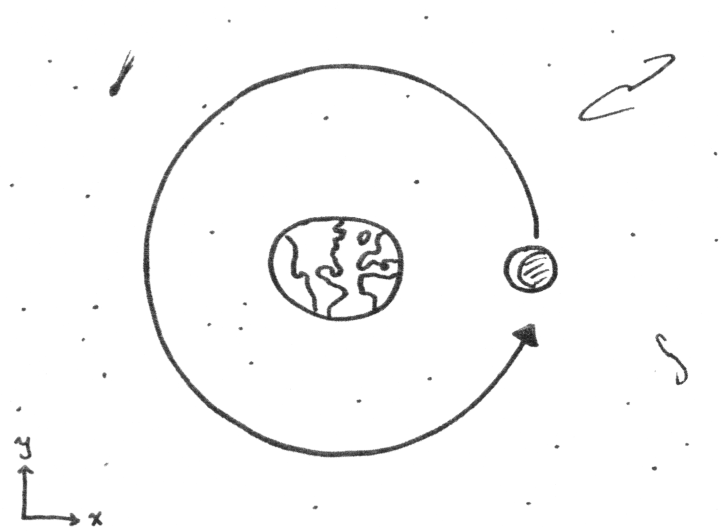

Figure 4: Graph of the moon's orbit around Earth.

sents the orbit as a curve in space and time, as in Figure 5. We see the moon's orbit now represented by a spiral; once it returns to its starting position, a month has passed.

Notice that by graphing records of observations, something wonderful has been done. The curve in Figure 5 represents measurements taken while something evolved in time, but the measurements themselves are timeless — that is, once taken, they do not change. And the curve that represents them is also timeless. By this means, we have made motion — that is, change in the world — into a subject of study by mathematics, which is the study of objects that don't change.

The ability to freeze time like this has been a great aid to science, because we don't have to watch motion unfold in real time; we can study the records of the past motions whenever we like. But beyond its usefulness, this invention has profound philosophical consequences, because it supports the argument that time is an illusion. The method of freezing time has worked so well that most physicists are unaware that a trick has been played on their understanding of nature. This trick was a big step in the expulsion of time from the description of

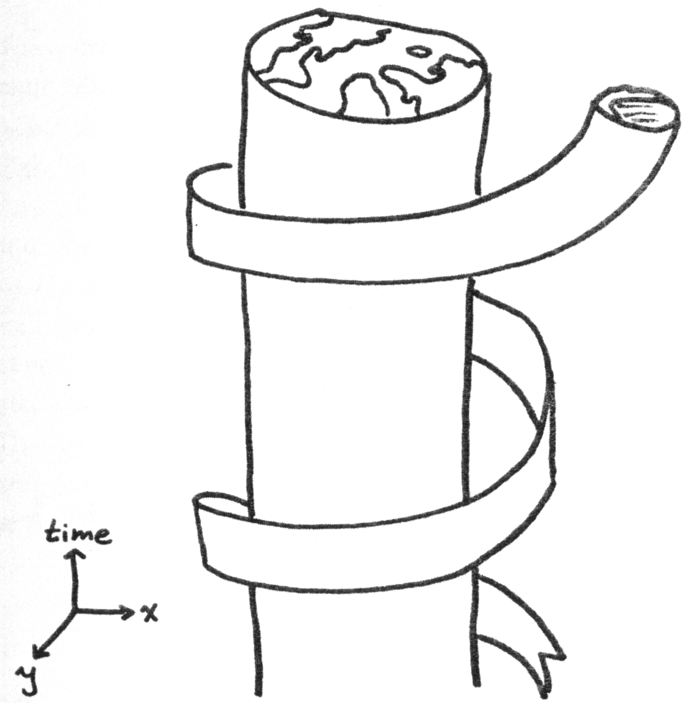

Figure 5: Graph of the moon's orbit as a curve in space and time.

nature, because it invites us to wonder about the correlation between the real and the mathematical, the time-bound and the timeless.

This correlation is so crucial that I want to frame it in an everyday example. All these weighty issues concern nothing more than what we can know about a game of catch.

◈

Around 1:15 P.M. on October 4, 2010, on the east side of High Park in Toronto, a novelist called Danny threw a tennis ball he had discovered that morning in his sock drawer to a poet he had just met, Janet.

To study Danny's throw through the eyes of physics, we do what Tycho and Kepler did for Mars. We observe the motion and record the ball's positions at a series of times; then we graph the results. To accomplish this, we need to give the position of the ball relative to some

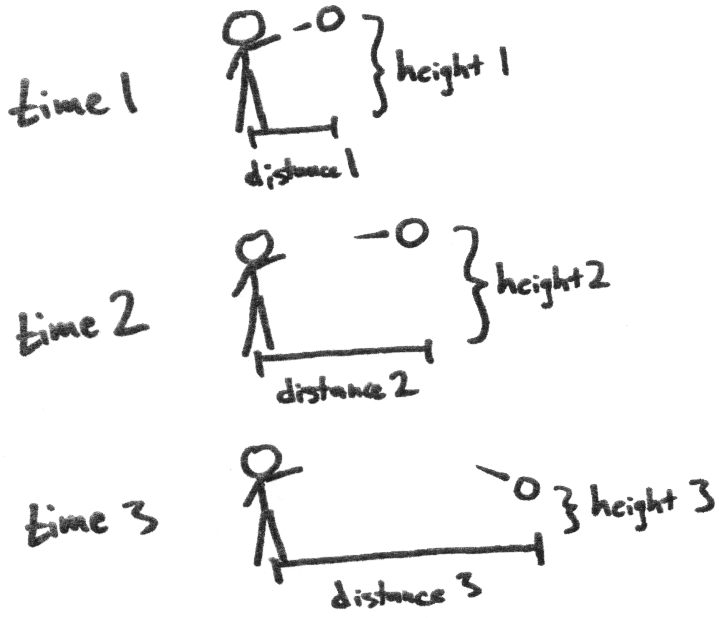

Figure 6: Danny's throw measured.

object, which we can take to be Danny himself. We also need a clock.

The ball moves fast, and this was a challenge for Galileo, but we can simply film Danny's throw and measure the position of the ball in each frame, along with the time of the frame. From the position of the ball in a frame we get two numbers, the ball's height off the ground and the horizontal distance the ball has traveled from Danny. (Space of course is three-dimensional, so we also have to describe the direction of Danny's throw. Other than to say he's throwing south, I'll ignore this complication here.) When we include the time of each frame, the record of the ball's trajectory is a series of triplets of numbers, one triplet for each frame of the film.

(time 1, height 1, distance 1)
(time 2, height 2, distance 2)
(time 3, height 3, distance 3)

And so on.

These lists of numbers are important tools if we are to study motion scientifically. But they are not the motion itself. They are just numbers, which are given meaning by the measurements we made of a ball in flight in a particular instance. The real phenomenon differs in several ways from the list of numbers that is its record. For example, many features of the ball have been neglected. We recorded only its positions, but the ball also has color, weight, shape, size, and composition. More important, the phenomenon unfolded in time: It happened just once and was gone, into the past. What's left is the record, and that is frozen, unchanging.

The next step is to graph the information in the record. Figure 7 is a picture of the path the ball made in space. We see that the ball flew on a parabola, just as Galileo predicted.

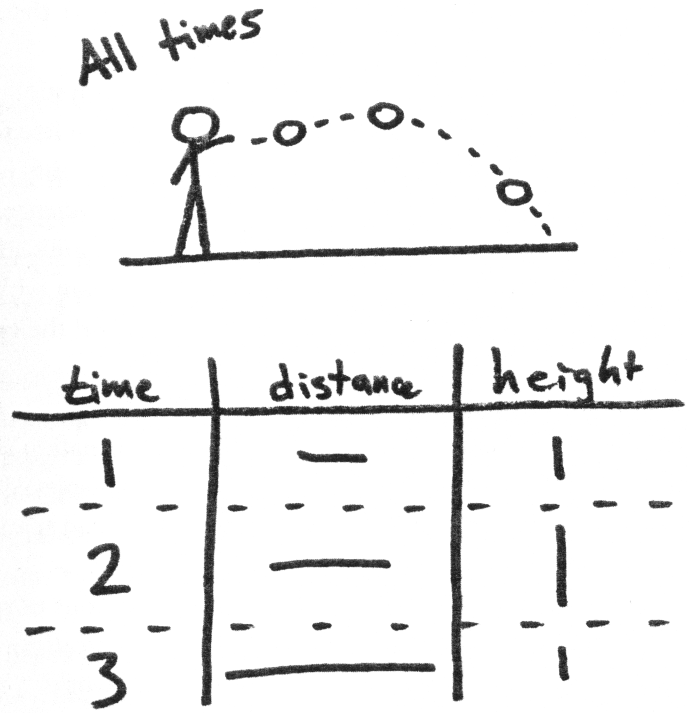

Figure 7: Danny's throw recorded and graphed.

We see again that the process of recording a *motion*, which takes place in time, results in a *record*, which is frozen in time—a record that can be represented by a curve in a graph, which is also frozen in time.

Some philosophers and physicists see this as a profound insight into the nature of reality. Some argue to the contrary—that mathematics is only a tool, whose usefulness does not require us to see the world as essentially mathematical. We can call these competing voices the *mystic* and the *pragmatist.*

The pragmatist will argue that there's nothing wrong with checking hypotheses about laws of motion by converting motion into numbers in tables and looking for patterns in those tables. But the pragmatist will insist that the mathematical representation of a motion as a curve does not imply that the motion is in any way identical to the representation. The very fact that the motion takes place in time whereas its mathematical representation is timeless means they aren't the same thing.

Some physicists since Newton have embraced the mystic's view that the mathematical curve is "more real" than the motion itself. The great attraction of the concept of a deeper, mathematical reality is that it is timeless, in contrast to a fleeting succession of experiences. By succumbing to the temptation to conflate the representation with the reality and identify the graph of the records of the motion with the motion itself, these scientists have taken a big step toward the expulsion of time from our conception of nature.

The confusion worsens when we represent time as an axis on a graph, as we did in Figure 5. In Figure 8, we see the information about the trajectory of Danny's ball including clock readings, displayed as if they were measurements made by a ruler. This can be called spatializing time.

And the mathematical conjunction of the representations of space and time, with each having its own axis, can be called *spacetime.* The pragmatist will insist that this spacetime is not the real world. It's entirely a human invention, just another representation of the record we

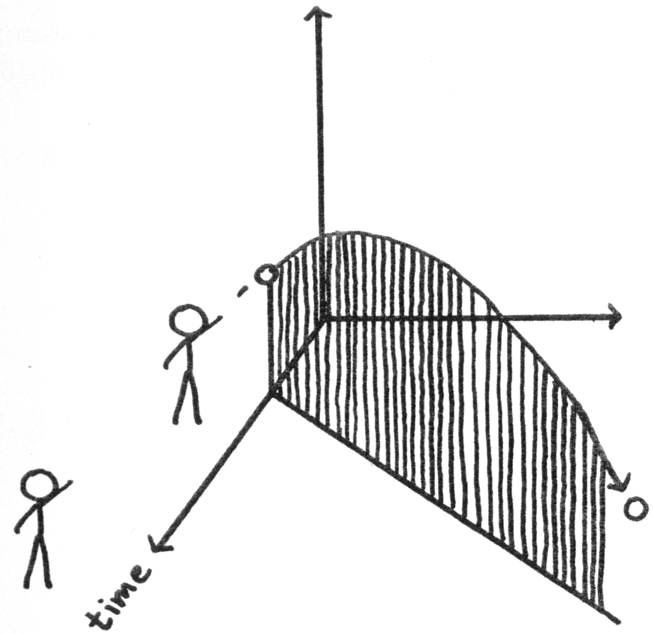

Figure 8: Danny's throw graphed as a curve in space and time.

have of the process of Danny throwing the ball to Janet. If we confuse spacetime with reality, we are committing a fallacy, which can be called the fallacy of the spatialization of time. It is a consequence of forgetting the distinction between recording motion in time and time itself.

Once you commit this fallacy, you're free to fantasize about the universe being timeless, and even being nothing but mathematics. But, the pragmatist says, timelessness and mathematics are properties of representations of records of motion—and only that. They are not, and cannot be, properties of real motions. Indeed, it's absurd to call motion "timeless," because motion is *nothing but* an expression of time.

There's a simple reason that no mathematical object will ever provide a complete representation of the history of the universe, which is that the universe has one property no mathematical representation of it can have. Here in the real world, it is always some time, some present moment. No mathematical object can have this particularity, because, once constructed, mathematical objects are timeless.[1]

Who is right, the pragmatist or the mystic? This is the question on which the future of physics and cosmology turns.

subjective, first-person experience of time!

It remains outside reach of science...

4

Doing Physics in a Box

IN HIGH SCHOOL, I was cast in Jean-Paul Sartre's *No Exit.* I played Joseph Garcin, a man locked in a small room with two women, all of us deceased. The play was an extreme version of a society in a box; it enabled the playwright to examine the consequences of our moral choices. In the climactic scene, I was to bang on the classroom door, shouting the famous line, "Hell is other people!" but the door's pane shattered, showering me with glass and bringing my acting career to a close.

Musical performance, like theater, allows a heightened examination of human emotion by isolating us in a controlled environment. As a teenager, I listened to a terrifying performance of my cousin's band, Suicide, in the basement of the Mercer Arts Center in Greenwich Village. The singer locked the doors and mesmerized the audience with a long aria of wanton murder, sung over a numbing repetition of the chords of the garage-rock classic "96 Tears." The atmosphere grew claustrophobic as the singer became increasingly menacing, but like the characters in *No Exit,* we were stuck. More recently, the insight-through-claustrophobia method has been adopted by conceptual art-