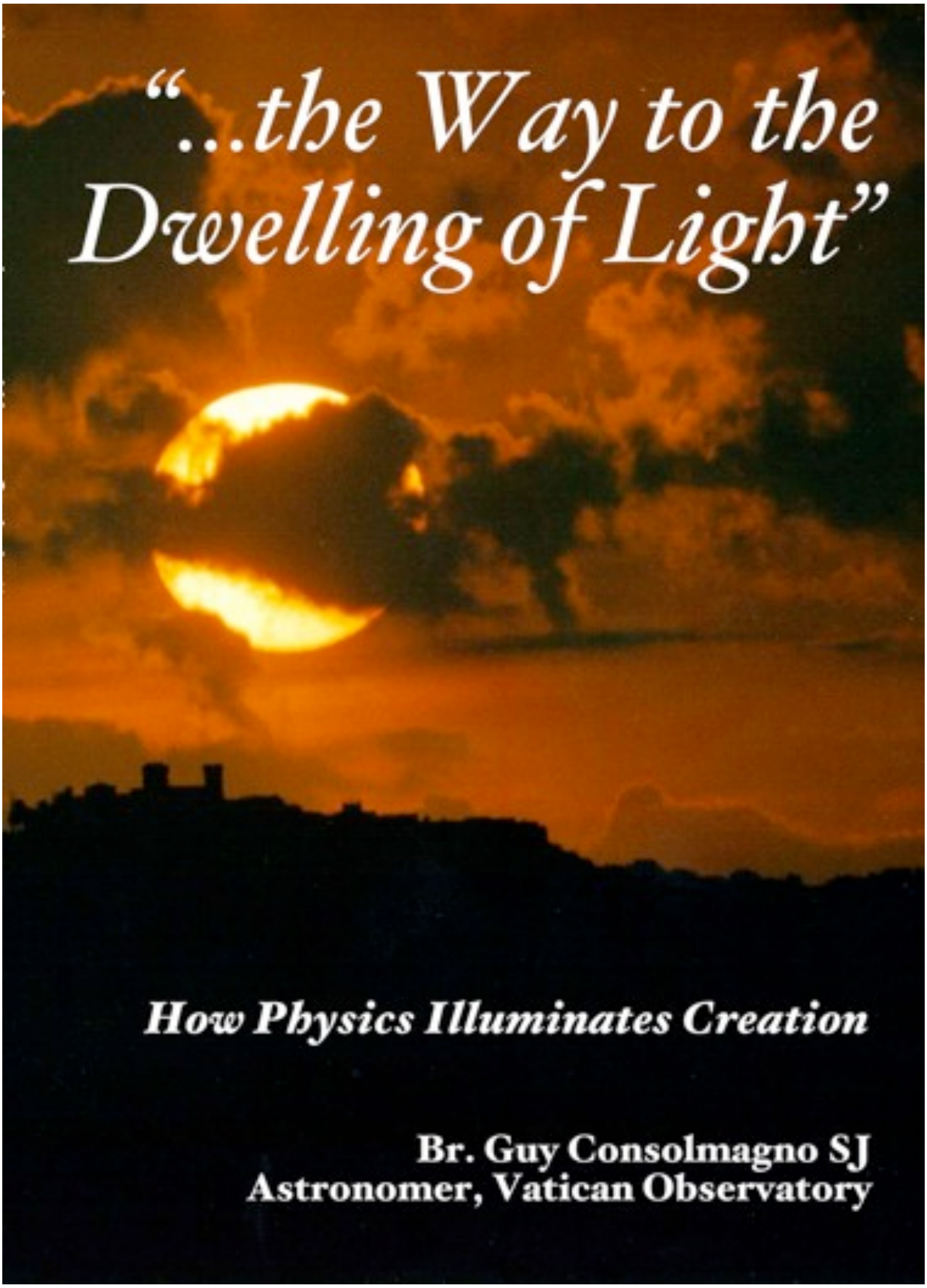

"...the Way to the Dwelling of Light"
How Physics Illuminates Creation
Br. Guy Consolmagno SJ
Astronomer, Vatican Observatory

The Way to the Dwelling of Light

How Physics Illuminates Creation

Guy Consolmagno SJ

Vatican Observatory

Illustrations by Su McTeigue, Pearl Bridge Studios

Second Edition, 2003

Published by the Vatican Observatory

Distributed (except in Italy and the Vatican City State) by:
The University of Notre Dame Press
Notre Dame, Indiana, 46556
USA

Distributed in Italy and the Vatican City State by:
Libreria Editrice Vaticana
V-00120 Città del Vaticano
Vatican City State

ISBN 0-268-01954-1

Contents

Foreword

This is not a book about religion. It's a book about physics.

However, you may well have found it in the "religion" section of your bookstore. That's because the people for whom it was written are precisely those who are more likely to browse the religion section than the science section.

Why did we put a physics book in among the religion books? Clever marketing, you might say — there are a heck of a lot more churchgoers in the world than there are scientists. But there's a deeper reason.

A lot of people talk about the "split" between science and religion, as if no scientist could believe in God and no religious person show an interest in Einstein. That's silly, of course. You obviously don't have to give up God to study His creation: most of the great scientists in history, including Copernicus, Kepler, Newton, James Clerk Maxwell, Marconi, and Einstein himself, called themselves believers.

Even Galileo endured the indignity heaped on him by his Church — my Church — and still remained a devout believer… though undoubtedly he was less than happy with his treatment at the hands of certain Cardinals and Popes. (Not all of them; immediately after his trial and recanting, he was welcomed as the honored guest of the Cardinal of Siena.)

And I discovered something surprising when I entered the Jesuit order, while still remaining a scientist; once my fellow scientists realized I was a churchgoer, many of them were delighted to talk to me about the churches *they* attended. In fact, from these conversations I'd estimate that the proportion of scientists in church on a given weekend isn't all that different from the general population as a whole.

But I have found that a lot of my fellow churchgoers are much less comfortable with science than the scientists are with religion. And, really, I can't blame them. Too often, the people working to popularize science have a non-science agenda that is distinctly anti-religious. I'm not talking about scientists in general here; I'm talking more about those who Want To Be Known As Scientists, which is a distinctly different breed of cat.

And that's a shame. I think we all, scientist or not, are curious about the world and how it works. And we're all curious to learn what the physicists are actually coming up with, presented without all the pseudo-philosophical folderol that you have to wade through in too many pop science books. (Most physicists have never had formal training in philosophy or theology, and it shows.)

If someone tells you that you have to choose between religion and science, and you are comfortable with religion but unsure of science, then I think it's pretty obvious which side you're going to take. And no one could blame you.

That's bad news for science — and not only because the churchgoers outnumber the scientists! It makes our job all the harder when we encounter this sort of hostility in the people we're trying to teach, both students and the people we're trying to advise on political issues involving space or ecology or defense, where scientific information plays an important role. And it makes it harder for us to recruit bright young people into the field.

But this popular division between science and religion is bad news for religion, too. Getting to know How God Did It ought to be a wonderful way of celebrating God's grandeur. It's traditionally been a form of worship that western religion had always embraced, until the late 19th century… when this canard of a split between science and religion took hold.

Science is too important to our lives to ignore. It's too much fun to leave to the atheists. And it's too Good not to be used as a way of getting to know the Lord Who Created Heaven and Earth; Who can be found, with His first creation, at home in the Dwelling of Light.

Introduction: Physics in the Pew

Then the Lord answered Job out of the whirlwind:

"Where were you when I laid the foundation of the earth? Tell me, if you have understanding! Who determined its measurements — surely you know! — or stretched the line upon it? On what were its bases sunk, or who laid its cornerstone, when the morning stars sang together, and all the sons of God shouted for joy?

"Have you commanded the morning since your days began, and caused the dawn to know its place? Have you entered into the springs of the sea, or walked in the recesses of the deep? Have the gates of death been revealed to you, or have you seen the gates of deep darkness? Have you comprehended the expanse of the earth? Declare if you know all this!

"Where is the way to the dwelling of light, and where is the place of darkness, that you may take it to its territory and that you may discern the paths to its home? You know, for you were born then, and the number of your days is great."

(The Book of Job, *Chapter 38: verses 1, 4-7, 12-13, 16-21)*

The poet Emily Dickenson, reacting against the arrogance of 19th century science, once complained, "Arcturus is his other name; I'd rather call him star… What once was 'heaven,' is 'zenith' now; where I proposed to go when time's brief masquerade was done, is mapped and charted, too." As she felt, there's something about the mystery, and the awesome power, of the physical world that can make trying to understand it seem sacreligious.

And isn't that what *The Book of Job* is telling us in the passage quoted above? "You know, for you were born then, and the number of your days is great." *The Book of Job* portrays a sarcastic God here, mocking Job's presumption. But underneath that sarcasm, I believe there's something more subtle going on. Ultimately, instead of merely laughing at us, the last laugh is one God shares with us. Because instead of letting us presume we know these things already, God is teaching us that only the humility of admitting our ignorance will ever allow us to approach the truth.

And here's the kicker: God is inviting us to do exactly what He says. We are invited to explore the length and depth and breadth of His incredible universe. After all, who created this universe, and what is it based on? I think God wants us to know. But can we really understand life and death, or comprehend the extent of space? Most definitely, yes — to some degree, at least. We were built to do just that, in order to appreciate its Maker.

We were given curiosity, and the ability to understand, by the Creator; and I suspect He would be disappointed if we didn't use our gifts. *"Ever since the creation of the world, His invisible nature, namely, His eternal power and deity, has been clearly perceived in the things that have been made,"* wrote St. Paul in the first chapter of his letter to his fellow Christians in Rome. By studying the things that have been made, Paul says, we study the Maker. The scientist who studies physics has accepted a divine invitation.

Studying the universe engages us in something bigger than ourselves. Science tries to describe, in terms we can only grasp intuitively, things that are beyond our intuition. Physics tries to make sense of the world, so that we might understand the Sense of the World.

What results from our study is a set of human-made descriptions of how the universe behaves. Like all human creations, it is certainly limited. But still, within those limits, it can be true; and the closer it is to the truth, the more beautiful it is.

And yet getting a handle on those theories about nature can be as intimidating as approaching nature itself. We've been raised in a culture that treats the practitioners of Science — especially Physics — as a priesthood of super-geniuses whose thoughts are far above us mere mortals. And even if we're willing to see beyond the foolishness of such misplaced reverence, we do have to admit that the physicists' way of looking at the world is so different from our ordinary day-to-day experiences that we have a hard time knowing even where to start, to try to understand what they are saying.

How can we relate to it? Where else in our lives do we attempt to deal with the ineffable, to describe the indescribable, to make sense out of a universe that at first glance can seem chaotic?

One place is in our religion. Those of us who live within a religious tradition can immediately recognize the parallels. Religion can be broken up into various components — the liturgies, the ritual practices where we encounter God like a scientific observer encounters nature; the theology, which tries to develop a theory (beautiful, but limited) of how God relates to us humans; and the moral laws which, like engineering, try to translate our theory into solving practical how-are-we-to-live-our-lives sorts of problems.

Within theology itself, we recognize that when we try to understand God, we're trying to describe the indescribable. We recognize that all our descriptions, ultimately, are inadequate. But we eventually develop an instinct for how far we can push a particular image of God and make it useful, and when we have to abandon that image.

These same instincts are exactly what we need to understand the universe through physics. Indeed, those of us who have been raised in a religion have a "leg up" on those who don't precisely because we have two important habits of thought that the more conventional skeptic lacks.

First of all, we have minds flexible enough to suspend our immediate scorn of any idea we hadn't thought of before ourselves, allowing ourselves to hear it out and see where it is trying to go, and what it is trying to do. We are familiar in our religions with ideas expressed in words that we know are inadequate; and so we are comfortable with knowing how far to push an analogy without being distracted when it (inevitably) breaks down.

Secondly, we have a belief, and an instinct, for truth. We believe that there is an objective truth out there, even while we recognize that we'll never grasp it completely. And we have encountered examples of that Truth; we have a sense of what it "feels like." It also means we have a sense of what non-Truth feels like, too; that gives us a bit of armor against ideas that are superficially appealing but with little basis in fact. Religion teaches us how to believe. And it teaches us to be wary of too much credulity.

(Many years ago I happened to attend a meeting of self-proclaimed "skeptics" where one of their own, an astronomy professor at the local community college, gave a lecture about The Big Bang. The reaction of the skeptics was amusing to watch. The idea of expanding space and time was more than their rough-and-ready common sense could handle. And they were outraged that the lecturer had to appeal to the "authority" of mathematicians and astronomers in order to make his case!)

I am an astronomer. I'm also a Jesuit Brother, a member of a Roman Catholic religious order. I write specifically as a Christian.

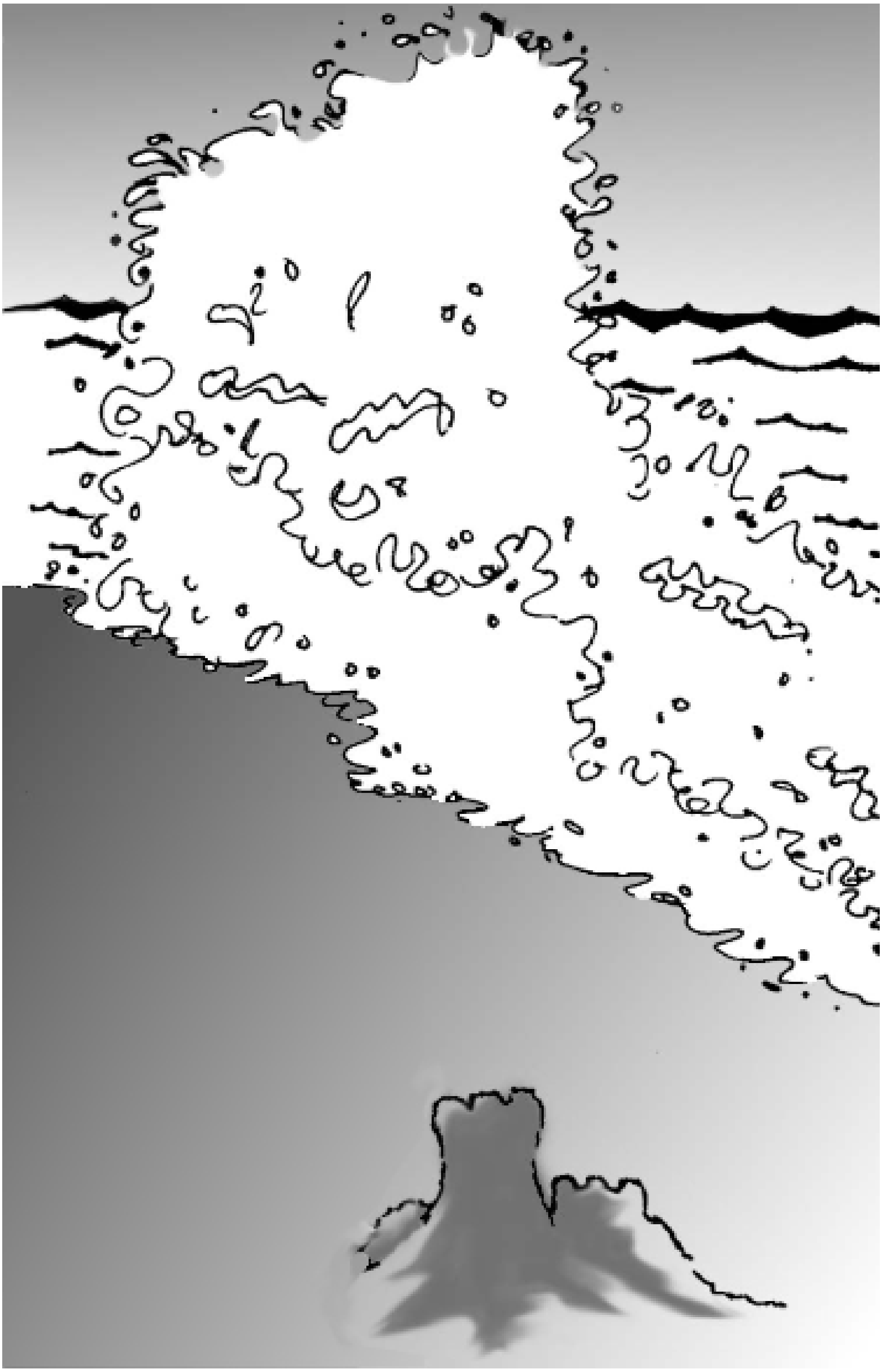
Within Christianity there are still many differing points of view; in this book I will try to stick to what C. S. Lewis once described as "Mere Christianity." By that, he meant the essentials held in common by most people who call themselves Christian, and not seen through a "scientific" filter or any other special point of view.

Most specifically, I do *not* espouse a "new" Christianity based on modern physics. Besides the patent absurdity of such a proposition, it would fly in the face of what I'm trying to do. I want to use common, ordinary Christian ideas as tools to understand physics; and like any good worker, the last thing I want to do is to bend my tools out of shape.

In the same spirit, what I'm presenting here might be called Mere Physics. This book deliberately avoids the fuzzy edges of the latest theories — "string theory" and "super-symmetry" and grand unified "Theories of Everything." First of all, they lie outside my expertise; in my twenty years as a planetary astronomer, I've rarely had to go beyond "mere" physics. Second, it's certain that half of what's being speculated about today will turn out to be wrong; the trouble is, we don't know yet which half. And finally, concentrating on the flashy new stuff misses the point of what I'm trying to do. What I really want to teach is not the facts of physics, but a way of looking at physics… indeed, a way of looking at the whole universe.

Besides, the concepts we will be getting into here, those parts of quantum theory and relativity and cosmology where there's a general consensus among scientists, are a strange enough trip all by themselves. It's a trip that will take us into the recesses of the deep… past the gates of darkness… to a comprehension of the expanse of the universe.

It's a way to the Dwelling of Light.



Chapter 1: Light

Part 1: "In the beginning…"

*God said, "Let there be light;" and there was light. And God saw that the light was good; and God separated the light from the darkness. God called the light Day, and the darkness he called Night. And there was evening and there was morning, one day. (*Genesis*, Chapter 1, verses 1-5)*

The book of *Genesis* opens with God saying "Let there be light." Now, a curious coincidence is that the modern Big Bang theory used by science to describe the origin of the universe also postulates that the universe started from a point of pure energy — light, if you will. (Light is energy.)

Seeing this coincidence for the first time, several generations of clever freshmen — I was one of them — have proposed that maybe the author of *Genesis* knew more than we credit him. Maybe God was whispering bits of astrophysics in his ear; or maybe ancient astronauts from another dimension visited Old Testament Palestine and let the cat out of the bag. How else could the author of *Genesis* have gotten the origin of the universe so right?

It's a cute idea; but it doesn't really work. For one thing, trying to force the rest of our scientific ideas for creation and the origin of life to fit the chronology of *Genesis 1* gets harder and harder as you step through the seven days of creation. But even worse, trying to play that sort of game runs the very real risk of missing the point that the *Genesis* author was trying to make. *Genesis* is not a modern science textbook; the author had no interest and no intention of writing such a book; and if that's the way you try to read it, it'll make about as much sense as reading the witches' scene from *Macbeth* ("… eye of newt and toe of frog, wool of bat and tongue of dog…") as if the play were a cook book.

What the author of *Genesis* tells us is that all the universe, even light itself, is God's creation. That's a profoundly important idea, with enormous implications both for our concept of God and of creation.

God as Creator of the Universe is in stark contrast to many of the other ancient religions neighboring Palestine. They suggested that the universe formed itself out of a formless chaos, or else existed in an eternal cycle without beginning or end. That's not what *Genesis* says. Rather, light was summoned at the beginning, by God… Who exists outside and above, beyond and before His Universe; Who is greater than His creation.

(Later books of the Hebrew Bible push their understanding of God as creator even further. By the time of the Books of *Maccabees*, a few hundred years before Christ, the writers explicitly understood that God did not merely rearrange things within an eternal, pre-existing chaos, but created everything out of nothing itself — *ex nihilo,* the medieval philosophers called it — simply by an act of the Divine will.)

It's this contrast with the surrounding pagan religions that is of interest to the writer of *Genesis*. He says: Our God is greater than other peoples' nature gods. Our God is *super*-natural.

And by choosing light as His very first creation, the author of *Genesis* is saying that God has nothing to hide. Unlike the gods of any other culture, the God of *Genesis* is inviting us all to watch, and understand.[1]

The point to remember here is that the Book of *Genesis* is not a book about nature, but a book about God. *Genesis* was written three thousand years ago, but it still has plenty of important things to say to us to-

---

[1]Is this why science has flourished especially in the West, among Jews, Christians, and Moslems? Obviously, it's more complicated that that. After all, good science was also done by the ancient Egyptians and Babylonians, and ancient Greeks, not to mention Hindus and Chinese and Mayan cultures. And there's excellent science being done in Japan and China today.

But if you have the pagan concept of a universe that's ultimately chaotic and meaningless, you won't see the point of "pure research" unrelated to making calendars or growing crops (an attitude still expressed in Congressional funding subcommittees nowadays). And if you don't have a sense of a supernatural God, then you might easily fall into the trap of worshipping science itself, much like the ancient Greeks (and 19th century Victorians) did.

day; and three thousand years from now, it will still tell humanity essential truths about God. In that sense, it is definitely not a physics book — I suspect that any modern physics book might look a little dated in the year AD 5000.

Indeed, *Genesis* never even attempts to answer the simplest questions of natural science. Such as: What *is* light?

Let's try an obvious answer first. Of course it won't work, but it's instructive to see just how it fails. We said it already: Light is energy. Fine. So… what's energy?

Energy is the motion of stuff; if something's moving, it has energy, and the more it moves the more energy it has. This definition is what the textbooks call "kinetic energy" and it's a perfectly fine definition. It's useful. You can quantify it. You measure the speed of something that's moving; and you can measure how much stuff you have moving. If you've got twice as much stuff moving at a given speed, it's showing you twice as much energy.

There are other kinds of energy, but you can always relate them to stuff in motion. For instance, heat is energy; but you can think of heat as the motion of molecules — each molecule is a little bit of stuff, and when you heat up the molecules you're just making each bit of stuff go a little faster.

Another kind of energy is called "potential energy"; we say a rock at the top of a cliff has more "potential" energy than another rock halfway down the cliff, because if you were to shove the rock off the top, by the time it fell past that other rock halfway down it'd be moving fast, not standing still — motion again. When it hits the bottom of the cliff, it stops, of course.

But in stopping, it kicks up a cloud of dust — more stuff moving. And it makes a loud noise — air molecules vibrating. And it gets hot — molecules moving again. (If you don't think it gets hot, ask a baseball catcher what his glove feels like at the end of an inning. Or try touching a nail after it's absorbed a few blows from an energetic hammer.)

What's more, something had to *move* that rock to get it to the top of the cliff in the first place… more motion again. Even electrical energy is generated by spinning turbines moving generator coils. Ultimately, all energy can be reduced to "stuff in motion."

The trouble with light is that we have the motion, but we don't have the "stuff." There's no way to measure the weight of light, because light has no weight. Very strange.

So what *is* light?

Most of the time, when we ask, "what is it?" we risk getting into all sorts of philosophical tangles about the meaning of existence, what does it mean "to be" and how do we know what we know… These are darn good questions, and a lot of fun to explore; but that goes beyond physics, and into metaphysics. (And it's the proper subject for a very different book from this one.)

To attack the problem directly, perhaps to say that light is an "essence" or an "emanation" or a "massless substance," is just to create a lot of meaningless words to substitute for the word "light." To be honest, that doesn't get us any farther, either.

Another way to attack the problem is to ask, "How does light behave?" This question is the way science actually operates. You never really understand physics, you just get used to it. So, once we can describe how light behaves, then we can look around and see if there's anything else in our experience, that we're already used to, that behaves the same way that light does. This at least will give us a way to get used to light.

You may recognize this trick. That's how religion teaches us about God. The Bible is not a theological tome that tries to define God in terms of an Essential Essence or The Source and Destiny of Our Transcendence[2]; it's a bunch of stories about God, set with commonplace backgrounds and familiar hooks, to let us get familiar with, and used to, this Being.

The best part of this approach is that in fact it does lead us to a much deeper and fundamental understanding of what light is. Physics eventually does have a surprisingly elegant answer to "what is light." Getting there from here will be an adventure, however… because in the process, we'll launch the revolution of quantum physics that completely changed the way modern physics understands the universe, and itself; and we'll open the door to Einstein's Theory of Relativity, on which all modern cosmology is based. By pushing with logic and common sense we wind up seeing that the universe, in its most fundamental parts, seems almost illogical and certainly far from commonly sensible.

It's a breathtaking ride; so, hang on.

Part 2: Light Bullets

The first thing we can say about light is that it starts from a source. Light doesn't just happen; we only get light when we have a fire, or a lamp, or a sun or star in the sky that gives off the light. Something to note is that, in our common experience, most of the things that give off light are *hot*. Heat is energy; so this fits the idea that the light emitted by hot things might be a form of energy, too.

Next, so far as we can tell it seems that light travels away from that source in straight lines. In our commonsense, ordinary day experience, we don't see light "getting tired" and coming to rest, or falling to the ground

[2] I'm not making this stuff up; German theologians are full of these kinds of words. Try reading Karl Rahner's book *Grundkurs des Glaubens* if you don't believe me.

This is not to make fun of Rahner; his books really are phenomenally profound. But they are deep reflections that come *after* you've gotten a gut-level intuitive familiarity with the subject. Most physics books are like that, too.

after a while, the way a rock behaves after we've thrown it. We don't see light pooling into ponds, like a fluid. We don't see it seeping through a crack into a dark room and evenly filling the whole room, the way that a substance like water or air might behave.

Instead, light travels in a straight line. When it meets a solid substance, it stops (though maybe it can get bent or reflected by a transparent substance, like water or a glass lens). And when it stops, it disappears, leaving only shadows behind.

Now, from our common sense experience we can compare the motion of light and the motion of other, more familiar stuff. Light is clearly different from a gas; a gas can fill containers, and it won't leave a shadow. Same is true of liquid, though at least with a liquid we can visualize the feel of a stream of stuff emanating from some source.

But what about a stream of solid particles?

Unlike solid particles, light doesn't fall… or does it? After all, if you throw a rock at a wall and then fire a bullet at the same wall, the rock falls farther than the bullet does before it hits. The rock is moving slower than the bullet, and so it takes a longer time to reach the wall, and so it has more time to fall. Maybe light does fall, a little, but it travels so fast that we can't see it. (In fact, this turns out to be true, but in a way much stranger than you'd ever expect at this point.)

Spray a bunch of bullets at a wall, and you can easily see that the wall creates a "shadow." Hide behind the wall and you won't get hit, but as soon as you stick your head above the wall — in the "line of sight," in fact — you become a target. So in that sense, bullets and light do seem to behave the same way.

After you've finished firing your bullets, you wind up with a wall full of lead. Do little bits of light stick to our wall after they hit?

A lot of them bounce off; sunlight pouring through a window makes a bright spot on the floor, but enough bounces off the floor that you can see into the rest of the room. Still, not all the light bounces off; the darker the floor, the less light there is bouncing into the rest of the room. So some light must be sticking to the dark floor. And the darker the floor is, the more light gets absorbed by it, just like bullets will bounce off steel but stick to wood.

Notice that the darker the floor is, the hotter it gets in the sunlight. A wood wall full of lead bullets gets heavier; a dark floor full of energy bullets gets hotter.

That's what gives us our first picture of light. Starting from a hot, energy-filled source, light acts like little bullets of energy that travel in straight lines (or virtually straight, anyway, depending on whether or not they fall) at very high speeds; bouncing off things or sticking to them, depending on how dark they are; and filling whatever they stick to with their energy.[3]

Now, a good scientific theory has to pass three tests:

First, it should explain the facts as they are known so far. Our bullet theory seems to do this quite nicely.

Second, it should be consistent with all the other theories of physics. We sort of guaranteed that, by the useful trick of taking a theory we already were comfortable with, how bullets work, and simply adapting it to the phenomenon on light.

So now we must ask the third test: Can we make predictions with our theory, and can we see if those predictions are true?

Our bullet idea predicts that, like bullets, light should bounce off surfaces with a predictable regularity. Shoot a bullet at an angle off a flat steel wall, and it will ricochet at the same angle that it hits. It's the same principle as a bank shot on a pool table (unless you're fancy, or sloppy, and put a spin on the ball). And light beams work the same way. That's how mirrors behave. Our theory works.

What about transparent substances? Well, somehow they must be things that our light-bullets can pass through, like a regular bullet goes through tissue paper. That also works. How come lenses can bend light?

---

[3] This theory of light dates back to Isaac Newton, around 1670, who referred to his "light bullets" by the much more scientific-sounding name of "corpuscles."

You can explain this, with a little arm-waving, by suggesting that our light-bullets get slowed down as they pass through glass; and by working out the different speeds of light in various substances, you can actually predict just how much light should bend, and how this bending should depend on the angle of the light shining on the glass.

And the predictions work.

Light is the way that energy is carried, in little packages, from one place to another. Light-emitters spew out little bullets of light; the light-bullets travel in straight lines through transparent substances (like air) or empty space until they hit something; and then they're either bent, or reflected, or absorbed.

Is this picture really true? Or, at least, has it brought us any closer to really understanding what light is?

It's such a crude concept… little "bullets" of energy (whatever energy is) whizzing about. One is tempted to be very skeptical of the whole image. We know what real bullets are like; seeing them, haven't we merely invented in our own mind a more ethereal kind of bullet, one that somehow can't be felt or touched, whizzing around at enormous speeds?

You can almost hear the sarcasm of a science skeptic: "Clearly, in the natural order of bullets, entities like 'light bullets' are impossible. They are the stuff of a fantasy writer's dream. Just because in some superficial way light seems to act like bullets doesn't at all prove that light really bears any resemblance to lumps of lead shot from guns. The idea is silly; it's childish; it's ridiculous. No practical grown-up person would believe in such a thing.

"And why bullets? What's this fascination with guns and violence? Probably there's something sexual beneath this image, something you're trying to repress!

"In fact, I wonder if there is any such thing as light at all. We may see things that are illuminated, or things giving off illumination; but who has ever seen pure light itself?" Thus speaks the agnostic of light…

In fact, our crude concept of light here really is just a first guess. It bears a certain similarity to the original pagan (and *Genesis*) picture of God as an angry father or an arbitrary king, powerful and terrible to behold, jealous and impatient, whose will was to be obeyed without question simply because He was indeed King.

Four thousand years of religion later, we understand God a little bit better than that. The Bible alone has given us many different pictures of God, some representing God like a person and others not; the very first image of God in the Bible, *Genesis 1.1*, pictures God as a breath of wind blowing across the waters. In fact we know that all such pictures can never be more than crude approximations. And yet, the image of God as Father or King still carries a great emotional punch even today. Such a picture still serves to point us in a good direction towards understanding God; one direction among many possible ways of looking at God, to be sure, but a useful and important one nonetheless.

The same holds true, in a much more mundane way, with our picture of little energy bullets of light. True, in one sense the picture solves nothing at all; we still don't know what the "substance" of light is all about, since we don't have any clue yet in this picture what any given "bullet" is made of. And yet the picture does tell us three important things.

First, it says that there *is* an entity called light. Our skeptic's arguments to the side, this picture insists that there's *something* that whizzes out of the Sun, bounces off the flowers, and winds up in our eyes so that our brains can grasp the size and shape and position and even the color of those flowers. That something, light, has an existence while it travels that is independent of Sun or flower or eye.

Second, we've learned some ready rules for predicting light's behavior: it travels in straight lines, it bounces off mirrors, it gets bent by lenses. The optician's technique of "ray tracing" to design mirrors and lenses comes directly out of this picture of light as whizzing bullets. It worked in the 17th century, and it works today. Ray tracing tells us how to shape our mirrors, where to put the lenses to make eyeglasses or telescopes that focus properly, and even lets us calculate ahead of time how much magnification we'll get from a given combination of lenses and mirrors.

And finally, and most profoundly, *this concept leads us to make predictions and ask new questions* about how light should behave, in ways that we might not have thought of before. When you're dealing with ineffable truths, even incomplete descriptions — even woefully wrong ones — can be important steps on the road to deeper knowledge.

For instance, let's use this picture and ask, why does light come in different colors? Why do rainbows and prisms always give us a spectrum of colors, always in the same order, ranging from red to yellow to green to blue, ending in deep violet? What is *this* trying to tell us about light?

Light is made of little *energy* bullets, according to our hypothesis; so perhaps, the difference between red light and blue light measures a difference in the amount of *energy* carried by any given bullet.

Indeed, inspired by this thought, a clever experimenter could put a thermometer into the rainbow of colors made by white light going through a prism. If you do so, what you find is that the blue light is hotter than the red light. The closer in the rainbow spectrum that any color of light is to blue or violet, the hotter it makes your thermometer.

You can even use this thermometer to detect invisible colors of light — infrared on the side of the spectrum before you get to red, ultraviolet on the side past violet — colors that the human eye can't see. Astounding! Who'd have thought there might be *invisible* light? You wouldn't have, without having a clever model for what light is. But carry out the experiment suggested by our model, and there you find it.[4]

---

[4] We use infrared lamps (which tend to shine red as well) as heating lamps; and as pieces of wood or metal get heated past the infrared point, they start glowing with visible red light. This leads to the popular idea that "red" is "hot." Compared to infrared, it is. But something glowing yellow or white is even hotter, and a metal that glowed blue would be dangerously hot. Indeed, ultraviolet light is so energetic that it can actually rip molecules apart; it causes sunburns, and skin cancer.

Let's push this idea even farther. Colors arise because one light-bullet carries a different amount of energy than another. How does one bullet become more energetic than another?

We know that if two lead bullets are traveling at the same speed, the one with more mass will carry more energy. So, do light bullets come in differing masses? No. As we mentioned above, light doesn't seem to have any mass at all. So that doesn't work.

On the other hand, given two identical bullets moving at different speeds, the one moving faster carries more energy. So perhaps "light bullets" of different colors travel at different speeds. Does that work?

Back in the 17th century, when many of these ideas were dreamt up by Isaac Newton, it wasn't possible to measure the speed of light with any high precision, so that was still a reasonable idea. By the end of the 19th century, though, when Michaelson finally came up with a very precise way of measuring the speed of light, he could show conclusively that this last idea doesn't work at all. All light, of all colors, travels at the same speed.

They have no mass; they all have the same speed; so how can different bullets carry different energies? Clearly, our model is beginning to fall apart here. We're missing something essential.

Michaelson's light-speed measurements in the 1870's might have been the death-knell of our light-bullet hypothesis, except that, by then, no one believed in this "corpuscular" theory of light-bullets anymore, anyway. Fifty years before Michaelson, another careful observation of light had shown that it behaved in a fashion very different from bullets. A new theory had taken hold to describe the behavior of light.

Part 3: Interfering Waves

*In the beginning… the Earth was without form or void, and darkness was upon the face of the deep; and the Wind of God was moving over the face of the waters…(*Genesis *1.1)*

When I was a small boy, I spent my summers on the shores of Lake Huron. The lake shore was a great place for a curious kid with a philosophic bent of mind. I spent hours every day digging in the sand, building castles and roadways and waterworks, then watching the waves come in and sweep them away. And I was filled with all those little-kid questions that somehow never get answered…

Where do the waves come from? Where do they go after they've hit the shore? And… for that matter… what is a wave, anyway?

A wave is a most peculiar thing. Indeed, it's not even a "thing." You can describe it, you can point to it, you can see the effects that it has on boats and sand castles; but if you try grabbing hold of a wave, all you're left with is a handful of water.

Waves are ephemeral; but very real. Energy, not substance. Ineffable. They travel silently from one place to another, and as soon as they arrive they disappear. Reminds me of a cat I once knew.

But does anything else come to mind?

By 1800 there was an excellent proof that light behaved exactly like waves: the phenomenon known as *diffraction*.[5]

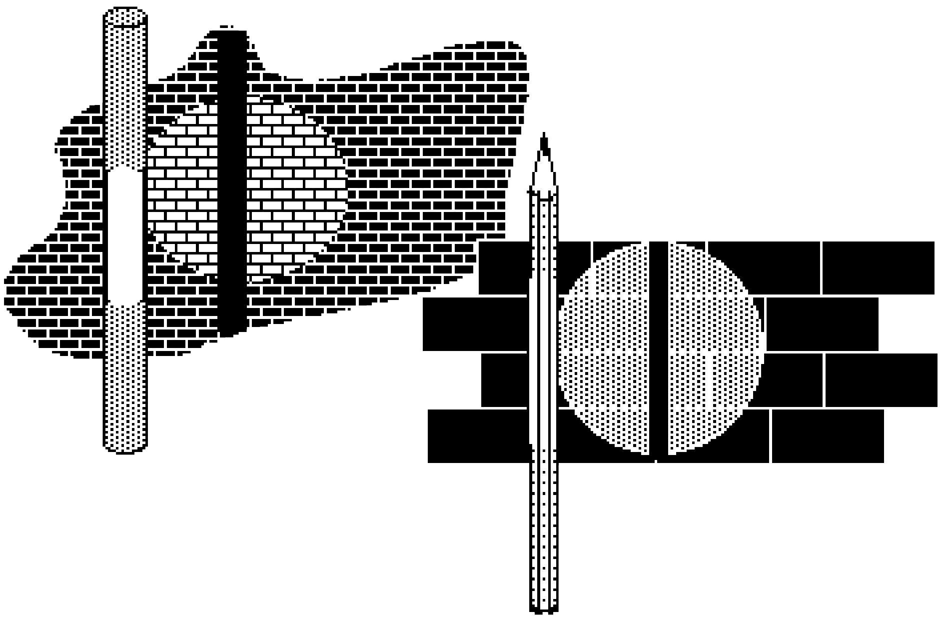
Here's how it works: Take a beam of light, with precisely one color and traveling in a single direction. (Nowadays, it's easy to produce such "monochromatic, collimated" light: just use a laser. Before lasers were invented, you needed a series of filters and lenses and narrow slits.) Now

---

[5] The name "diffrecation" was given to this phenomenon by Fr. Francesco Maria Grimaldi SJ. His book, *Physico-mathemesis de lumine (The Mathematical Physics of Light)* came out in 1665, predating Newton's work by nearly a decade. He talked about the diffraction of light — *diffractio luminis* — and interpreted its cause as "*undulatio minutissima criptata*" or "very tiny hidden waves." It took another 150 years for the rest of science to catch up.

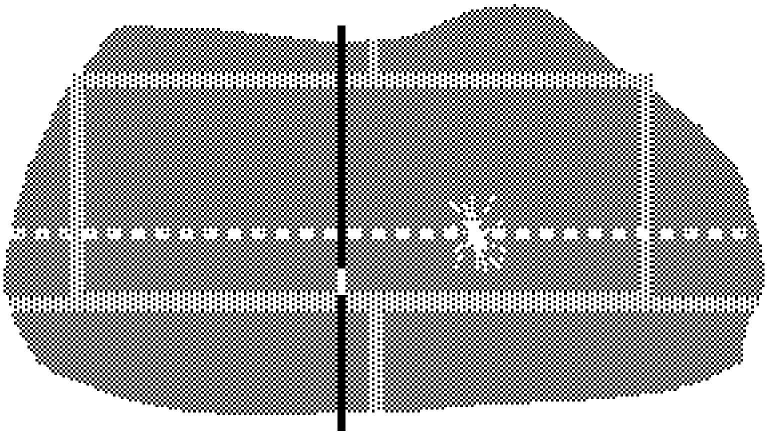
take this beam of light into a dark room, and shine it on the wall. You'll see a spot of light where it hits. So far, so good.

Now, put something in the beam of light to cast a shadow. If you have a big spotlight shining past a lamppost and on to the wall, you'll get a thick black streak on the wall where the lamppost casts a shadow. Shine a flashlight-sized beam on a pencil, and all you get is a thinner shadow. This much can be explained easily by our old "light-bullet" theory.



Next, pluck a hair from your head and hold it in a very narrow beam of uniform, monochrome light. If light behaved like a stream of bullets, you'd expect to see a very thin black shadow where the hair interrupted the beam of light, just like with the other cases.

But that's not what you see. Instead, the light is broken into a zillion little stripes, alternating light and dark, that spread away from either side of the hair:



This phenomenon is called *diffraction*. And it drove physicists to distraction. What in the world was going on?

Ah! you say, trying to rescue the bullet theory. Maybe the trouble now is that the bullets are hitting something that's smaller than they are; maybe they don't get absorbed but just bounce around the hair.

Two things are wrong with that idea, however. First, it doesn't really explain the alternating stripes. And second, if we take a card, make a slit in it that's as narrow as the hair, and shine the light on the slit, the light can still pass through. So much for our idea that the bullets are bigger than the hair or the slit.

And, incidentally, the light that does get through our narrow slit makes exactly the same pattern of light and dark stripes as the light passing around the hair. Very strange, yes?

Actually, it's not so strange. In fact, turns out it's exactly what you would predict — if light were made of waves rather than bullets.

First, go back to the shadows made by the pencil and the lamppost. Look very closely at the edge of the shadow. You'll notice that, while it looks like a sharp edge from a distance, when you look up close the edge of the shadow is never perfectly sharp.

Try looking at any shadow with a magnifying glass and you'll see what I mean; it's always a little bit fuzzy. In the same way, the image made by a lens is never perfectly in focus. You might think that's just because we never have perfect lenses… but what's imperfect about a shadow?

Maybe the shadow is fuzzy because the edge causing the shadow is rough? Well, look at the shadow of the smoothest, sharpest edge you can find… a knife edge, a razor's edge. Still fuzzy.

Maybe the trouble is that the light comes from a source that's too spread out? Well, try to use a light source as small and as far away as possible, down to a tiny brilliant point.

You may want to work in a completely dark room, with just one beam of light.

You may take a powerful desk lamp, cover that lamp with a thick piece of cardboard, and only let one pinprick of light through.

Maybe aim the beam of light with a lens, and then pass it through a series of cards with pinholes lined up exactly in a line. (This is the same thing as the making the light "travel in one direction" as we mentioned before.)

You can go to great lengths to get a finer and finer beam of light. And…

Well, if you do all that work, you're beginning to get a flavor of what an experimental scientist goes through. You're trying to study one particular effect, in this case the shadow of light; and all sorts of other things keep getting in the way, like rough edges or light sources that are too big.

So you keep trying to simplify your system. You get rid of anything that will interfere with the one particular effect you're looking at — whether it is extraneous sources of light, or the family cat that you keep tripping over. Rather than sitting in an easy chair, thinking about light, you're running around trying to create an extremely artificial set of conditions to see a particularly odd and hard-to-see phenomenon.

Only extreme situations are clear and simple enough to give clear and simple and unambiguous results. Of course, that's life. It's easy to be a hero — or, at least, to know what a hero ought to do — when faced with a massive Evil. It's doing the right thing day after day, when that involves nothing more romantic than taking out the garbage or being nice to your in-laws, that actually is a whole lot tougher.

So, you take heroic measures and set up your light experiment with painstaking detail. If you get everything right, all the conditions as perfect as humanly possible, it all helps. But you'll still never see a perfect shadow. It's always a little fuzzy. The light seems to bend itself just a little bit around the edge of the razor.

Why? What's going on here?

Go to a yacht harbor during a storm. (The great proponents of the wave theory were Englishmen and Dutchmen, well familiar with small harbors and stormy seas.) The waves of the ocean come crashing against the harbor wall, but inside the harbor the ships are peaceful. The harbor wall shields them from the waves; the boats are in the shadow of the wall. And yet the wall is not a perfect shield. The waves do tend to bend, a little, around the wall, and send their ripples throughout the harbor.

Light as waves? Okay, let's think about this for a minute.

Think of the waves on a beach… each wave is a line of spume, approaching the shore. Now think of waves in a still pond, where you've just dropped in a pebble. Here, the waves are perfect circles. So which is it — do waves radiate out from a point in circles, or travel in long straight lines? Obviously, they do both. How?

Christian Huygens, a 17th-century Dutchman, saw a clever way that you could turn circles into straight lines. Let's assume, he said, that a wave radiating away from a point will create circles. If you have a line of points, each radiating circles, then all those circular waves will run into each other. In principle, there are two extreme cases. If two identical crests run into each other, you'll get a new crest that's twice as big. (Same happens when two troughs add together; you get a trough twice as deep.) But when a crest meets a trough, they cancel. You wind up getting no change at all. This phenomenon has the name "interference."

This interference business can be a pain at times. If you aren't careful how you design an auditorium, you may wind up with spots in the audience where the sound waves bouncing off the walls cancel themselves, so the person sitting at such a spot can't hear what's happening on stage; or, even worse, they can reinforce some wavelengths and cancel others, so that some musical notes get overemphasized while others disappear.

On the other hand, wave interference can be useful, too. On an island off the coast of New Jersey there is a forest of antennae, first erected in the 1920s, designed to send radio signals to Europe. Why so many antennae? The problem was that it takes a lot of power to send a radio signal 8,000 miles across the Atlantic. But a single radio tower would send that signal not only to Europe, but also to the North Pole, to the Amazon Jungle, to Hawaii, and to any other spot within 8,000 miles of New York. What a waste of power!

But with even just two antennae, much of the problem is eliminated. How? Both would broadcast the same radio waves; but if the antennae were placed east-west one quarter of a wavelength apart, and the east one sends out its signal a quarter of a wave time after the west one, then waves heading east (towards Europe) would be reinforced while waves heading west (towards Hawaii) would be cancelled. Adding more and more antennae, at just the right locations, allows the direction of the beam to be tightened more and more.

In particular, if you add up an infinite line of points, all cheek by jowl next to each other, Huygens reasoned that all the crests traveling in nearly all directions would eventually run into troughs that cancel them out. The exception is where the crests meet along a new line, parallel to the first line, one wavelength downstream.

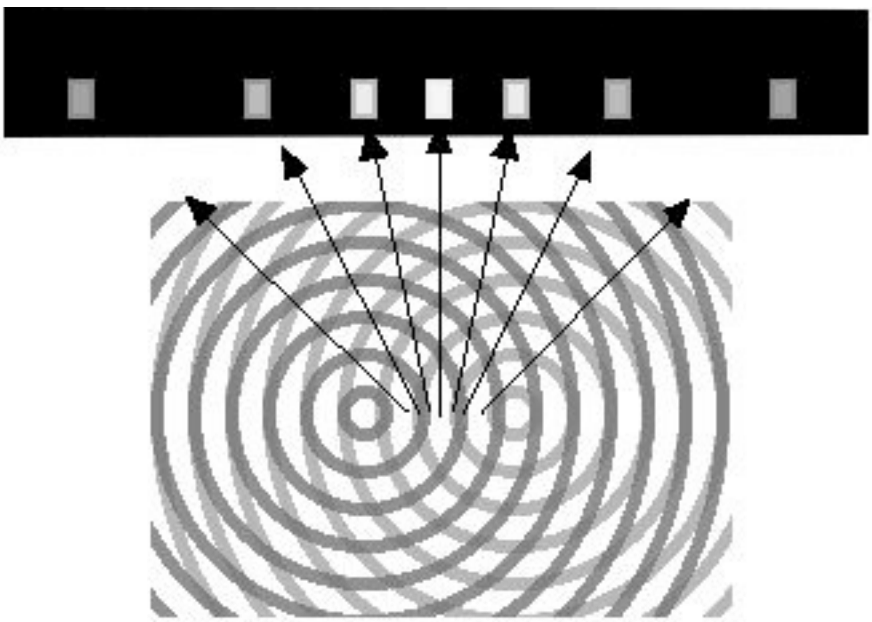
Waves add; and all waves, no matter how the wave front is shaped, can be described as if they were made of a string of point waves, radiating in circles and interfering with one another.

So what happens when a line of waves are cut off by an obstruction? All the points in the middle of the line still work together to make the line of the wave propagate downstream; but the point at the end of the line, the place where the line's been cut off, is still radiating like a circle. So the waves, just from those regions at the ends, make a wave that bends around the wall.

In fact, if you have two points — if the line has a beginning as well as an end — Huygens' idea predicts an even stranger phenomenon. The two points will both be radiating waves in circles.

If the two points are close together, there will be streams radiating out from the points where their waves interfere either to make bigger crests and troughs, and between them other streams where the waves cancel each other out.

When these streams hit a wall, you get a spot with lots of waves; then, next to it, no waves; then, another spot with waves…



That's exactly what we saw when the light went past our hair.

The hair cut the beam of light, and there were two points — one on each side of the hair — where the light propagated out in circles.

In fact, you can draw the paths of the light very carefully (or calculate, using high school trigonometry) and eventually, by knowing how thick the hair is, how far away the wall is, and how far it is from spot to spot, you can calculate what the wavelength of the light must be. For visible light, it turns out to be incredibly small… less than a thousandth of a millimeter.

So the Huygens model of waves does a great job of explaining something about light that our old light-bullet theory couldn't do. It explains why shadows aren't sharp, and why narrow hairs or slits don't leave shadows but instead produce diffraction patterns.

It even does more than that. If light is made of waves, then all the waves may be traveling at the same speed — as we know light does — but the *wavelengths* could vary from wave to wave.

If each wave crest has the same size, intuitively it ought to take more energy to push a lot of wave crests past a point at a very high frequency (so that the length from wave crest to wave crest is very short) than to push just a few well-spaced waves. So the energy in a wave — the color of the light — should be telling us the wavelength.

We can test that, too. Our experiment with the thermometers in the last chapter said that blue was hotter than red. Our theory predicts that, therefore, blue light should have a shorter wavelength than red light. We can put beams of different colored light past our hair by the wall, and see if the diffraction spots occur at different positions.

The answer? They do!

And their positions change just the way you'd predict, for blue light does indeed have a shorter wavelength than red light.

So, we've done it. Light is an ineffable something that we can see traveling, but that we can never catch; that bends around corners, ever so slightly; that forms diffraction patterns. And waves are ineffable somethings that we can see, but never catch; that bend slightly around corners; that form diffraction patterns. Obviously, light is a wave.

Obviously.

There's a huge flaw in this logic. Can you see it?

Common sense has always told us that, if something looks like a duck and quacks like a duck, then… you know what. But even common sense knows that looks, and quacks, don't *prove* that it's a duck. That's why we also look for duck eggs.

Or, put it another way: the Sun is yellow, and bananas are yellow; does this prove that the Sun is a banana? Of course not. Just because two objects have some properties in common, it doesn't prove they're the same thing.

Well, then, what does prove it?

The answer, surprising as it is, comes from a branch of philosophy called epistemology: the study of truth. Pontius Pilate was no dummy when he asked Jesus, "What is truth?" It's the biggest question any philosopher can ask. And, like a lot of other philosophers, Pilate didn't get an answer that satisfied him.[6]

So, what would *prove* that light is made of waves?

Nothing.

It's impossible. You can convince me, with an overwhelming pile of similarities, that light acts so much like a wave that there's no reasonable doubt about it; but that's still not proof.

And one piece of evidence that shows light behaving in an un-wave-like fashion is enough to topple the whole pile of evidence. After all, recall that's how we decided the light-bullet theory didn't work. Even though there certainly were lots of ways that light acted like bullets, it couldn't explain the diffraction patterns we saw.

Science never knows anything with complete certainty. And everything, *everything*, assumed and asserted by science is open to correction.

Remember that, the next time you hear someone claim that thus-and-so has been "scientifically proved." Anything that's been "scientifi-

[6] "Quid est veritas?" asked Pilate, assuming he spoke Latin to Jesus. A clever monk once noted that if you rearrange the letters of this question you can get the answer, "Est vir qui de sta' "— which is bad Latin for, "It is the man who stands before (you)."

cally proved" is therefore by definition open to question and doubt. That's how science and its proofs operate. Science by its nature does not deal with absolute truth.

Does this mean that anything's possible? Not so. Science *can* tell us when a specific statement is false. Perpetual motion machines are impossible, depend on it. But even these statements of falseness are only qualified, not absolute. The idea that "mass is neither created nor destroyed" we now know must be altered to account for nuclear energy, for instance.

Our understanding of the physical universe is always incomplete. And, of course, science can *never* prove or disprove matters of taste or beauty, of love or hope or faith.

By contrast, the nature of light is a subject of the physical universe that science is expressly designed to deal with. By observing behavior like diffraction, we can make state with great confidence that light is *not* made up of bullets.

But whatever theory we do come up with had better be able to explain all the things the bullet theory explained… and more. And that, at least, is what our wave theory does. So that's why we'll stick with it.

Until it fails.

Every scientific theory fails eventually; after all, science is merely a human creation. (That's why it's so foolish to base religious faith on science.) But, ironically, once we know exactly when and where and how our theory fails, we'll have all the *more* confidence in it.

Huh?

Well, yes. Because, once we know just where it fails, then we'll know where it *doesn't* fail. That's why opticians still happily trace the rays of light bullets, knowing exactly the circumstances where their calculations will be good enough, knowing ahead of time how much error to expect when they neglect interference and diffraction.

That's useful. That works. And *that's* science.


TODAY'S
SERMON
FROM THE
BOOK OF
NUMBERS

Chapter 2: Electricity and Magnetism

Part 1: Field Equations, or Loving the Math More than Nature

You may want to skip most of this chapter. It has math in it.

If you refuse to even look at math, then by all means skip to the end of the chapter, and start reading again at Part 4, the section with the title, “It’s OK To Read From Here.”

But first let me at least talk a little bit about math. The role of math in science has been very controversial for a very long time, and with good reason. Among the earliest of the Greek philosophers, 2500 years ago, were a bunch of mathematicians who followed Pythagoras (yes, the one with that theorem about triangles). They found math so simple, and pure, and delightful, that they imagined all of nature must be based on mathematical principles. So far, so good.

But this splendid idea immediately got corrupted into the notion that you could predict things about nature just from looking at magic numbers. For instance, followers of Pythagoras argued that the distance from the Earth to the Moon must be some simple number, like twice the circumference of the Earth, because any other value would be… well… inelegant.

Aristotle, a Greek philosopher who actually observed nature instead of speculating about how it ought to be, replied that this was nonsense and superstition — numerology, we call it nowadays. Saying the universe is based on magic numbers makes as much sense as saying that your future is determined by the number of letters in your name, or that the value of your bank account must equal your car’s license plate.

Aristotle saw that nature was just a little more complicated than that. In fact, he insisted that a good scientist is one who is willing to accept nature exactly as it is, not as some gorgeous and elegant theory tells us it ought to be. Aristotle banished numerology from science.

Unfortunately, his followers also corrupted *his* insight. They carried that principle too far in the other direction, and insisted on banishing all arithmetic and math from natural sciences. This attitude became especially prevalent after Aristotle was reintroduced to the Western world in the middle ages. By the late 16th century, the proper role of a scientist had been reduced to describing nature in words, not in equations.

Galileo upset that idea. That's what made him an object of scorn to the philosophers of his day. By insisting that he could describe motion in terms of measurements and numbers, it sounded to them like he was going back to the bad old days of Pythagoras and numerology.

And Galileo's opponents may have had a point. Kepler (a contemporary and a fan of Galileo) insisted, for instance, that the universe could only have seven planets because mathematically there were exactly six, and only six, regular solid polyhedra known to geometry, which he assumed fit into the spaces between the seven planets. Six was a magic number, to him. Shades of Pythagoras!

Even worse, by the 19th century (the heyday of classical physics) many scientists began to confuse the things that equations were trying to describe, with the equations themselves.

This tendency reached its extreme in the early 20th century, when a group of mathematical physicists and philosophers (most notoriously Ernst Mach and Bertrand Russell) developed a mathematical philosophy of science called Logical Positivism that asserted that the rules of mathematics and logic were the only basis of all truth.

Fortunately, one of the first things you can do with logical positivism is to use it to prove all the shortcomings of logical positivism.[7]

It's like worshipping a religious statue, or even the Bible, instead of worshipping God. Statues, Bibles, beautiful prayers, hymns, elegant ser-

---

[7] To give just one example, there's the problem of inductive reasoning. Every mathematician knows that simply because, in the past, event B has always followed event A, that's no *proof* that B must necessarily always follow A in the future. But that's exactly what you must assume when you rely on experimental science. When you flip the switch, the light always comes on; though that's not good enough for the logician, it's good enough for the scientist.

mons, gorgeous liturgies… all these things can be wonderful, so long as they're used to direct the congregation towards Him Whom we would worship. When they become ends in themselves — watch out.[8]

Science can have the same problem with math.

Nonetheless, given that warning, math really can help us understand nature and it's a fool who would not take advantage of its help.

To show what I mean, I'm going to tell a story of four scientists and their equations. The story is not, strictly speaking, exactly true in all its historical details; their history, like most truth, is far too complicated for me to do justice to in a few short paragraphs. But these people did exist, and they did come up with the laws that bear their names. So just take the following as a sort of mini docu-drama.

Furthermore, I warn you, it's going to feel like a shaggy dog story, with digression piled on digression. Be patient. It all ties together at the end.

We'll start first with a Frenchman, Charles Augustin de Coulomb. Like his contemporary, Benjamin Franklin, he was fascinated by static electricity. By his time (the late 1700s), it was perfectly fashionable to try to describe nature with numbers and equations; so he gave himself the task of finding out the mathematical formula that would describe the strength of the force that one got from a static electric charge.

Wait for a dry day (indoors in the wintertime works great) and rub an inflated balloon against a carpet — or your hair — to charge it up. Then take a small piece of newspaper, and rip it into tiny bits. Carry the paper scraps in the palm of your hand over towards the charged balloon. Far away from the balloon, nothing happens; as you get closer, scraps may stand up on their edges; then, when you get really close, they'll fly out of your hand onto the balloon. The closer you get, the stronger the force; the farther away, the weaker the force will be.

---

[8] It's a temptation in all philosophies. The Jesuit theologian Frans Josef van Beeck has written a book about it, based on an old Jewish folk tale: *Loving the Torah More than God*.

Well, after lots of careful measurements, taking those heroic efforts that all experimentalists make to take clean data, Coulomb convinced himself (and everyone who's tried it since) that the force of an electric charge drops like $\mathbf{1/r^2}$, where $\mathbf{r}$ is the distance between the charge and the object it's attracting. Go twice as far away, the force drops by 4; three times as far away, the force drops to 1/9 of its original value. We call this sort of force an "inverse square" force.

Coulomb and Ben Franklin and a host of other like-minded experimenters also found that the more charge you had, either on the balloon or on the scraps, the stronger the force would be. And, finally, they found that in fact there were two kinds of charges, which they called *positive* and *negative*. Positives and negatives attracted each other, while positives repelled positives and negatives repelled negatives.

In fact, here's the way that a modern physicist would write the equation to describe Coulomb's electric force law: $\mathbf{F = kQq/r^2}$. Which gives us the chance to spend a minute talking about equations.

Equations are sentences; sentences in a foreign language. As languages go, it's actually pretty simple; but it certainly is foreign. Let's translate this one, slowly.

Like any simple sentence, this one has a subject, a verb, and an object. The subject is the letter **F**; it stands for Force. The verb is the equal sign. So **F =** can be translated as, "The value of the Force (the electric force between two charges, we understanding this to mean) is found by doing the following operations…"

The object, the stuff on the right hand side of the equation, is where all the action is. If I want to calculate the force, it tells me that I need to know the distance between the two charges, **r**; the amount of the two charges, **Q** and **q**; and a constant number **k**, which we'll get to in Part 2 of this chapter. And it tells me what to do with these quantities. If I know **Q** and **q**, I multiply them together. If I know **r**, I square it and divide the result into the product of **q** and **Q**. And then I take the whole mess, and multiply it by **k**. And that gives me the strength of the electric force.

One of the nice things about this formula, incidentally, is that it doesn't matter what order I do all this multiplying and dividing; I'll wind up with the same answer at the end. That's not true of all formulae, but it's true in this case.

So, say I'm sitting in a big church with a lump of charge, **q**, in my lap. Up on the pulpit, where the preacher stands, is another lump of charge, **Q**. I can use this formula to calculate how much force there is between my charge and the charge on the pulpit: $\mathbf{F_1 = kQq/r^2}$. If I get up and leave my seat, you might come by with a different lump of charge, **z**. When you sit in my old seat, you'll see a different force: $\mathbf{F_2 = kQz/r^2}$.

Now, because it doesn't matter what order I calculate this force, I can rearrange the terms so that all the parts that don't change are lumped together: $\mathbf{F_1 = q[kQ/r^2]}$ in my case, $\mathbf{F_2 = z[kQ/r^2]}$ in your case. In fact, if I look under the pew seat, I'll see taped there a little piece of paper with a number, equal in value to $\mathbf{[kQ/r^2]}$, and the instructions: "To find the Force, multiply the amount of charge you have in your lap by this number."

The number on that piece of paper depends on the amount of charge **Q** sitting up in the pulpit, and on the location of the pew where the paper's taped. Every seat has a different number, just as every pew seat is a better or worse location for hearing the preacher.

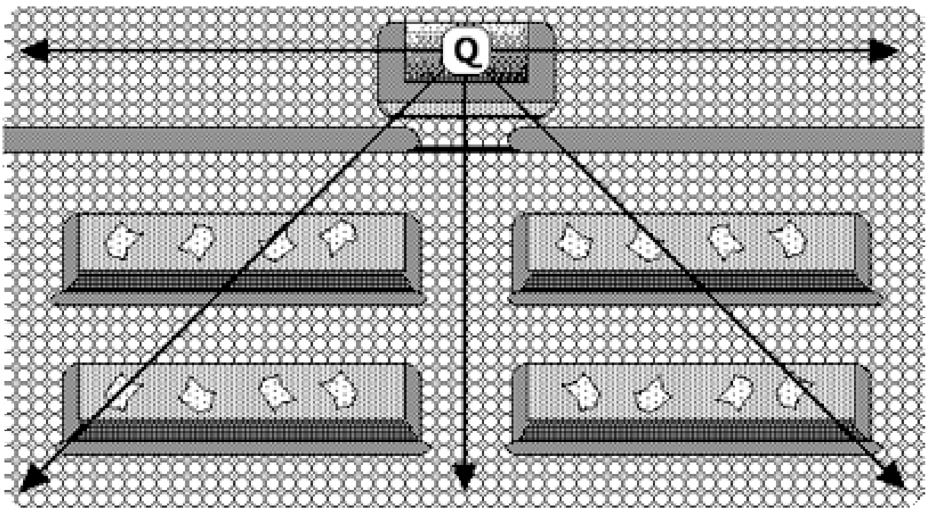
That number stays the same, no matter who's sitting there. Just as we all hear the same sermon, we all experience the same charge **Q**. But the force we actually feel from that charge **Q** depends on how much charge **q** we each bring to our seats, just as the effect of the sermon varies from person to person, depending on what experiences (and how much attention) we bring with us to church.

Every seat has its own number; and these values all taken as a whole, this set of numbers everywhere, is called a *field*; in this case, an *electric field*. Just as every seat in my church has its own number pasted to the pew, or every point in a wheat field has a unique blade of wheat, every point in space has its own number describing the electric field at that point.

Actually, my electric field has something else besides a number. If I look at that piece of paper on my pew seat, I notice it also has an arrow drawn in. It's pointing directly away from the pulpit. This particular field tells me both the strength and the direction of the force that I will feel. (A field with both magnitude and direction is called a *vector field*.)

Obviously, the orientation of this arrow will change from seat to seat as I move around the church, changing my location relative to the pulpit. In fact, I can do something new to show these directions. Taking a piece of colored chalk out of my pocket, I approach the pulpit (much to the surprise of the preacher!) and then draw on the floor straight lines, radiating away from the pulpit, crossing over aisles and pews. Every time I encounter one of those little pieces of paper, I notice that the arrow is pointing along the direction of the chalk line I have drawn.

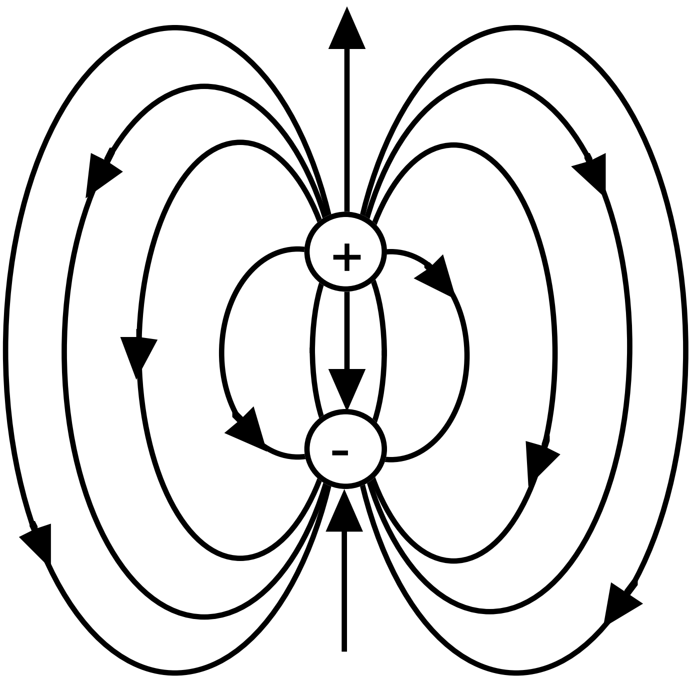
Eventually, the floor of the church looks like:



These lines are called "field lines." They're a pretty picture to help me visualize the field. The convention is that positive electric charge is drawn with field lines pointing out; negative charge has field lines pointing in. If I have a positive charge in my lap, it will be pushed in the direction of the arrows; a negative charge goes opposite to the direction of the arrows. (Assuming we all have positive charges in church today, including the preacher, I see from this picture that we're all being repelled away from the sermon. He's probably complaining about the collection again.)

For an inverse-square field, it can be shown[9] that the strength of the field is exactly proportional to how tightly packed the lines are. Thus, for my point charge, I can see from this picture that the field feels strongest in the seat nearest the pulpit. No surprise.

The place where field lines really get useful, however, is where I have more than one charge interacting. What if I have a positive and a negative charge a fixed distance apart from each other? Detailed calculations confirm what my simple picture intuits: we have a very complicated looking field, called a dipole field:



[9] A German mathematician named Karl Friedrich Gauss did this, around 1800.

Regardless of what kind of charge I bring, I'll be attracted to one end of the dipole, and repelled from the other. Using this picture I can even guess how I will want to move, following along the field lines.

One final thing you should realize about the electric field. It's just a mathematical gimmick. We've invented it. The real thing, the thing in nature that we actually experience and can measure, is the force. The "field" is just a shorthand way of letting us visualize and calculate that force. But obviously, a field is not something that actually exists, the way that time and space, or matter or forces, exist…

Is it?

Part 2: A New Force to Reckon With

Recall our equation for the Electric Force: $\mathbf{F = kQq/r^2}$. Say I know a value for the charges **Q** and **q** at two different points, and I know **r**, how far apart these charges are. I plug the numbers into the equation, and do the multiplying and the dividing like the equation tells me to do. What does this get me?

It gets me a number. This number represents the strength of the force. But what does this number mean? How can I compare it to other forces, like the weight of a pound of butter, or the push of a Space Shuttle rocket?

Now, there are lots of scales I could use to describe my force. For many years, English engineers measured the strength of a force by calculating how many pounds of stuff it could lift against gravity. The modern way is to calculate the speed (in meters per second) to which the force can accelerate one kilogram of stuff, starting from rest, in one second.

This definition comes out of Isaac Newton's laws of motion, and so a unit of force strong enough to accelerate one kilogram in one second to a speed of one meter per second is said to have the value of one Newton.

(It takes a little over four Newtons to a lift a pound of butter — or a pound of anything else, for that matter.[10])

Well, if I want **F** in my equation to be in Newtons, I have to be careful what values I put into the object part of my equation. Clearly, measuring the distance **r** in inches will give me a different number than measuring the distance in miles.

That's where the constant **k** comes in. The value of **k** depends on what units I decide to use to measure **r**, **q**, and **Q** in. It's the conversion factor that changes the number I get from $\mathbf{Qq/r^2}$ into Newtons of force.

For instance, say I stick strictly to the metric system and measure the distance **r** in meters. How do I measure charge? **Q** and **q** could stand for the number of electrons on each of my objects; that could be billions and billions of electrons, so **k** would have to be a very small number to counter the very large values I'll have from multiplying **Q** and **q**.

Another way is to use Coulomb's equation itself to define a unit lump of charge: by saying that **k** is exactly 1, then my basic unit of charge will be whatever charge I need on both **Q** and **q** to give me exactly one Newton of force at a distance of one meter. In fact, such a system is used in many advanced textbooks, and such a lump of charge is named an "electrostatic unit."

However, it turns out that the most commonly used unit of charge is defined in a very different way.

André Marie Ampère, another Frenchman, was one of the first to work with flowing electricity: charge in motion. One of the things he noticed was that if you move charge through parallel wires, the wires feel a force pulling them together (or pushing them apart, depending on which way the currents flow).

If you have two very long wires running parallel to each other, one meter apart, with equal amounts of current in each, an amount of current

---

[10] According to the internationally recognized standards for scientific nomenclature, the names of units taken from the names of persons, like the Newton and the Ampere, have no plural. But everybody I know happily refers to "Newtons" and "Amperes" anyway.

is defined to be one Ampere when you get a tiny force of 2×10^{-7} Newtons (0.0000002 Newtons) acting per meter of wire; one Amp, we usually say.[11]

And one Amp of current is equal to a certain amount of charge passing through the wire every second; we call such a lump of charge, one *Coulomb*. So this defines both our units of current, the Amp, and our units of charge, the Coulomb.

As it turns out, one Amp is a mighty strong electrical current, and so one Coulomb is a pretty large lump of charge… more than six billion billion electrons (about 6.25×10^{18} — 6,250,000,000,000,000,000).

If we measure **Q** and **q** in Coulomb's equation using Coulombs, the force between two such huge lumps of charge will be very big, and so we need the constant **k** to be a very large number: about 9×10^9 in fact.

In other words, two lumps of charge of one Coulomb each, situated a meter apart from each other, would attract (or repel) each other a force of nine billion Newtons. That's enough to lift a million tons. No wonder electricity can do so much work for us.

Now wait a minute, you may be asking. Why should two wires with current running through them have a force between them? My house is full of wires; why haven't I noticed this before?

You never noticed it, because (in contrast to the electrostatic force of two Coulombs) the force between two Amperes is tiny… 2×10^{-7} Newtons is about the same force as the weight of a speck of dust less than a millimeter in size. But the force is there. If you pump up the current, and bring your wires close together, you can actually see them jump when you turn on the juice.

So what is this mysterious force between the wires? Whatever it is, it's pretty weak. In order to make it a little stronger we could push ever

---

[11] Why 2×10^{-7}? Don't ask. The way units are defined turns out to be much messier than it would seem possible, at first blush. And in fact, nowadays the value of one Ampere has been arbitrarily defined, along with the value of one meter, one kilogram, and one second. We no longer depend on measuring ridiculously tiny forces on wires to define our electrical units.

more current through the wire, but pretty soon the wires would start to smoke, and then start to melt. Not a good idea.

A more clever trick is to use the same wire over and over again. Instead of having one long wire, why not bend the wire into a loop? Each ring of the loop will produce the force of the original wire, but if we have hundreds or thousands of loops we'll increase the total force by hundreds or thousands of times. In fact, we can do this to both wires, and we wind up with two coils of flowing electricity.

Now we're getting somewhere. We find that the two ends of the coils have very different behaviors. The end where the electricity is entering the coil feels a force pulling it towards the end of the other coil where the electricity is coming out. But it's repelled by the opposite end of that coil. One end attracts; the other repels.

Sounds like the dipole we described in Part 1. What happens if we bring an electric dipole up to this coil?

Nothing. Whatever the force is, it's not the electric force. Are there any other dipoles we can think of? Does any other material behave in this way, so that you can bring two objects together and they repel each other in one orientation, but attract each other in the opposite orientation?

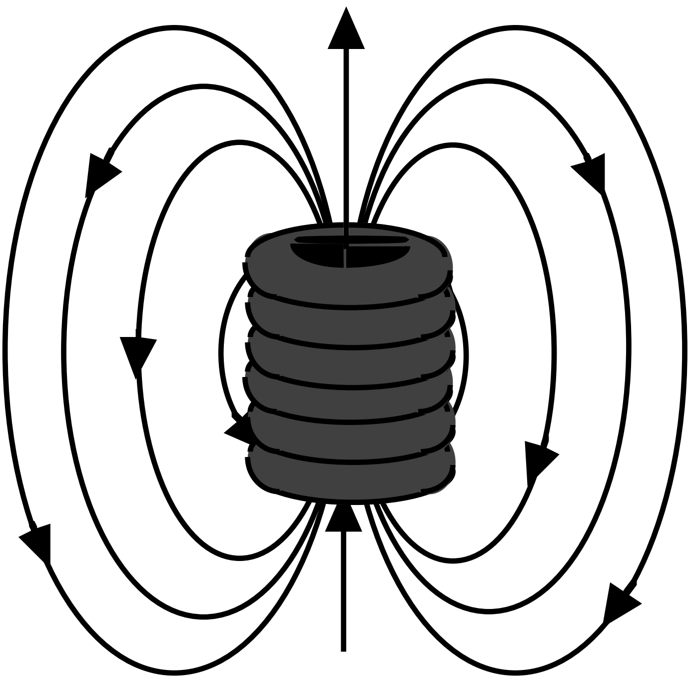
If you thought "bar magnets," you get the prize. In fact, a bar magnet behaves exactly like one of our coils. Bring a bar magnet up to one end of the coil, and one end of the magnet is attracted while the other is repelled. Repeat it at the other end, and the opposite behavior occurs. Our coils of wire with electricity flowing through them — what we call electromagnets[12] — have created a *magnetic field*.

---

[12] In fact, all magnetism, even in a natural bar magnet, is the result of some sort of electrical current flowing in loops… or more complicated directions. Usually, kids are told that an electromagnet acts just like a bar magnet. This is backwards. The more true statement is that bar magnets act just like coils of wire with current flowing through them. And if you can see the difference between these two statements, you have the makings of a first-class theologian. It's a question of which one is fundamental, and which is the imitation.

Field. There's that word again. What does the field of our coil look like? We suggested it already, and it turns out careful measurement confirms that the field lines of a coil of wire look just like a dipole's.

Because we can use a magnet as a compass on Earth, with one end seeking the north and the opposite end pointing south, we can call the two magnetic poles a "north pole" and "south pole."



We know that an electric dipole is made up of two separate charges, positive and negative; does this mean that a magnetic dipole made up of two magnetic "charges"? Maybe. Only trouble is… no one has ever been sure they've found a single magnetic pole, a "monopole."

The search for magnetic monopoles has been going on pretty intensively for several hundred years now; but the fact is, in nature it seems you never get just a north pole or just a south pole sitting by itself. It would be like having a coil with only one end; common sense tells us that every object with an end also has a beginning. Still, it's nice to know if we ever do find a magnetic monopole, we have the names ready.

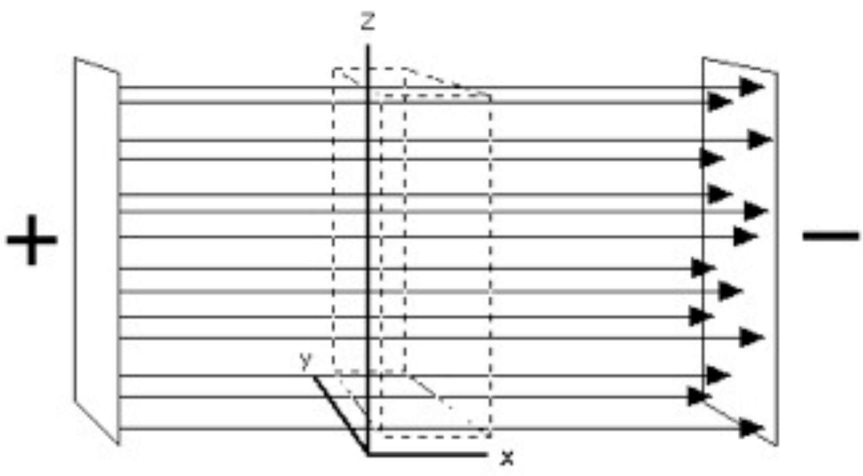
We also have the equations ready. All the equations that describe electrical charges, plus and minus, in all the possible orientations it comes in, should also apply to magnetic "charges" as well.

After all, all those equations are based on Coulomb's law; the dipole, after all, comes from applying Coulomb's law twice, to two poles. If monopoles exist, they should obey their own version of Coulomb's law and so they should obey all the other rules we've developed over the years about how electric forces, and electricity, behaves.

That, it turns out, is very important. Why? Well, let's look at some of those electric laws.

The electric field, remember, was a field that could be represented by field lines. That means that we can draw pictures, and infer things from the geometry of those pictures, without resorting to a lot of messy equations. (Hah! The truth finally comes out — we physicists don't like equations any more than regular people do; especially since we have to live with them. Anything we can do, any trick that works to simplify or eliminate equations, we'll do it.)

So let's imagine the simplest possible electric field… where all the field lines are straight, and parallel, running side by side in what we'll call the "**x**" direction.



Is such a field possible? Sure. Imagine two enormous flat panes of glass, side by side. One has positive charge painted on it, the other negative charge. The field lines will run straight from one to the other. No reason for them to bend around at all (except near the edges, but we'll stay away from edge effects).

Now picture an imaginary flat box sitting in this field, as illustrated. Obviously, the number of lines going into the box from the left equals the number of lines going out of the box to the right, traveling through the box a distance we call $\mathbf{\Delta x}$. (This would change if you hid some extra charge inside the box, but we aren't going to do that.) Incidentally, the Greek letter Delta, $\mathbf{\Delta}$, is the symbol that means "the change," specifically the change in whatever's represented by the symbol sitting next to the Delta.

So we take this sentence from English, "The change in the electric field (which is running in the x direction) over a distance in the x direction, is zero," and translate it into math: $\mathbf{\Delta E_x / \Delta x = 0}$. Now, this statement is obviously true for the simple geometry we've cooked up here. You'll have to take my word for it (or better, Karl Friedrich Gauss's word; it's called Gauss's Law) that this result is merely a simplification of a more general result that's true for *all* electric fields.[13]

What about magnetic fields? Well, if they can also be described by field lines, then this rule must also be true for them. And, trust me, Gauss can show that it's true for *all* magnetic fields, too. We don't even have to worry about the case of charge hiding inside the box, since there are no magnetic "charges" as far as we know.

Because it'll be useful later on, let's have our magnetic field lines running in the **y** direction, so we can say, "The change in the magnetic

[13] Halliday and Resnick, in their freshman College Physics textbook, take twenty pages to explain Gauss's Law. Even then it usually takes a year or more of working with this stuff, long after the course is over, before the typical undergraduate finally understands it… or at least gets used to it.

field (which is running in the **y** direction) over a distance in the y direction is zero," which in math looks like: $\mathbf{\Delta B_y / \Delta y = 0}$.

What good does all this do us? Well, besides introducing us to Mr. Coulomb and Mr. Ampère (and Mr. Gauss, as well) it lays the foundation for the wham-o finish to our story, in Part 3. There, we'll meet our last two characters, Michael Faraday and James Clerk Maxwell.

Mentioning Michael Faraday, I'm reminded of one final anecdote. It seems he was demonstrating some of his early experiments in electricity to assembled members of the British Parliament when Disraeli, ever the practical politician, stopped him in mid-sentence.

"Electricity?" fumed the Prime Minister. "What possible good can it do us?"

Faraday didn't miss a beat. "Why, Mr. Prime Minister," he's reported to have said. "Some day, you'll be able to tax it."

Part 3: "And Maxwell said…"

When I taught physics to college students, I'd always start the semester on electricity and magnetism by wearing a tee-shirt I'd picked up during my own student days at MIT. It was put out by MIT-Hillel, the Jewish Student Association at MIT. It read:

And God said…

$$\nabla \bullet E = 4\pi\varrho$$

$$\nabla \times E = -(1/\mathbf{c})\, \partial B/\partial \mathbf{t}$$

$$\nabla \times B = (1/\mathbf{c})\, \partial E/\partial \mathbf{t} + 4\pi \mathbf{J/c}$$

$$\nabla \bullet B = 0$$

…and there was light!

The fun, of course, was that to my beginning students these equations looked like just so much Greek; but by the end of the term, they'd know exactly what each squiggle meant. (Some of them really are Greek.) And then they'd get the joke.

They're called Maxwell's Equations, and we'll meet Mr. Maxwell soon. I won't make you learn all the squiggles. But, at the end of the chapter, I'll share the joke.

Actually, you already know more than half of what is in these equations.

The top one is the all-purpose generalized calculus version (and in electrostatic units, not Coulombs) of our old friend $\mathbf{\Delta E_x / \Delta x = 0}$, while the last one is likewise the generalized version of $\mathbf{\Delta B_y / \Delta y = 0}$. I hope you can see now why I'm sparing you the general derivation, and looking only at my simplified cases.[14]

And in fact, you've already seen the science behind part of the third equation, too. In that equation, **J** stands for a kind of electric current, and in some complicated way this equation is telling us that a current **J** on one side of the equation can produce a magnetic field **B**, as seen on the other side. That was Ampère's Law. However, in our super-simple system of constant parallel fields, whatever currents are generating my magnetic field $\mathbf{B_y}$ are all far off-stage and so **J** should be set to zero here.

Enter Mr. Faraday. As we mentioned in the last section, he was fond of playing (I should say, experimenting) with electricity and magnetism. He knew about Ampère's Law; and if an electric current could produce a magnetic field, he realized there must be some connection between the two fields.

One startling discovery he made was that, not only could an electric current produce a magnetic field; but a magnetic field could likewise pro-

---

[14] Notice that **E** stands for Electric Field, and **B** stands for Magnetic Field. Why **B**? Why not **M**, say? Well, that's puzzled me for years. Finally, I looked up James Clerk Maxwell's original paper in the *Transactions of the Royal Philosophical Society* for 1865 (there's a complete run of this journal, going back to the 1600's, at the Vatican Observatory). It turns out that Maxwell talked about a number of electric effects and forces, which he labeled simply "**A**" through "**F**". For some reason, using the letter "**B**" (by which he meant what he called the Magnetic Force) has stuck.

duce an electric current! The situation was a natural one to come up in the lab. Faraday had his coil of wire, with a delicate meter attached to see just how much current was flowing through it.

He then brought over a strong bar magnet from the other side of the lab. As he approached, the ammeter would suddenly register a current in the loop — even though no batteries had been connected to the wires.

As soon as he stopped moving, the current stopped.

The faster he ran over towards the wires, the bigger the current. If he jerked the magnet away from the coil, a burst of current would flow in the opposite direction. Pretty soon, by waving the bar magnet back and forth, he was able to get a current flowing back and forth through the wires.

He had just discovered the principle behind electric generators, and in the process invented alternating current.

Now, this isn't power for nothing… it takes energy to swing those heavy magnets back and forth, and they have to be kept moving all the time or else the current stops. But modern generating stations, using turbine engines powered by coal or gas or nuclear heat, or simply using the weight of a huge lake of water behind a dam, have big magnets spinning around huge coils of wire, and they produce the high voltage alternating current that powers our toasters and television sets. A bicycle lamp that gets its power from the spin of the bike wheel works the same way.

Faraday, as it turns out, was not a big fan of mathematics. In fact, he managed to write three volumes about electricity and magnetism without using a single equation. His results, in words, can be summarized as: a changing magnetic field produces a current in a coil; and the amount of current you get depends on how big the coil is — the bigger the area enclosed by the coil, the more current you get — and on how fast you change the strength of the magnetic field.

Faraday may not have liked math, but we do. (Don't we? I do; and if you've stayed with me this far, you must, too.) So let's express this rule in a mathematical form.

Start with our magnetic field, $\mathbf{B_y}$. Remember that the little y means the field is directed in the **y** direction. If we have a loop of wire perpendicular to this field, it'll run in the **x** and **z** directions. Make it a little square loop; then its area is $\mathbf{\Delta x}$ times $\mathbf{\Delta z}$.

The change of the magnetic field with time we can write as $\mathbf{\Delta B_y / \Delta t}$ so we wind up with the change in the strength of the magnetic field, times the area of the coil loop, being expressed as: $\mathbf{\Delta x\ \Delta z\ \Delta B_y / \Delta t}$. This produces a little bit of current.

Now, a current is a bunch of charge in motion. What makes charge move? An electric field. We know that; we even know that the direction the charges flow tells us the direction that the electric field is pointing. The amount of extra current you get depends on how much you've changed the field, and how far you're trying to push the charge.

And obviously, if the current is increasing, the (positive) charge must be moving in the direction that the field is increasing. So if we're producing a new current, we must be upping the strength of the electric field over the length of the coil: $\mathbf{\Delta E_x\ \Delta x + \Delta E_z\ \Delta z}$. Thus our equation gives us:

$$\mathbf{\Delta E_x\ \Delta x + \Delta E_z\ \Delta z = \Delta x\ \Delta z\ \Delta B_y / \Delta t.}$$

Now we fast forward to the middle of the 19th century, and the last of our characters: James Clerk Maxwell. The story is about to reach its climax.

Maxwell, a math whiz, assembled all the equations he knew about electricity and magnetism and tried to imagine in his mind what they were really saying. You see, rather than just using equations as simple calculation rules, he believed them. He took them literally. He said, in effect, "If the equations say the electric field is produced when you change the magnetic field, then by golly that's what happens…" [15]

[15] For the purists in the audience… I'm telling a story here, not reciting history. In fact, the attitude I am attributing to Maxwell is more properly the attitude of twentieth-century field theorists. What Maxwell actually did was assume that electricity and magnetism transmit their effects through some sort of hypothetical elastic medium called the ether; by assuming the reality of this ether, which we now know does not exist, he derived his famous equations.

Now remember our warning from back in part one. The electric field is a calculating fiction, not something real. The result of Faraday's law only makes sense when there's a real wire with real charges in it really feeling a real force. At least, that's our commonsense gut instinct.

But that's not how Maxwell thought. To him, the fields themselves were real. "Even when there's no coil of wire around to actually have a current in, there should still be an electric field produced when the magnetic field changes. If there *were* a coil, it *would* have a current; and that's good enough for me!"

It's a matter of trusting your instincts. Look at the prophets of ancient Israel. There were many devout Jews in those days; but only a prophet, someone slightly crazy — or maybe, a little bit saner than the rest of the world — would take all the words and rituals and ideas that their culture had repeated so often they'd lost their meaning, and suddenly believe them for exactly what they said.

For example, many pagan religions so feared the gods that their priests would offer human sacrifices (chosen from somebody else's family, of course); but only an Abraham would have been willing to sacrifice his own son — and trusting enough to hear and obey God when God said "No!"[16] Only a Moses would be crazy enough to think, if God is God then I can depend on Him to back me up when I lead my people from slavery.

And only a Maxwell could think, if the equations talk about an Electric Field, then maybe I should assume that there actually is a real electric field there.[17]

So let's follow Maxwell, and apply his reasoning to our simple field. We had an electric field in the $\mathbf{E_x}$ direction, and a magnetic field in the $\mathbf{B_y}$ direction. So, for simplicity, let's concentrate only on what's happening to $\mathbf{E_x}$, so we can drop the $\mathbf{E_z}$ term and simplify everything to:

---

[16] A liberal is someone who wouldn't have believed in God enough to think of making such a sacrifice. A conservative is one who wouldn't have believed in God enough to stop.

[17] Is this to suggest that scientists have their own private "hot line" to God? Well… yes. But so do we all, each according to our gifts, as St. Paul says.

$\mathbf{\Delta E_x \, \Delta x = \Delta x \, \Delta z \, \Delta B_y / \Delta t.}$

Now divide away the $\mathbf{\Delta x}$ and we are left with:

$\mathbf{\Delta E_x = \Delta z \, \Delta B_y / \Delta t,}$ or $\mathbf{\Delta E_x / \Delta z = \Delta B_y / \Delta t.}$

This says, in English, that if I start changing the strength of the magnetic field in the **y** direction, I'm going to start seeing a change in the electric field (in the **x** direction) as I travel in the **z** direction. Uh, right.

Assuming you're a Maxwell, you can keep all that in your head and carry it even farther. Because then, Maxwell did something even stranger.

He knew that the electric fields and the magnetic fields seem to behave in very similar ways, and seemed to follow similar equations. He had just worked out this new simple equation for producing a change in the **E** field by varying **B** with time.

Would it be possible that a change in the **E** field with time would also produce a **B** field? Could there be some equation that would look like $\mathbf{\Delta B_y / \Delta z = \Delta E_x / \Delta t}$?

Almost.

You get a **B** field when you've got current passing by; a current is a lump of Coulombs per second; and a lump of Coulombs produces an **E** field. So an **E** field that changes with time implies a lump of Coulombs changing with time, which implies a current, which ought to produce a **B** field.

But remember our problem with units. To change from an **E** field, to a lump of Coulombs, we have to divide away our factor of 9×10^9. And then, remember that to find the force of the magnetic field from two parallel wires, we had to reduce the number by a factor of 2×10^{-7}. Here, we have in effect only one wire, so we reduce it to 1×10^{-7}.

Thus our equation looks more like:

$\mathbf{\Delta B_y / \Delta z = (10^{-7} / 9 \times 10^9) \, \Delta E_x / \Delta t.}$

But even this has a problem.

If I increase the **B** field, I increase my **E** field. If I change my **E** field, I change my **B** field. Obviously this starts to feed back on itself. Do I wind up with ever-increasing **E** and **B** fields? Common sense tells us that's impossible. And common sense is right.

Instead, what happens must be that increasing the **E** field must cause the **B** field to get smaller, which then drops the **E** field, so the **B** field gets bigger… and these two fields can bounce around between reasonable limits, without ever becoming infinitely big. So to show this, I must add a negative sign to my equation:

$$\Delta B_y / \Delta z = - (1 / 9 \times 10^{16})\ \Delta E_x / \Delta t$$

Notice that I've also simplified things by doing the multiplication inside the parentheses. The number I wind up with, 9×10^{16}, may look pretty big. It is. But more than big, it's important. Do you recognize it? I thought not.

But Maxwell did…

And so he did one more very strange thing. He asked, in effect, what happens if I look at the change of both sides of this equation, as time changes? I get

$$\Delta(\Delta B_y / \Delta z) / \Delta t = - (1/ 9 \times 10^{16})\ \Delta(\Delta E_x / \Delta t) / \Delta t$$

Then I can shuffle the pieces around until it looks like

$$\Delta(\Delta B_y / \Delta t) / \Delta z = - (1/ 9 \times 10^{16})\ (\Delta^2 E_x / \Delta t^2)$$

But I know what the part in the parentheses on the left is; I can substitute, and get

$$\Delta(\Delta E_x / \Delta z) / \Delta z = - (1/ 9 \times 10^{16})\ \Delta^2 E_x / \Delta t^2$$

or:

$$\Delta^2 E_x / \Delta z^2 = - (1/ 9 \times 10^{16})\ \Delta^2 E_x / \Delta t^2$$

And when Maxwell saw this equation, he must have flipped his wig. It was incredible! Astonishing! Words pale at the immensity of this breakthrough. It was the most frighteningly powerful new thing to happen in physics in 300 years.

All right, it may just look like gibberish to most of us. But not to Maxwell… because he was able to read the language of mathematics. And he had seen this equation before. He knew what it meant. It meant…

Part 4: It's OK to Read From Here

If you've skipped over the math part, you skipped all the tedious derivations that finally led James Clerk Maxwell to an amazing mathematical discovery.

People before him had talked about electric and magnetic fields, though in fact these "fields" were mostly just calculating devices to make it easier to compute the strength of electric and magnetic forces. But Maxwell treated these fields as if they were real things.

He assembled all the laws of electricity and magnetism known to that time, expressed them in mathematical form, and recognized that there must be a missing piece to the equation… just from the "symmetry" of the equations he had before him. Then, by manipulating these equations, he came up with a stunning new equation.

It said that if the electric field changed from place to place, the change of the change in space matched the change of the change in time… divided by a big number, about 90,000,000,000,000,000 (nine followed by sixteen zeros), in metric units.

He knew this equation. It's the equation for a wave.

Think of it. If you stand in one spot in a lake, you'll watch the waves go up and down as time passes. Or, if you freeze time (and the lake) then move along from place to place in space, you'll see the water going up and down *in just the same way.*

The fact that the wave goes up and down is a change; the curve to the wave is the change in the change, the way the change changes; and the shape of the wave in space is the same as the shape of the wave in time.

That equation about the change of the change in space matching the change of the change in time, is the mathematical formula that describes the shape of a wave. And it doesn't matter what's waving. This equation said that the electric field should have the shape of a wave. The electric field ought to behave like a wave. It ought to *be* a wave.

What's more, this equation gives the speed of the wave. That big number we had, 90,000,000,000,000,000, sits in the part of the wave equation that tells you the speed of the wave; only in the wave equation, it turns out that it's actually the speed, squared. So what's the square root of 90,000,000,000,000,000? It's 300,000,000 meters per second. That's darn fast. In fact, it's the speed of… The speed of light.

Now, what acts like a wave, and travels at the speed of light?

How about, a light wave?

Maxwell realized that the same thing he did for the electric field he could do for the magnetic field. Both of them would wave together. The electromagnetic wave, he called it. And that's light.

Out of his purely mathematical equations, describing a purely invented thing called a field, Maxwell had derived the meaning of light. He'd found the answer to the question we raised in Chapter 1, Part 1, that we thought we'd never get an answer to. Now we know what light is! It's a wave in the electric and magnetic fields.

But… but… but… Those fields were just mathematical fictions, calculating shortcuts; they weren't actually real, were they? That laughter you hear in the background is Maxwell, saying over the centuries, "I told you so! I told you they were real! The equations don't lie!"

Not everyone was convinced. In the previous section, we looked at how the inspiration of a scientist was similar to the inspiration of a prophet. But scientists, like prophets, must be very cautious. Not every stirring of the soul has God as its author, and not every flash of "insight" turns out to be correct. Just as salvation history is littered with false prophets, so too the history of science has been peppered with neat ideas, from phlogisten to cold fusion, that sure sound great but just don't work.

For better or worse, we look to the authority of Organized Religion to ward off wacky theologies; for the same reason, Organized Science has checks and balances to put new ideas to the test. Just to get a paper published, a scientist must win the approval of a journal editor and any number of referees, fellow scientists who review and challenge each new idea under the protection of anonymity. And even after it's published, the paper is open to both theoretical and experimental challenge. And so it was with Maxwell's equations.

In Germany, an experimentalist named Heinrich Hertz scoffed, "If these equations were true, then I should be able to make such waves by building an oscillating electric dipole. It shouldn't take more than a few minutes to rig one up at one end of a table, and then stick a wire all by itself at the other end. I sincerely doubt I'll actually see electric currents flow back and forth in that other wire. See? It's… um… hmmm… well, isn't that interesting…"

It actually worked. And to this day, the frequency of oscillating electromagnetic waves is measured in units called "Hertz" in his honor.[18]

In Italy, Marconi figured out how to make powerful oscillating electric dipoles, with tall towers to broadcast the waves, radiating in every direction, each direction like the spoke of a wheel, a circle's radius… a *radio*, as they say in Italian. Pretty soon everyone was stringing up wires on tall towers[19] to catch Morse Code, and then actual human voices and music, and eventually pictures — broadcast in Living Color!

Radio waves are light waves; it's just that, compared to visible light waves, the oscillation frequency for radio is pretty low and so these waves are pretty long. At the other end, x-rays and gamma rays are also just light waves, but with extremely high frequencies — they pack a lot of energy. Between ordinary radio and light are the frequencies we call microwaves. When a technician building radars for the Raytheon Company during

---

[18] And, actually, Hertz understood better than Maxwell that his equations could be true even if the theory behind them wasn't. "Maxwell's theory is Maxwell's equations, that's all," he once said. And he was right.

[19] These tall towers had an impact on another field of human endeavor. About the same time as radio, airplanes were also being invented. People studying airplane wings suddenly realized that sailboat sails were really just airplane wings stood on end; and that, like airplane wings, a tall, thin shape was more efficient than the big square sails people had been using for thousands of years. In order to support a tall, thin sail you needed a tall mast, just like a radio tower. And so, to this day, this kind of sailboat rig is called a "Marconi" rig.

World War Two noticed that he could heat up his lunch by sticking it inside a radar transmitter, the microwave oven was born.

Electromagnetic waves confined to wires are what make telephones work; even the alternating current that is used to send electricity from power stations into our homes is a kind of very long wavelength light wave. And all the electronics that goes on inside a computer has its basis in Maxwell's equations. His fooling around with some mathematics, back in the 1860's, is what invented the 20th Century.

Maxwell had no idea what he was unleashing, of course. Just the fact that he knew what light was, was amazing enough for him. And if physicists at the end of the 19th century had a sort of overweening pride, well, maybe that was understandable. A seemingly intractable problem had been solved. Theory had tied together several fundamental parts of physics, and made a prediction. And it worked. Everything else looked like just tying up loose ends.

At this point, I'd like to stand back and pause for a second, to appreciate from a human point of view what it must have been like for Maxwell. How does it feel to make an incredible scientific discovery? What's the sort of emotion that goes through your mind?

The fact is, of course, that Maxwell had no concept of radio and computers and modern electronics, and what they'd all mean a hundred years later to the average Joe in the street. What he had done, however, was something far more sublime and far more satisfying than making microwave ovens possible.

I can only speak from my own experience; and as a scientist I'm not quite in Maxwell's league, any more than my spiritual life approaches that of, say, Theresa of Avila. But I've played a little minor-league science, and some sandlot spirituality; enough to feed my imagination of what the big leagues must be like. And it's from my own experience that I keep coming back to that parallel between religious and scientific inspiration.

The emotion of scientific discovery is very much like the emotion of prayer — that rare kind of prayer when you suddenly realize that God has actually been speaking, and you've just gotten the slightest glimpse of What It's All About. I assume that you have experienced such prayer. At least, I hope so. C. S. Lewis described it beautifully, in one of his books, as being "surprised by joy."

A scientific insight, like an answered prayer, is something you only notice after it's happened. Often it's a long time after.

Sometimes you need someone to point out to you that you've had an insight, because while it was happening it seemed like the most natural thing in the world and you can't imagine that no one else has ever realized before what you've just understood. Isaac Newton once happened to mention in passing to his friend Edmund Halley that, many years earlier, he'd worked out the law of gravity from the motions of the planets; it was only Halley's repeated pleas that got Newton to finally write his *Principia,* the foundation of modern physics. More recently, I know one astronomer who only realized the implication of one of his discoveries when it was written up (to his surprise) in the astronomy magazine *Sky and Telescope*. It took a journalist to see the context that the so-called expert had totally missed.[20]

You'd think that making a big discovery would make you feel smart; but generally it's just the opposite. Once you understand something, you feel so stupid that it took you so long to get to what now seems obvious. It's just like how anyone who's heard from God feels — not blown up in a feeling of importance, but rather blown away by an intense feeling of being humbled (not to say, downright scared). Every Old Testament prophet was reluctant;[21] even St. Paul, after getting knocked off his horse, spent

[20] Okay, so I'll confess — that idiot of an astronomer was me. Fact is, that was the first and last time I ever did any science interesting enough to get into *Sky and 'Scope*. July, 1975, issue, as I recall.

[21] So much so that someone finally wrote a hilarious satire about the ultimate reluctant prophet: the Book of *Jonah*. Jonah is the prophet who gets it all wrong and succeeds in spite of himself. But then, isn't that usually the case? Even Jesus didn't hesitate to compare Himself to Jonah.

the next ten years hiding out in Tarsus making tents before Barnabas finally dragged him off to Antioch to start spreading the Gospel.

Once you've had a visit from the Real Thing, all your self-importance just looks shabby and ridiculous. Likewise, a real scientist (as opposed to the media stars you see on TV) is often reluctant to talk in public about his or her work, because there's nothing like standing on the edge of the unknown to make you look insignificant and stupid.

But finally, as C. S. Lewis described it, there's a feeling of great joy. Sure, some of it is ridiculously self-centered — "I'll get tenure! I'll show up those guys up at State U.! And maybe I'll get my name in *Sky and Telescope!*" A deeper part of the experience is a sense of reassurance — God really is listening, and you're not a total failure as a scientist after all.

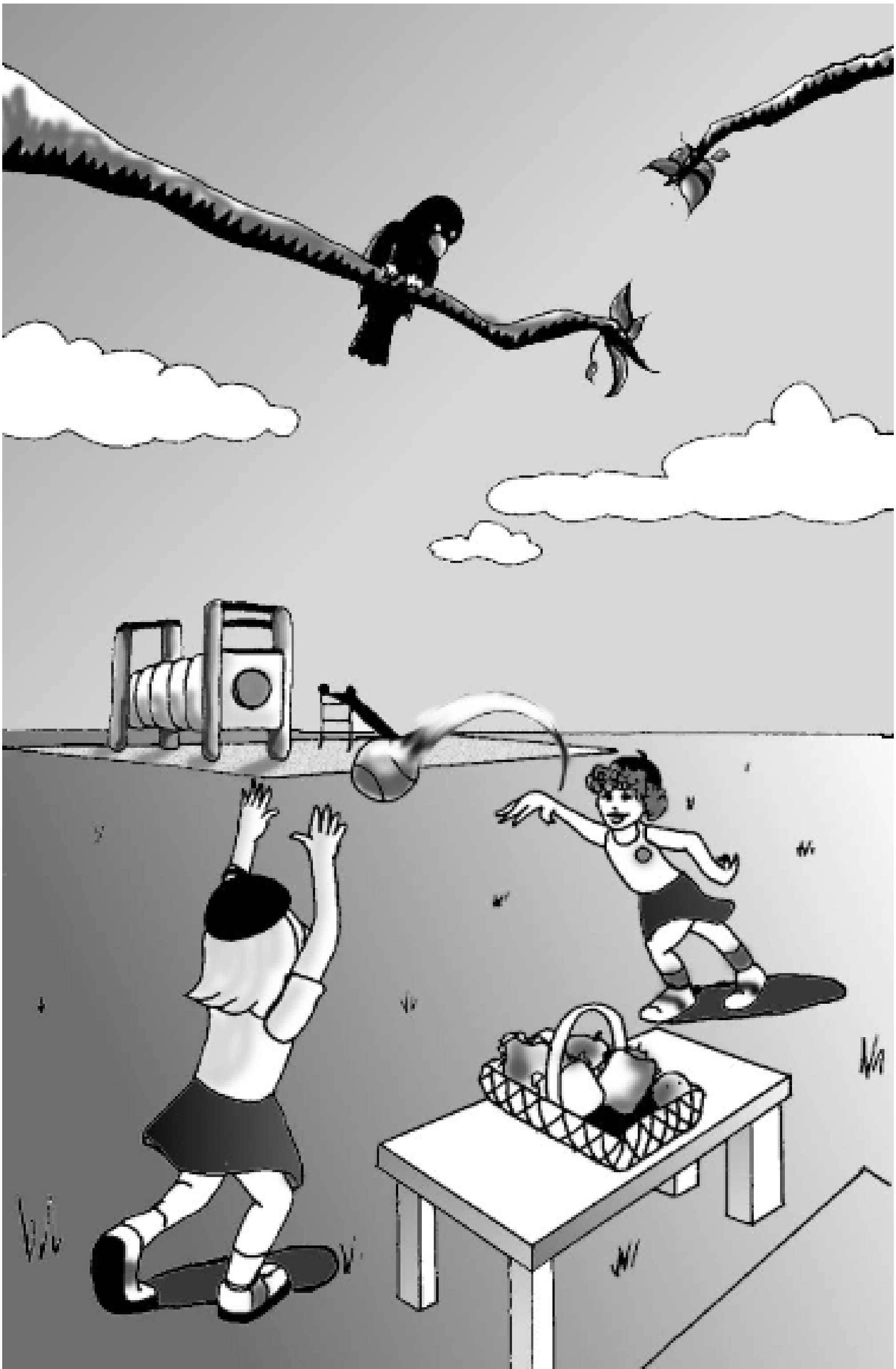
But underneath it all is a real sense of pure rightness that comes from having touched your own personal piece of the truth. It's not complete; it's probably not even completely true. And certainly it's not all yours alone; you couldn't have done it without a lot of other scientists before you. But it's yours; it happened to you; and no tax man or tenure committee can take it away.

I don't know for sure, of course; but I'll bet that's what Maxwell might have felt. His was a glorious achievement, an insight that tied things together in a way that no one before (or since) has ever done. It was, so far as light was concerned, The Answer.

So what's left for us to do? Where do we go from here?

In fact, Einstein's Theory of Relativity comes straight out of Maxwell's equations. That might seem like a logical place to go from here. We're not going to do that, however; because there's something more fundamental, more important, to talk about first. In fact, it's so important that it's this other work, not Relativity, that earned Einstein his Nobel Prize.

Because, you see, it turns out that Maxwell was wrong.



Chapter 3: The Photon

Part 1: Stuff and Nonsense

One question may already have occurred to you in our discussion of light. Why did we insist, back in Chapter One, that light traveled as individual bullets? How did we decide that it wasn't one continuous bar of solid "light stuff," whatever that might be?

In fact, this question is just one instance of a continual problem for classical (nineteenth-century) physicists. Is nature continuous — can we think of it as made up of smooth, slowly varying entities, which physicists called "fields"? Or was it made up of discrete lumps — "corpuscles" or "atoms"? Is "stuff" continuous, or atomic? Is energy continuous, or atomic? Is light continuous, or atomic?

Each possibility carries some serious implications for what reality is all about. If matter came in discrete lumps, then what happens if you somehow split the lumps into smaller lumps — how long can this splitting process go on? Even more disturbing, if energy and space and time itself come in discrete indivisible pieces, what marks their boundaries? And what is the meaning of the space between the lumps?

On the other hand, if matter and light and energy are smoothly continuous, sort of like pudding or Jello, then what gives solid things their shape? What defines their boundaries? Why doesn't everything sort of flow into one another?

Many physicists in the 19th century, especially after Maxwell's triumph, definitely preferred the continuous universe theory. The mathematics that Newton invented to describe gravity and other forces, the math we call Calculus nowadays, assumed that nature was continuous, not discrete; and this mathematics seemed to work just fine, as far as anyone could see. Though Newton championed the corpuscular ("bullet") theory of light, he also proposed that gravity somehow created a continuous field of force that allowed the Sun to pull the farthest planets into their orbits.

If a force like gravity fills space with a continuous field, shouldn't energy also be continuous?

Maxwell's equations, based on the assumption that electricity and magnetism are continuous fields just like gravity, showed that light should be continuous, too. And if everything else is continuous, why not matter?

Chemists of the same era had a very different opinion, however. By the beginning of the 1800s, people like Dalton and Lavoisier had begun to have great success first in isolating the basic chemicals of nature, which they called *elements*, and then in figuring out how all the stuff of the universe was made up of these basic elements.

All in all, it turned out that there were about ninety different kinds of basic elements found in nature. Now, ninety sounds like a lot of different elements (the ancient Greeks had been happier with just four — earth, air, fire, and water[22]); and yet, ninety is still a limited number. There may be lots and lots of ways to put ninety elements together, but lots and lots is *not* the same as infinitely many.

There's not an infinite number of different elements. Chemistry may be complicated, but if you work at it long enough eventually you ought to be able to figure everything out.

And one of the first things they figured out was that the relative weights of different elements in common substances were surprisingly simple. Look at water, for instance. By running an electric current through

---

[22] The Greeks, in their commonsense way, were perfectly correct; only now we refer to four *phases*, rather than four elements. Earth, water, air, and fire correspond to solid, liquid, gas, and plasma.

Of course, the difference is that any substance can come in any phase; it all just depends on its temperature. Heat up a chunk of iron, for instance, and eventually it'll melt; then vaporize; then if it's hot enough, like on the surface of the Sun, even the iron gas can become ionized and glow. On the other hand, chemical elements don't change (short of a nuclear reactor). As far as 19th century chemistry could tell, no matter how hot you heated it, iron would always still be iron; it'd never change itself into oxygen or bromine or gold.

water, you can split it into two gases.[23] What becomes obvious, right away, is that the volume of hydrogen you get will always be twice the volume of the oxygen (assuming both gases are kept at the same temperature and pressure). Exactly two to one. Then if you weigh the gases you find that a given volume of oxygen weighs eight times as much as the same volume of hydrogen.

Hydrogen peroxide is slightly different; split it up and you get equal volumes of hydrogen and oxygen. But the weights of the equal volumes keep the same proportions, eight to one. Ammonia is slightly different yet; the gases in this case are hydrogen and nitrogen, and the volumes of the gases come off at three hydrogens per nitrogen. The weights of given volumes of nitrogen and hydrogen come in a ratio of seven to one.

It works for water, peroxide, ammonia, and nearly every other common substance. (The exceptions are organic compounds, which turn out to be much more complicated; it was another hundred years before chemists really started understanding them.) One to one, two to one, three to one by volume; eight to one, seven to one in weight… what could be going on that allowed such simple ratios of stuff?

If nature were made up of continuous fluids, you'd expect a continuous range of substances. Indeed, you can dissolve peroxide or ammonia into water in any proportions you want, but the properties of that mixture showed that these mixtures were not pure substances[24]. You can't make a pure substance somewhere between water and hydrogen peroxide

---

[23] It's real easy to do. When I was a kid, I used to rip apart the innards of D-cells to get at the carbon rods inside, and then hook these rods up to the transformer from my electric train set. Soon I'd have hydrogen and oxygen bubbling away into glass jars. The whole thing was incredibly dangerous, now that I think of it; if these two gases had ever mixed back together they could have gone off with quite a bang.

[24] What's "a pure substance"? Among other things, a pure substance melts at one specific temperature, but a mixture melts over a range of temperatures; think of the way mixing salt into snow produces slush, not pure water. And when the mixture refreezes, you often find the two substances become unmixed in the process.

that has a chemical formula of 1 part oxygen, 1.414213562… parts hydrogen. Hydrogen to oxygen comes either one to one, or two to one.

The easiest way to explain this is to propose that the elements come in discrete lumps, which the 19th century chemists called *atoms*, a word first proposed by the ancient Greeks. Molecules of water, for instance, are made when you physically link together one atom of oxygen with two atoms of hydrogen. And each oxygen atom weighs as much as eight hydrogen atoms.

The more you split substances up, the more you can make sense out of how they are put together. As the chemists found great regularities in the chemical properties of these atoms, the Periodic Table was discovered.

Meanwhile, geologists studying minerals had come to the same conclusion. A perfect crystal can only be sliced in a few given directions. If you try to split a diamond or a hunk of quartz at some unnatural angle, it won't break clean; at best you'll get a rough surface that shows, under a microscope, alternating faces stepping back and forth at the crystal's natural angles.

The mineralogists also explained this in terms of atoms. If crystals are made of atoms arranged in a regular, repeating order, like layers of ping-pong balls in a box, then you can get flat surfaces like crystal faces only if you split the crystal along one of these layers. The "continuous fluid" model for matter implies you should be able to carve up stuff in any direction you want, but in fact that just doesn't happen in nature.

The evidence from chemistry and geology was overwhelming: matter comes in lumps. But physicists didn't necessarily pay attention. "Real science is *physics;* the rest is just stamp-collecting," sniffed one prominent 19th century physicist (Lord Kelvin, who should have known better). The regularities found by those "stamp collectors" who merely observed the elemental abundances of chemicals or the physical shape of crystals, without benefit of Calculus or other higher mathematics, somehow didn't count.

So what changed their minds?

Let's go back again to James Clerk Maxwell. Back in the 1880's, after his triumph with light, he turned his attention to the place where physics, mathematics, and chemistry meet: thermodynamics.

A story goes that one day Maxwell was visited by the president of Yale University. In those days America was still a backwater, and so the president of Yale had come to Europe to try to hire the best brains available, to raise the tone of his university. After Maxwell turned him down (he had his own lab at Cambridge University and no desire to leave it), the esteemed president then looked for Maxwell's advice on who else to try. "What's the hottest topic in science today?" he asked. (Or words to that effect. This isn't history, it's story telling.) Maxwell, excited about his latest work, replied, "Mathematical thermodynamics."

"Oh," said the president of Yale. "Uh, huh. Right. So, tell me… who besides yourself is tops in this field right now?"

Maxwell replied without hesitation, "J. Willard Gibbs. He basically invented the field; I'm just following in his footsteps."

"Great!" said the head of Yale. "I'll offer anything he wants! We need him at Yale! Where can I find him?"

Maxwell paused, not sure how to break the news. "I'm sorry… I guess you didn't realize." The president looked back, fearing the worst. "You see," Maxwell continued, "he already teaches at Yale."

What Gibbs started at Yale, and Maxwell pushed forward in Cambridge, was a mathematical way of predicting what should happen when chemicals react.

Think of a gas; it will have a certain temperature, and pressure, and density, all quantities that can be assigned numbers and related to each other using equations. If you add energy, you can raise the temperature and that will in turn change the values of the other quantities. If you have two gases and let them react, that can release energy in some cases, or absorb it in others. So a reaction can change the temperature, releasing heat (energy); or pumping heat (energy) into a system can change the temperature, driving a chemical reaction. Reducing all of this to math and equations was just the sort of thing that 19th century physicists loved to do.

But there were several commonsense aspects to chemical behaviors that were surprisingly hard to explain in terms of math. If I dump a hot lump of iron into a cold bath of water, and don't let any heat escape, the water will heat up and the iron will cool down, but the total energy stays constant. Fine. So why is it that energy never spontaneously leaves the water and flows back into the iron, turning the water cold and the iron hot? This doesn't violate any law of energy — either case has the same amount of energy. But we know darn well it never actually happens. How come?

Here's another example: say I have a bottle of red gas in a room full of green gas. They're both at the same temperature and pressure. When I pull the cork, the two gases mix. And they stay mixed. No matter how long I wait, the red gas never all finds its way back into the bottle. How come? There's been no change in energy, or in volume, or in temperature; but something fundamental changed when I uncorked the bottle, because that gas has escaped and there's no simple way to get it all back.

The mathematical thermodynamicists discovered they could give that something a name — they called it Entropy — and they could even write down equations that would relate it to the other quantities like heat and temperature. (The most famous are the "Maxwell Relations," not to be confused with "Maxwell's Equations" for electricity and magnetism that we ran into last chapter.) But it was darn hard to figure out just what the heck Entropy actually was measuring.

The atomic theory provided an answer. Whether it was the right answer was another question; but at least it was a possible answer. To see how it worked, let's pretend for a moment that our atoms are pretty big (or our bottle pretty small) and that there's room for only two atoms in the bottle, and room for only four atoms in the room. How many different ways could one shuffle these six atoms?

Well, let's chart out all the possibilities. I come up with the following possible combinations (where R stands for a red atom, G stands for a green atom, and the first two spots in each line represent atoms inside our bottle):

RR/GGGG (This represents the bottle full of red atoms.)

RG/RGGG, RG/GRGG, RG/GGRG, RG/GGGR (In these cases, the bottle is half red and half green.)

GR/RGGG, GR/GRGG, GR/GGRG, GR/GGGR (Just the same as the second line, after shuffling the atoms in the bottle.)

GG/RRGG, GG/RGRG, GG/RGGR, GG/GRRG, GG/GRGR, GG/GGRR (Here, both the red atoms have escaped into the room.)

There's fifteen different ways the atoms can be shuffled (assuming that you can't tell the difference between any given two red atoms or any two given green atoms). But of all those ways, only one combination has both red atoms in the bottle.

That means that there's only one chance in 15 that you'd actually see the bottle full of red atoms. In other words (if the atoms move in and out of the bottle reasonably quickly) during the course of an hour the bottle will have at least one green atom for fifty-six minutes, while there'd be a bottle full of red atoms for only four minutes. Unless you happen to look at the right time, you won't see it.

Now, if instead of six atoms we have six gazillion atoms, counting up the possible combinations gets a whole lot more complicated. Fortunately, a French mathematician (Pascal) with some gambler friends had worked out the laws of chance — probability and statistics — back in the 17th century, so you can figure out these sorts of odds mathematically if you really want to.

I think it's pretty obvious, however, that the more atoms you've got, the smaller the odds are that you're going to get all of them back in the bottle. With gazillions of atoms, you have just gazillions and gazillions of possible combinations, but still only one of those combinations has all the atoms in the bottle; so the odds become gazillions and gazillions to one.

For instance, if we had 60 atoms total, 40 green and 20 red, the odds of getting all 20 of the red ones into the bottle at one time become about one in 4,200,000,000,000,000. Not likely. When you consider that a real bottle is likely to contain, not 2 or 20, but something more like twenty thousand billion billion (20,000,000,000,000,000,000,000) atoms… you get the idea.

Entropy, then, according to this picture, is something related to the odds of finding atoms arrayed in a certain way. The more random the arrangement, the harder it is to tell one such arrangement apart from another, and so the more likely that you'll see something that looks like such a random arrangement occurring. The more special, the more structured, the more specifically defined a situation, the less likely you are to actually see it happen by chance.

That, in turn, depends on the number of particles you're dealing with. The more cards you have, the more different ways they can be shuffled. It's hard to tell one shuffled deck of cards from another; but if, after several shuffles, you deal and discover that the cards are coming out ordered by rank and suit, then you have every right to be surprised (and suspicious).

This description of entropy was one of the grand successes of the atomic theory of gases. If gases were made of atoms, you could consider density to be a measure of the number of atoms in a certain space; temperature would be a measure of how quickly those atoms were jostling about; pressure would represent the force imparted by all those atoms-in-motion bouncing off the walls of the bottle holding the gas. And now, entropy could be construed as a measure of how mixed up different atoms had become.

Even the problem of why heat flows from hot to cold could be explained: hot atoms and cold atoms (fast atoms and slow atoms) have more ways to exist, mixed together, than to exist, separated; and so just like with our red atom/green atom case, it's far more likely to find hot and cold atoms mixed together than separated.

Physicists like Ernst Mach, however, were not convinced. Just because the atomic theory happened to explain everything well, didn't prove that it was true; it might be, as Ernst Mach stated, merely a mathematical coincidence that it worked.[25] This, mind you, came from the same guy who promoted Logical Positivism, insisting that all truth could be proved mathematically! When it came to mathematics overthrowing his pet theory, however, he was all too positive (in his own self-rightness) and less than logical; which is to say, he was only human.

The foremost proponent of the atomic theory, who came up with the mathematical formula for expressing entropy in terms of the numbers and arrangements of atoms, was Ludwig Boltzmann. By all accounts, the battles between Boltzmann and Mach were heated and furious.

Mach was the better debater, and during his lifetime he was held in much higher esteem. However, it's interesting to look him up in modern books; he seems to be associated with all the great claims of classical physics that have subsequently been found to be wrong. One after another of his ideas and assertions have fallen into disrepute.[26]

By contrast, Boltzmann was personally devastated by the attacks from Mach and his followers; he eventually committed suicide. Needless to say, it's his science that lives today… his formula for entropy, fundamental to the modern science of Statistical Mechanics, is engraved on his tombstone.

What made scientists eventually decide that Mach was wrong, and Boltzmann was right? What was the critical experiment that made people realize it was atoms that properly described "stuff," and that Mach's ideas were "nonsense"?

---

[25] "All mystical experience is coincidence; and vice-versa, of course." Or so says Tom Stoppard, anyway.

[26] He is remembered today mostly because his name is used for the "Mach number" that describes the speed of sound; airplanes flying at twice the speed of sound are said to be flying at "Mach 2."

Probably the single clearest demonstration of atoms is a phenomenon called Brownian Motion. Brownian motion has nothing to do with Brownies (it's named for the Scottish scientist, Robert Brown, who first described it); still, let's imagine a church playground filled with about a hundred active 8-year-old girls in their Brownie girl-scout uniforms.

In this playground there are a number of rubber balls; as you can imagine, the balls get tossed around quite a bit and move from one end of the playground to another more rapidly than the girls themselves move. There's also a jungle gym, set in concrete. It's not moving at all.

But in between, imagine there's a picnic table. The table would be much bigger and heavier than any particular girl. But it is still small enough that, as the girls run around in the playground, the table would get constantly jostled. Over the course of an hour it might be shoved back and forth by as much as several feet.

Remember that table.

If Ernst Mach were right, our playground would not have individual girl scouts in it; instead, it would be filled with a uniform essence, say a pudding of sugar and spice. (More spice than sugar, judging from the 8-year-olds of my acquaintance.) If so, then no matter how long you waited, you'd never see the table move an inch. It's only when it's hit by individual kids that it starts to creep, irregularly, back and forth across the ground.

Now, a bird high up in the sky over the playground may be too far away to ever see the rubber balls. It may be too far away to see the individual Brownies themselves. But it might, just possibly, be able to make out the table, which is after all bigger than any of the girls. If so, it can see the table moving back and forth, irregularly, and realize that something interesting is going on in the church playground.

Back in 1827, Brown had turned his microscope on a drop of fluid from a pollen grain and saw tiny particles less than a thousandth of an inch across, jiggling back and forth as if they were being hit, randomly, by fast-moving atoms[27] too small to be seen directly. Fifty years later, when the atomic debate heated up, people remembered this Brownian motion and tried to calculate if the sizes and speeds of the atoms that the theory predicted could actually be responsible for the motion Brown saw.

It took a hefty amount of calculation, and some very careful observations of many tiny particles. But the answers were unambiguous. The atomic theory worked just as predicted; it explained exactly, in all details, the sort of motion in the dust particles Brown had seen. And there was no way that Mach's field theory could explain the same phenomenon.

So, finally, the physicists were convinced. Energy and electricity and gravity and light may be continuous fields, but matter was different: it came in atoms. However, in the process of coming up with an atomic theory of gases, a subversive new notion had snuck itself into the physicists' way of thinking: *Probability*. Instead of calculating the way the universe *had* to be, they were now reduced to merely figuring the odds on the *most likely arrangement* of the universe.

To understand the impact of this change, consider by analogy the way we understand right from wrong. According to thousands of years of tradition, right and wrong are moral norms that speak to something in each human being beyond the mere material, from the vague stirrings of a conscience to the specific laws handed down by a transcendent God on Mount Sinai. It's knowing right from wrong, we believed, that made a human being different from a mere animal.[28]

Now, suddenly, imagine that we're told that right and wrong have no intrinsic meaning, but rather they're just a description of the most probable behavior of the majority of people, the majority of the time. (In fact,

---

[27] Or, as we now know, atoms linked together into a molecule

[28] A favorite cartoon: A cat is staring at an open can of tuna fish on a table, where it knows it is not allowed. Over the left shoulder, a little devil cat is whispering in its ear: "Go ahead, steal that fish." Over the right shoulder, an angel cat responds: "Sounds good to me!"

that's a grossly simplified description of some moral philosophies at the end of the 19th century including, ironically, the logical positivists.[29])

But the difference is even stronger. Say, instead of a Sigmund Freud or a John Stuart Mill telling us our morality is not based on absolute truths, imagine that God Himself were to return to Mt. Sinai, plop down Moses among us to replace the Ten Commandments with Utilitarianism.

That's what it felt like. The change of physics from one of absolute calculations to one of probability was based, not on some outside radical philosopher, but on the very same conservative experimental data that had been the foundation of deterministic physics for the previous two hundred years.

Of course, there seemed to be an out.

The fact that our description of physical systems was reduced to merely finding the odds of the most likely outcome was, the 19th century scientist insisted, only an expression of a temporary ignorance.

We wouldn't need to use probabilities if we actually knew the specific position and speed of every atom in the universe, and a complete description of every force acting on each atom. Obviously at the present, that was impossible; but with better and better measurements, they still insisted, one could get closer and closer to such a complete description of the universe.

The triumph of deterministic science, though somewhat delayed by our realization of just how many atoms and how many forces there were to consider, nonetheless was merely a matter of time. Progress was inevitable; this was, after all, the 19th century.[30]

Determinism is a deceptively logical way to view the world: Everything that exists (it begins), exists as atoms; every atom must obey the

---

[29] Quite possibly this similarity is just a coincidence. Still, I wonder which idea gained currency first?

[30] A teen-ager, arguing for permission for some questionable venture, accused his parents of being behind the times. "Oh, come on! It's 1995!" he said. "Why is it," his mom replied, "when you want me to let you do something stupid, you always tell me what year it is?"

laws of physics; the laws of physics are rigid, explicit, and completely predictable; and therefore (in theory, even if no human being is smart enough to know the exact order of every atom in the universe and the exact value of all the forces acting on those atoms) the universe itself must be precisely determined by exactly those forces and that arrangement.[31]

Every thought we have, is expressed by a slight change in the arrangements of the atoms in our brains. Every emotion we feel, is precisely related to (and indeed could be exactly replicated, if only we were clever enough) the levels of certain chemicals in our bloodstreams. Our senses of right, wrong, good, evil, beautiful, ugly, love, indifference, are all the states of flesh-and-blood human beings.

And the more we understand the physics that goes on in our bodies, the chemistry of our flesh and blood, the more we should be able to recognize and predict and even manipulate those concepts. Or so the Determinists believed.

It also means that, by knowing exactly how everything is arranged at any given instant and exactly all the forces that are acting on every thing, it should be possible to use the laws of physics to predict exactly everything else that ever will happen in the future… including the thoughts and feelings of every person that ever lived, or ever would live.

Naturally, in such a world-view there's no room for free will. But there still might be room for God (or so many Determinists thought, at least into the start of the 19th century). After all, there were certain nagging things that physics didn't get exactly right, or couldn't explain.

Some were obvious — how did the universe get started, and who set the initial conditions on which everything else was determined? Some

[31] As it happens, each of the assertions of 19th century Determinism described here came under serious, and sometimes fatal, challenge in the 20th century. Where the temptation of the 19th century was for science to believe that Determinism proved there was no God, the temptation of the 20th century has been to think that the collapse of Determinism is the basis of a new kind of theology. Neither one works.

were more detailed — why were the motions of the planets slightly different from Newton's predictions?[32]

Those of us living in a religious tradition will immediately recognize, however, that a God who did nothing more than set the world going, and then fudged a few laws here and there to cover over the things physics couldn't quite get right (yet), was a God very different from the personal God of our own religious experience — and certainly different from the God of Scripture. He'd been reduced to a God that fit in only where our physics needed Him: The God of the Gaps.

And by the nineteenth century, a lot of those more detailed questions were beginning to be solved. The gaps were being closed; and it looked as if God were getting squeezed out of the universe.

Part 2: The Photoelectric Effect

Most of us are familiar today with the modern picture of an atom; the smallest bit of stuff, but itself made up of protons and electrons and other things with exotic sounding names.

Electrons, at least, we're used to; they're the stuff of electricity. Robert Millikan, around the turn of the century, showed conclusively that electrons come in individual lumps by putting a static electric charge on tiny oil drops and measuring how strong an electric field you needed to hold the droplets up against gravity. It was clear that the amount of electric charge sticking to the droplets came in discrete lumps, not a continu-

---

[32] The French mathematician Laplace finally solved the orbits problem, just by doing the math more carefully; Newton really had assumed God intervened. When asked by Napoleon about whether God controlled the planets, Laplace is said to have replied, "I have no need for that hypothesis."

The 19th century atheists were delighted at dismissing "God" as a mere "hypothesis," and an unnecessary one at that; 19th century theologians were appalled. Both sides missed the point. Laplace was right. What he had rejected was both bad physics and bad theology.

ous gradation. This fit well into the newly-adapted atomic theory of matter (and it won Millikan the Nobel Prize).

As scientists got more and more familiar with electricity, spurred on by the commercial possibilities that financed power companies and electrical consumer goods (the bankers who paid for Thomas Edison's experiments made back their money by setting up General Electric), the properties of electricity were seen as a wonderful way of probing the mysterious world inside matter.

And one of the great mysteries, that surely was telling us something about how matter was put together, was the fascinating way that certain materials could generate electricity when they were placed in a beam of bright light. Light is the flow of energy; electric current represents a flow of energy; clearly something interesting was going on inside these materials to change one kind of energy into another.

Since you were converting one kind of energy into another, there are certain general rules of behavior that you might expect to see in this "photoelectric effect." First, you'd expect the brighter the light, the more electricity you'd see. And, if you got any electricity at all out of your experiment at all, then upping the brightness gave you more. No surprise there.

Second, recall that blue light had more energy than red light; so you'd expect blue light to give you more electricity. And certainly there was nothing inconsistent with that, when you ran the experiment.

Finally, you know that the energy you get out in the electricity is probably going to be a little less than the energy of the light you put in; life is never 100% efficient.

Well, when they did the measurements it turned out that there was a certain "threshold" inside the material; you had to have a certain amount of energy in your light hitting the photoelectric material before you'd see any electricity coming off. The threshold varied from material to material; some stuff gave you electricity more easily than others. No surprise there; that's how we're used to nature working.

Except… "if you got any electricity at all"? Why would it work some times, and not others? The blue light experiment showed you "nothing inconsistent" with the theory? What kind of weasel words are those? And finally, what was this about a "threshold" effect? That sounded a little bit different from nature just being less than 100% efficient.

What precisely did those experiments show?

Even the faintest blue light, of a certain wavelength, vibrating at a certain frequency, gave you a small amount of electricity; but light below that frequency, light more red, never gave any electricity at all. No matter how bright the red light. No matter how long you waited. You went from reasonably good efficiency, to absolutely zero, all in one step.

Savor this paradox another moment. Bright red light won't release a single electron, but dim blue light releases plenty. If light interacts "rationally" with matter, shouldn't we see some gradual change as we gradually change the wavelength? But we don't.

Even stranger, when you put a blue light on the stuff you got electricity right away. The material didn't seem to be "storing up" a certain amount of energy before an electron was kicked off. It either happened right away or else, if the light were too red, it didn't happen at all.

Who could explain all this?

Back in 1905, there was a recent physics doctoral graduate working at the Swiss patent office, who had a string of neat ideas. The first concerned Brownian motion, which we discussed above. The second idea was his explanation of this "photoelectric effect;" it won him the Nobel prize. (And a third idea, in case you haven't guessed by now, was Relativity. The patent officer was, of course, Albert Einstein.)

How did he do it? Really, Einstein invented nothing new here. Instead, he took a bunch of ideas and equations that had been sitting around for a while, and just thought about them in a new way.

The key work, in fact, had been done some ten years earlier by Max Planck. It all centered around a number, now called Planck's Constant, that connects the frequency of a light wave to its energy.[33].

The photoelectric effect raised questions about the nature of light, because there were puzzling things about how light could produce electricity. But the fact is, there were equally puzzling things as well when you went in the other direction, and tried to explain how electricity could produce light — in other words, in the physics behind that invention called the light bulb.

The principle behind the light bulb was no secret. If you pass enough electricity through anything, it gets hot. And if it gets hot enough, it will start to glow. Thomas Edison's genius (besides talking bankers into supporting his experiments) really was 99% perspiration, just like he'd said: the hard work of simply finding the best stuff to run the current through. His carbon filament was cheap, long-lived, and relatively efficient.

But once we had light bulbs — that was Edison the engineer's job — the question arose, just how the heck did they work? *Why* does something that's hot emit light?

Classical 19th century physics thought it could handle that problem.

Heat and light are both energy. And energy is stuff in motion. A hot object has all of its molecules, and the atoms in its molecules, and the protons and electrons that make up those atoms with their positive and negative charges, all jiggling around. The hotter it gets, the more they jiggle. When protons and electrons jiggle, they set up a jiggling electric field. And an oscillating electric field, according to Maxwell's equations, creates an oscillating magnetic field; the two, together, propagate themselves as wave of light.

That even explained (sort of) the color of the light. As stuff warms up — coals in a fire, or pigs of iron in a furnace, or the wire in a bulb — it

---

[33] When I was an undergraduate at MIT I met a Wellesley College student who was the great-niece of Max Planck. Her father was also a physicist; I imagine it must have been hard to live up to a name like Planck. So maybe he can be forgiven his odd sense of humor. He named his daughter Constance.

first emits mostly infrared light; then as it gets hotter, a dull red colored visible light; and then the colors progress on through the rainbow getting more and more blue as the glowing stuff gets hotter and hotter.

Sure, that makes sense — the more heat, the more things jiggle; the faster they jiggle, the more frequently the electrons will bounce back and forth, and thus the higher the frequency — the color — of the light they emit.

Sort of. Again.

Because, you see, this leads to a much more subtle problem. Anyone who's taken indoor pictures with color film knows that ordinary light bulbs cast an orange tint — i.e., they're much cooler — compared with daylight.

But notice, it's really only a difference of shading. You still have all the colors of the rainbow coming out of a light bulb, just as you have in sunlight. Granted, a light bulb gives you these rainbow colors in different proportions, more orange than blue. But if the main color of a light bulb is orange, if the atoms in the wire are vibrating at orange's frequency, why are you getting *any* blue light coming out at all?

When you heat up a wire to a specific temperature, it radiates light at a whole range of colors; yes, one color more than the others, but all the others are also included. Okay, so you have a "distribution" of temperatures, maybe, centered around some average temperature. But…

First of all, classical physics couldn't come up with any reason why one wavelength should be favored over another. (Adding heat should just be changing the intensity, not the color.) And if all wavelengths were equally likely, while the energy of the light depends on the wavelength of the light, then all the energy in the hot filament would be carried away by a few very high frequency light waves, out in the ultraviolet.

In fact, if all wavelengths are possible, then why couldn't you make some light with infinitesimally small wavelengths, carrying away infinitely large amounts of energy? Yet we just don't see that happening.

In any event, such infinite energies would be absurd, of course. But they seemed to be predicted by the equations. And no one could find a mistake in the math. The scientists of that era called this problem, "The Ultraviolet Catastrophe."

Instead, what you see is the distribution of energies spread out over many frequencies, but very little of the energy going out at really short or really long frequencies. Most of it is distributed around some in-between frequency.

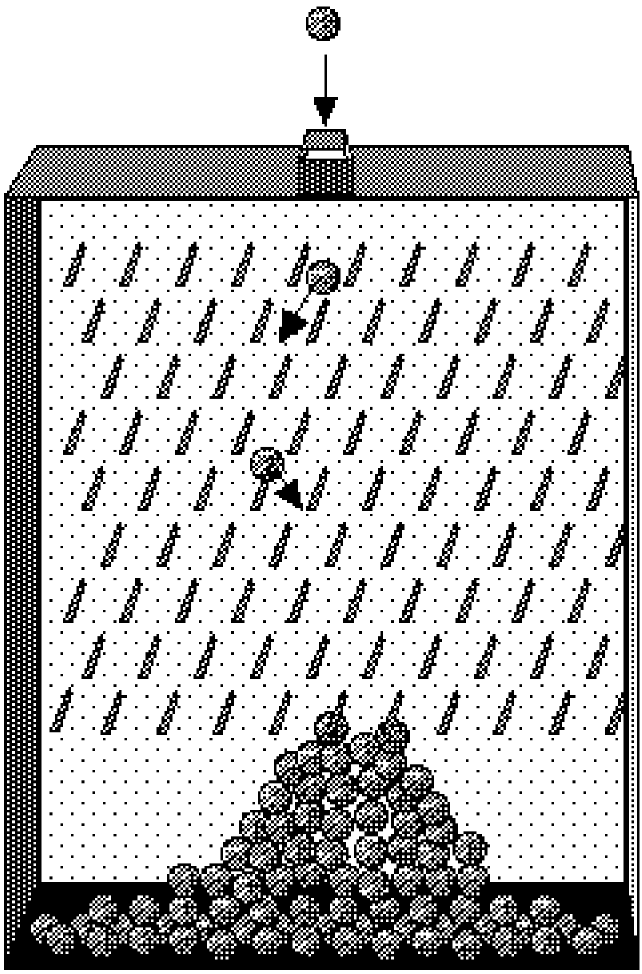
And this raises an even nastier problem that comes up when you assume that all the different frequencies of light follow some sort of distribution curve. It comes from the nature of these curves themselves.

Imagine the different energies were spread out into a smooth "bell-shaped curve" like the "grade curve" in a class, where the "A" students represent the light with the most energy and the "D" students those with the least, while the "B" and "C" students have their energy, their heat, sitting on either side of the average temperature.

This kind of curve, technically called a "Maxwellian distribution" (you guessed it, the same James Clerk Maxwell again) arises all the time in nature. It's what you get when you've got a *probabilistic distribution of individual particles*.

Maybe you've seen what I mean demonstrated in a junior-high science fair. You take a board, and hammer a bunch of nails halfway into it. The nails, each sticking out about an inch, are spaced an inch apart in many overlapping rows. Lean the board up on its side and then, starting at one spot at the top of the slope, pour a bag of marbles down through the nails.

The marbles bounce back and forth in this primitive "pinball" machine (the nails are the "pins") until they reach the bottom. Hold the marbles at the bottom of the board, and let them pile up. When they do, you'll find the balls distributed in a "bell-shaped curve." The top of the curve is placed right under the place where you poured in the marbles, but the pile tails off in either direction.



Now, instead of marbles, if you poured water down the board you wouldn't see any such curve. A continuous fluid like water does not go through a series of discrete, individual, probabilistic events like a falling marble hitting a nail and bouncing either one way or the other. Instead, a stream of water at the top will still be just a stream of water at the bottom. If you catch it, it'll spread out evenly into all positions. In neither case does it create a bell-shaped curve.

The point is this: getting a bell-shaped curve indicates that you've got a collection of individual, probabilistic events, not a smooth continuum. (The Maxwell distribution of atoms' speeds was another death blow to Ernst Mach's ideas.) If you're seeing light come off with a bell-shaped curve to its energy levels, then you're seeing light come off with discrete units of energy, some more probable than other.

Mind you, any wavelength is still possible. But it turns out that the amount of *energy* collected in any given frequency comes in discrete lumps. Each lump (the Latin word is "quantum") is equal to the frequency of the light wave, times that number named after Max Planck: Planck's Constant. At any given frequency, you might see Planck's Constant times that frequency's worth of energy; or twice that amount; or three times; or four, or five, or 2,345,678 times; but you'll never see one and a third times that amount of energy (or 2,345,678.12 times, either).

Well, maybe that's not so surprising. The light is coming from the vibrations of discrete atoms in the light bulb filament (or wherever); maybe that's why we're getting discrete amounts of energy. If one atom gives off Planck's Constant-times-frequency energy, two atoms would give off twice as much, and so on. That solves the ultraviolet catastrophe.

But it doesn't solve the problem of how a photoelectric cell worked. So Einstein looked at it in a very different way. The Planck law could just as easily be describing, not the energy of individual atoms, but the energy of *individual bits of light*.

Mach had argued, incorrectly, that all nature must be continuous; Einstein suggested instead that along with matter, light too might very well be made up of individual lumps. He had revived the old "light-bullet" theory.

It explained the photoelectric effect. Each chunk of light, each *photon*, had a certain energy. If the energy was enough to knock off an electron, then as soon as it hit, that electron was history — one photon in, one electron out. If that one photon didn't have enough energy, then ten million just like it wouldn't, either… a whole string of cream-puffs would have no effect. Thus bright red light would be useless in making electricity, whereas even a weak blue light succeeds.

But, wait a minute. Doesn't this just bring us back to where we started in Chapter One? After all, there was a good reason why the old bullet theory had been rejected — light clearly behaved like a wave. Remember all those interference and diffraction experiments? And Maxwell proved that light obeyed a wave's equation, didn't he?

The classical physicists who adopted the new photon theory were able to finesse their way out of that one, however. Each lump of light, they said, was just a short piece of a wave.

In fact, that helped them squirm out of an embarrassing flaw in Maxwell's wave theory that no one had really had the guts to face before. You see, the equation for a wave that Maxwell solved for, is the equation for a wave that has no beginning and no end. But real waves aren't like that at all. You can turn a light on and off. And the equation for a wave that starts and stops is quite a bit more complex than the equation Maxwell had come up with.

It turns out there's an easy way to solve part of the problem. You can break up a single, infinite wave into an infinite string of tiny, repeatable bits by combining that first wave with a second wave at a slightly different frequency.

Remember interference, from the first chapter? If two wave crests meet at the same time, the crests add together and you get a bigger crest. If a crest meets a trough, they cancel each other out and you get nothing. Well, if one wave is slightly longer than another, for a while their crests will add together, but eventually one will get ahead of the other and the two waves will cancel for a while. Then, yet later, they start to add together again.

Piano tuners know this effect well. If you ring a tuning fork and then pluck a piano string that's slightly out of tune, the two tones will get loud for a bit, then soft, then loud again, in a waa-waa-waa sound that is called a "beat" frequency. (You can try it yourself with two strings of a guitar, played slightly out of tune.) Only when the piano string exactly agrees with the tuning fork will the beat sound go away.[34]

Thus an infinite string of Einstein's photons could just be thought of as a beating of two light waves, at slightly different frequencies. But how could you go from an infinite string of beats, to one individual photon (or at least, to a string of photons with a definite beginning and end)?

Well, instead of a real piano, think about a stereo loudspeaker playing a recording of a piano — better yet, a concerto with piano and orchestra. A typical loudspeaker is just a metal plate glued to a cone of heavy paper, made to vibrate by an electromagnet. Yet isn't it amazing that a single cone of paper can vibrate, simultaneously, at all the different frequencies that go into the music of a full orchestra? It seems so unlikely that a lot of people thought Edison's first phonograph was a fake.

But then, for that matter, how can our single eardrum vibrate in a way that matches all the different vibrations of all those different instruments — on tape or live — so that we can hear them all playing at the same time?

The answer had been worked out in the early 19th century by a French mathematician named Jean Baptiste Fourier. He proved mathematically that any squiggly curve, no matter how complex, can be broken into an infinite series of different perfect simple waves, each with a different frequency, added together…

Maybe you'd have a whole lot of one wave at one frequency, and not very much of another at a different frequency. But if you adjusted all

---

[34] In fact, that's how Marconi got his radio to work. The radio waves were traveling at a frequency of hundreds of thousands of cycles per second, far too fast for the human ear to hear. But Marconi built a rig inside his radio that produced a constant frequency just a little bit different from the frequency of the radio waves. The "beat" between these two waves could be adjusted until the loudspeaker (or headphones) vibrated at a beat whose range matched a frequency that the human ear could hear.

The same principle is still used in ordinary radios to this day. In fact, the weak signals put out by such oscillators can be detected by someone outside your house, or car, who can use them to find out what station you're tuned to.

the amplitudes of an infinite series of such waves, you could eventually get a result that looked just like the incredibly complex squiggle.

This mathematical theorem seemed, like most mathematical theorems, to be utterly useless to anyone except another mathematician. But, like many such mathematical theorems, it turned out to be incredibly useful to scientists and engineers a hundred years later.

Music is an incredibly complex set of sound waves, all added together. But even though a simple loudspeaker cone vibrates only like a pure wave, it can vibrate at any possible frequency, and at many different frequencies at the same time; thus it can reproduce any possible complex wave.[35]

Now, think back to Einstein's photon. It only lasts for an instant. It's a wave that doesn't exist for most of space, then blossoms up suddenly at one frequency for just a little bit of time, then doesn't exist anymore. Fourier's math said that this was really a wave made by adding together lots of waves at lots of different frequencies, all interfering with each other and canceling each other out, everywhere, except at the one spot where we see the photon.

The photon, while it lasts, has one particular wavelength; but *because it is finite, it also has a little bit of lots of other wavelengths added in as well*.

It turns out that this obscure little mathematical fact has a radically profound implication for the entire nature of reality. But if you don't immediately see why, don't worry… Einstein didn't notice it, either. Neither did anyone else, for about twenty years.

---

[35] Well, okay, real loudspeakers can't do *any* possible frequency; but they can do a lot of them. And the more frequencies they can cover, the closer they can come to reproducing the full fidelity of the music. That's what audiophiles are paying for when they buy expensive speakers.

Part 3: The Strangeness of Polaroid

We started out thinking that light was made of little "energy bullets;" but we found out that, among other things, it was hard to understand the colors of light with this theory. Then we decided that light was made of waves; but this ran into the problem of the photoelectric effect. The compromise that seemed to satisfy both was the photon, a little bundle of waves that could act like a particle. But even this compromise had a couple of real problems.

The one that attracted the most attention at the end of the nineteenth century was the question of a "medium." If light is a wave, what's doing the waving? Water waves are waves in water; sound waves are waves in air. But what are light waves waving?

The answer you get from taking Maxwell's Equations literally was that it was the electric and magnetic fields that were waving, but this didn't satisfy many classical scientists who, it turns out, were right but for the wrong reason. The arguments about a hypothetical "ether" that carried light and radio waves[36] were influential in developing the theory of Relativity; and Relativity pretty much did away with any such "ether" by about 1915. Of course, that didn't stop two more generations of bad journalists[37] and cheesy science fiction writers from talking about sending radio messages "through the ether," even into the 1950's…

We'll spend a good chunk of Chapter Five talking about the ether, and relativity. For now, I'd like to point out a much more damning flaw in the wave/particle theory of light. It's a problem that was well known but

---

[36] Maxwell himself believed in the ether. Its existence was disproved by, among other things, the Michaelson-Morely experiment; we mention it briefly in Chapter Five, and you can look it up in any popular book about Relativity. After that, other scientists carried on with the idea that the fields themselves were real and explained the existence of light.

[37] Journalists are subjected to merciless criticism by nearly everybody nowadays, but that may not be fair. Simple, honest, accurate, straightforward reporting must be extremely hard to do. Why else would it be so rare?

widely ignored in the late 19th century. It's a phenomenon of light that you can demonstrate yourself with a couple of pairs of sunglasses. And it's one that can be easily explained away mathematically. Because it was so mundane, and so simple to describe mathematically, its fundamental bizarreness basically went unremarked for about fifty years.

Light, you see, has three measurable properties. The first, *intensity* — the brightness of the light — was easily explained by the bullet theory. More light meant more bullets. The second, *color,* required a wave theory. The shorter the wavelength, the bluer the color.

The third property is something called *polarization*. Like a lot of physics terms, this has a popular connotation somewhat different from what a physicist means by it. We think of people being "polarized" when they are sorted into two hostile irreconcilable camps, like liberals versus conservatives, or dog owners versus cat owners. But light waves with different polarizations can coexist with no problem, and indeed they are so similar in every other way that separating different polarization states can often be difficult.

For the physics meaning of polarization, let's try a different analogy. Imagine an order of religious nuns who live together and pray together, but who are engaged in two different sets of work. One group goes out every day into the streets of the inner city, teaching and doing social work. The other group is cloistered, and stays at home praying.

The two activities are about as different as you can imagine, but they're not what you would call "opposites." The opposite of working is not working or, arguably, working to do evil instead of good; the opposite of praying is not praying, and living a life of self-indulgence instead. Working and praying are both good things. They're as different as can be, but they are hardly opposites.[38]

---

[38] Like all analogies, this one breaks down if you push it too far. Anyone who's ever done social work in the inner city will tell you that if you don't also have an active prayer life you'll soon run into despair and burnout. Likewise, contemplative prayer is hard work, offered to help the same humanity that the social workers are working for.

Or consider two old friends, a philosopher and a scientist, strolling down an empty street late at night, contemplating the meaning of life. “Time flies like an arrow,” muses the philosopher. “But fruit flies like a banana,” replies the entomologist. They aren’t contradicting each other; we say, instead, that they are talking at cross purposes.

Likewise, it turns out that light can exist as the mixture of two different varieties, every bit as “crossed” as the two friends. These varieties each move at the same speeds in the same directions, each with its own intensity and color, never bothering each other; but under certain special circumstances they can be sorted out into two separate beams.

The wave theorists thought they could explain these different polarization states; it followed naturally from the nature of waves. And they got so much of it right that they didn’t see the little piece that didn’t fit.

To understand how to picture the “crossed” nature of wave polarization, go back to the lakeshore and watch the water waves coming in. See how, as the wave moves towards you, the water goes up and down.

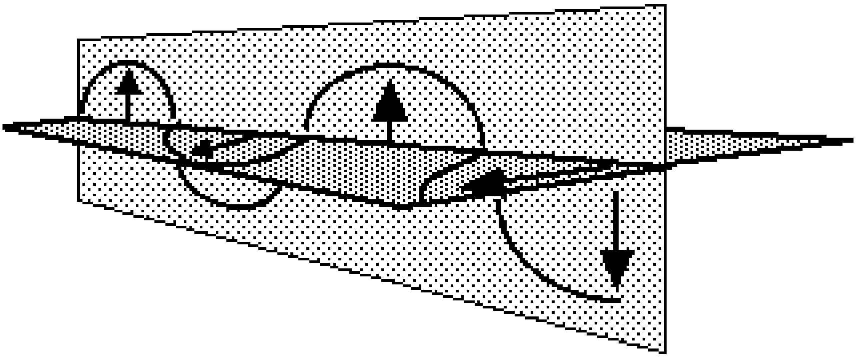
Along the beach, imagine a couple of children with a jump rope. One of the kids runs into the lake and starts to shake her end of the rope, making a pulse that runs down towards the other kid on the shore. With the right thickness of rope, maybe she can even get the wave on the rope to run at the same speed as the wave on the lake.[39]

So the rope can imitate the lake. But the rope can do more.

Suddenly our kid in the lake, instead of shaking the rope up and down, now starts shaking it left and right, back and forth. The rope can have a left-and-right kind of wave, which is something obviously impossible for the wave on the surface of the lake.

---

[39] The waves in the lake, and the waves in the rope, are both moving in towards the shore; but remember, of course, that neither the lake itself, nor the rope itself, are going anywhere. They just have pieces that are moving up and down as the wave moves past.



The up-and-down versus left-and-right motions of the rope are called the two states of its *polarization*. The lake is always polarized into an up-and-down sense. The rope can have waves polarized in any direction, but any angle of polarization can be thought of as a combination of up-and-down plus left-and-right polarizations — or any other two directions crossing at right angles.

Light, too, can be polarized. The beam of light from the sun setting over the lake can have an electric field jiggling up and down; or the electric field might be jiggling left-and-right, instead. But interestingly, the light that glances off the lake — or the hood of a car, or the surface of a slick road — will only have left-and-right motion.

You see reflected light when the electrons in the lake, or the road, get pushed back and forth or up and down by the incoming light, and thus set up their own waves as they themselves move hither and yon.

But when light comes in at a glancing blow, there's not much chance for electrons to move up and down; that would mean that they'd have to jump in and out of the surface of the lake or the road. So that kind of light can't get reemitted very easily. On the other hand, the left-and-right light can make electrons go left-and-right on the lake surface or road surface just fine. So the glare of light off a lake or a street tends to be strongly polarized in the left-and-right sense.

If you had sunglasses that only allowed one polarization of light to get through, you could block out the left-and-right light bouncing off the road. They'd be mighty useful, cutting down on glare while you're driving.

How can you get something that transmits only one polarization of light? Turns out, some minerals are natural polarizers; but mineral crystals have rigid facets and so don't make particularly good lenses. What you need is a polarizer that could be formed, like plastic, into a lens shape.

And back in the 1940's, an inventor in Massachusetts figured out how to mass produce just such a plastic. He took a kind of plastic that was made of long, tangled molecules, and stretched it out in one direction. Electrons tend to move along the molecules, so by pulling all the molecules until they were lined up in one direction, he succeeded in getting all the electrons to confine themselves to that one direction, too.

Like the road surface, those molecules would reflect, or absorb, any light polarized in their direction. It is only the light polarized at exactly *right angles* to those molecules that can pass through the plastic. Edwin Land made a nice little pile of money with his plastic, which he called Polaroid.[40]

What happens, you might ask, to light that comes in neither lined up with, nor at right angles to, the direction of the Polaroid? That's the trick question.

The mathematician's way of looking at it, is to assume that you can divide any direction of polarization into two parts, at right angles to each other. If you are traveling northeast, you're really going a little bit north and a little bit east. So if the Polaroid filter cuts off the northbound component of your light, you'll be left with just eastbound light. In other words, some light should pass through, but not as much as before; and whatever light does pass through, should be purely east-ways polarized.

---

[40] But in the 1940's and 50's he spent it all on some crazy idea for making cameras that developed their own pictures…

Mathematicians call this "resolving a vector into its components." If you're a mathematician, you learned the rules for how to do this in your freshman math classes — it just involves sines and cosines — and so you can blindly apply these math rules to predict how much light should get through for any given twist of the Polaroid.

So how does this mathematical prediction match what is actually observed? Well, the direction of the polarization is exactly what you'd predict: if you shine a light through a Polaroid oriented east-west, you can tell that the light coming out is purely east-west polarized because a second Polaroid filter, twisted at exactly right-angles to the first, cuts off all of this light. None gets through. Thus, the first Polaroid must have absorbed all of the north-south component of the light. Try it yourself, with a couple of sunglasses lenses (the Polaroid kind, of course).

Furthermore, here's the key point: if the second Polaroid is not exactly north-south, but at some other angle compared to the first Polaroid, you'll find some, but not all, of the light getting through. And, to the joy of the theorists, the mathematical theory of vector components, using those sines and cosines, exactly predicts just how much light you will get for any given angle between the Polaroids.

So where's the problem? An elegant mathematical theory makes a prediction; and the prediction is confirmed. What more could you ask for? Well, there's really just one trouble with this theory.

Physically, it just doesn't make sense.

Remember, you are dealing with Mr. Einstein's photons.

Think of it. A stream of photons come merrily along, each one carrying a little bit of wave that has precisely one kind of polarization. Once they've passed through one Polaroid, then all the photons must have the same direction of polarization. Now they're heading straight at a second piece of Polaroid.

If the Polaroid works the way we think it does, these photons should be like keys approaching a keyhole, or coins aimed at a piggy bank slot.

Unless the Polaroid is twisted in exactly the right way, none *of the photons should get through at all.*

And yet, with a partly-twisted Polaroid, some light does get through.

Ok, so maybe there's some kind of complicated interaction going on at the second Polaroid. Maybe the Polaroid absorbs each photon, and uses its energy to emit a new, slightly weaker photon at the new direction of polarization.

The only trouble is, photons don't come in "weaker" and "stronger" varieties. That was the whole point of Einstein's explanation for the photoelectric effect.

The energy of a photon is *fixed* at its wavelength times Planck's Constant. A brighter light means *more* photons, not stronger photons. To say that photons get weaker going through a Polaroid is like saying an off-angle nickel can get through the piggy bank slot, but it comes out worth only three cents. That's nonsense — you can't change the value of a coin by forcing it through a slot; and you can't change the value of a photon by forcing it through a piece of Polaroid.

The fact that we're getting fainter light means that some photons are getting through, and others are not. But every photon with a given wavelength is exactly the same as every other photon.

So how does the Polaroid decide which photons to let through? (Pause here for ominous music!) Is it just random chance that decides which photons shall pass? As far as we can tell, it appears to be just that. There's no way of telling ahead of time which photon will get through, and which one won't.

The truly bizarre nature of polarization is most dramatically shown if you've got a third piece of Polaroid. Take two of them, side by side, and turn them until their polarization directions are at right angles to each other, so that no light gets through. Now slide the third piece of Polaroid in *between* these other two, and turn it about. When its polarization direction is lined up with either of the others', no light gets through. But if it is turned at an angle in between the other two, you will suddenly see light coming through the whole sandwich of Polaroids.

That's utterly bizarre. You've got two pieces of plastic that cut off all light, and now you add yet another piece and suddenly light can come through once again! But of course, notice that it depends on where you put the third piece; it only works if it is between the other two pieces.

All right, so how do we explain all this?

Well, the answer is utterly trivial. The answer is: There is no answer.

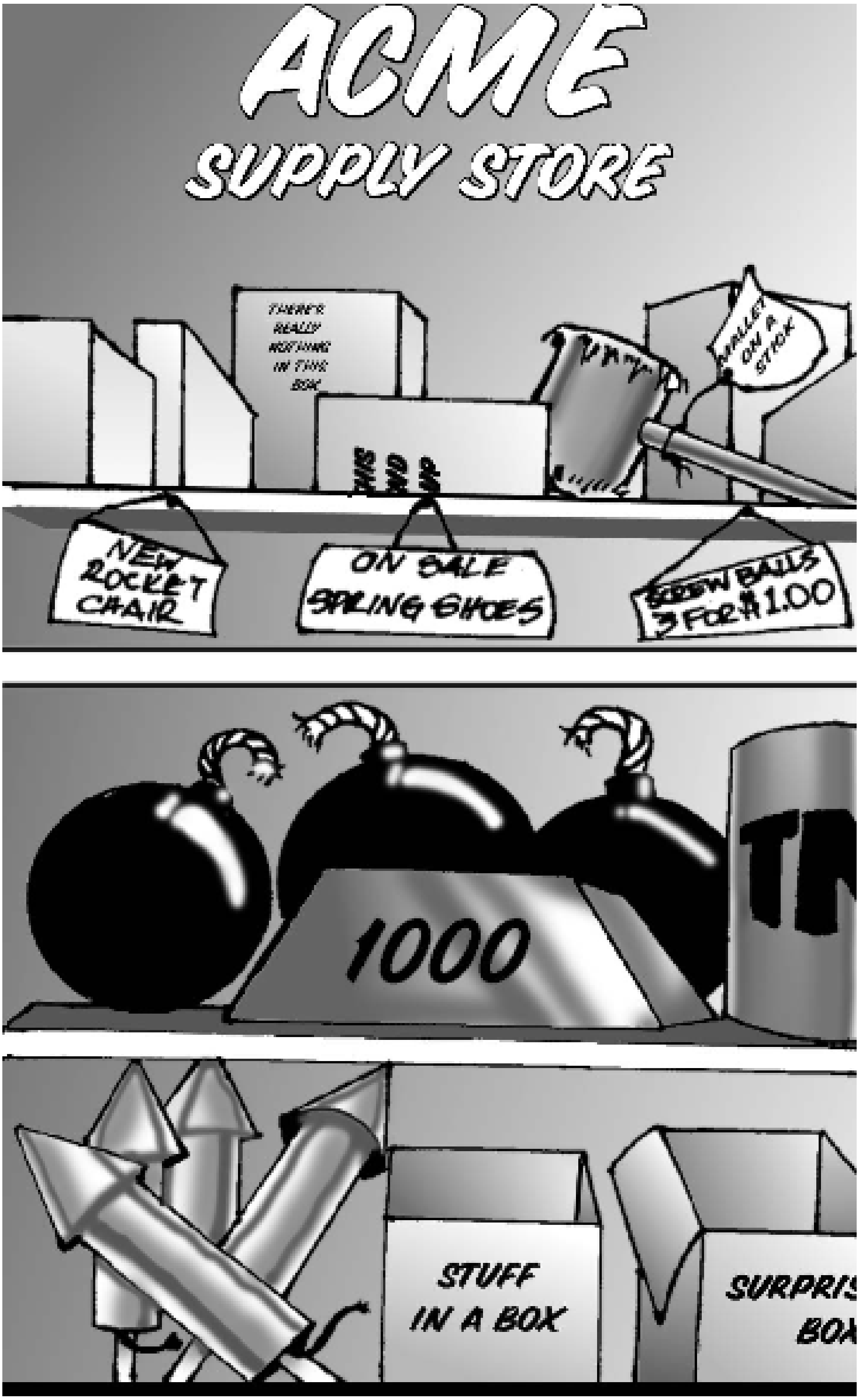
We could explain the intensity and direction of light beams by thinking of little bullets of light. We could explain colors and diffraction patterns by thinking of little waves of light. The photoelectric effect could be explained if we somehow had bullet-like pieces of a wave. But there simply does not exist a simple, familiar, commonsense analogy to explain why some photons go through a Polaroid and others don't.

Light obeys all the bullet equations. But it is not a bullet.

Light obeys all the wave equations. But it is not a wave.

Light is something so completely different that we can never totally picture it in our minds. We have equations that describe its behavior with tremendous accuracy. We can predict it, we can control it, we can make it bounce off mirrors and jump through hoops and do all sorts of useful work for us. And we experience light all the time. It's real.

But understand it? Reality is just a little bit stranger than our common sense can accommodate. You can see why the 19th century materialists didn't like this kind of talk.


ACME
SUPPLY STORE
THERE'S REALLY NOTHING IN THIS BOX
MALLET ON A STICK
NEW ROCKET CHAIR
ON SALE SPRING SHOES
SCREW BALLS 3 FOR $1.00
1000
STUFF IN A BOX

Chapter 4: Particles

Part 1: Dwelling of Light, Next Exit

By the end of the nineteenth century, the atomic theory had triumphed. Everyone knew that matter was made up of atoms. But what were atoms made of?

By bouncing high-speed electrons off atoms, Ernest Rutherford determined that most of the space of the atom was pretty much empty — most electrons passed right through, interacting only with other electrons. However, every now and then one rare electron would bounce off a nucleus, a tiny bit of mass and charge at the center of the atom.

The relative size of the nucleus to the rest of the atom is similar to the relative size of a speck of dust inside a movie theater. Yet virtually all the mass of the atom is in that speck of dust.

In 1915, Niels Bohr commented offhand that the electrons might go around this nucleus like planets go around the sun. Within a year, he realized that this picture really didn't work; but that hasn't stopped eighty years worth of badly written and badly illustrated popular books from repeating the mistake, so that nowadays the "planetary" picture of an atom is one of the most firmly fixed misconceptions in the popular imagination.

Really, we can't blame these books too much. The appeal of this model is that it is easy to visualize. No other model can be pictured quite so quickly. However, reality (as we're beginning to see again and again) is much harder to get a grip on.

In 1895, when Henri Becquerel discovered radioactivity — he'd stuck a piece of uranium ore on top of a photographic plate, and when the film was developed the place where the rock was sitting looked as if it has been exposed to light — he produced a new tool for probing the inside of atoms. And what was seen in the behavior of radioactive substances was all too reminiscent of the polarizer problem with photons.

Some elements contained unstable atoms, he discovered, that decayed into smaller atoms while emitting the "rays" that fogged his film — different types of rays that were given Greek letters for names: alpha rays, beta rays, gamma rays.[41]

If you had a handful of these unstable atoms, you could predict with great accuracy the *average* number of rays you were likely to see at any given time, and you could also predict the *average* energy of the rays. But you could never predict just precisely which atom was going to decay and emit a ray next. And you could never predict just precisely how much energy that particular atom was going to give its ray.

How do they know when to decay? Every atom is just like every other atom; how do they decide which one will decay next? How do they know how much energy to emit now, and how much will be emitted by another atom later on, so that the emitted energy over a long period of time keeps the same average? Either there's something deeper going on — a "hidden variable," some factor we aren't aware of yet — or else, as in Einstein's famous bit of sarcasm, God is playing dice with the universe.

Well, is there a hidden variable? Does the atom know something we don't? Does the photon know something that determines whether or not it passes through a polarizer? That seems like a reasonable assumption; after all, we know we don't know everything.

But think back on those photons. If we start with polarized light, and pass it through another polarizer at some angle to the light's polarization, only a few photons get through.

Let's say those are the magic photons that have the right value of the hidden variable. Then, presumably, that Polaroid has filtered out all the photons that don't have the magic value of the hidden variable, and so if we pass this light through yet another polarizer, its behavior should be different from how it behaved before we filtered out all those "wrong" values.

---

[41] Eventually they figured out that alpha rays are really the nuclei of helium atoms traveling at a high speed; beta rays are high-speed electrons; and gamma rays are very energetic photons.

But that doesn't happen. Just by looking at polarized light, there's no way to see how many times it's been twisted around by other polarizers. There's no way to sort photons (or, at least, keep them sorted) according to their hidden variable. The simplest conclusion, instead, is that there is no hidden variable.

Likewise, for any given pile of unstable atoms, if you wait one half-life you'll find that half of them have decayed. But the half that remain will continue to decay such that, after another half-life, only half of them remain. The ones that decayed first don't remove the "decay now" hidden variable from the system; because eventually, at some later "now," the undecayed ones will decay, too — with exactly the same half-life.

And if you make up a new batch of unstable elements in your local atomic reactor, and mix the newly-created unstable atoms with the leftover undecayed guys from the last batch, the overall half-life of the mixed bunch of atoms doesn't change. Half of them, some old and some new, will decay in one half-life. You can't sort out slow-decayers from fast decayers. If there is a hidden variable, it's mighty well hidden.

The fundamental principle of the classical, clockwork, determinist universe is that everything that happens has a cause; and this cause is physical, predictable, understandable, and immutable. But photons and decaying atoms refuse to behave that way.

This discovery completely destroys the purely mechanistic view of the universe. And it drove the classical physicists nuts. But what was more subtle was the danger that these discoveries would have for the Theists.

So there are places where physics no longer can make specific decisions. Are these the places where God acts in the universe — playing dice or not? This is a tempting hypothesis.

But it smacks of the "God of the Gaps" all over again… and once more, our conception of God becomes hostage to future discoveries in physics, to the risk that something like those hidden variables may some day be discovered (even if we're pretty sure right now that those hidden variables don't exist.)

A different temptation is to throw away not only determinism, but realism as well. Instead of accepting that a single real universe does exist (even if it is not completely deterministic) one might instead embrace a universe that is mystical, paradoxical, and utterly unpredictable. This leads to a sort of New Age anything-is-possible, there-are-no-rules way of living.

But of course, that's not at all what the physics is saying. There are rules. Though I cannot tell you which atom will decay, I can predict with utter certainty precisely what fraction of them will decay, right on schedule. Though I cannot tell you which photon gets through the polarizer, I can predict with incredible precision how many photons will get through. Though, as we'll see, ultimately and fundamentally I have no idea what an electron really is, that doesn't stop me from wiring my house so that every time I flip the switch, the light comes on.

No, our problem is not a lack of rules. It's a problem of understanding *why* the rules work.

A third temptation, one popular for many years, has been to somehow identify the "Strangeness-of-Physics" with the "Strangeness-of-Eastern-Mysticism." Actually, about the only thing these two fields have in common (besides the terminology given to subatomic particles by certain physicists with puckish senses of humor) is that those who connect the two don't understand either of them.

A sort of bastard stepchild of this temptation is the Famous Physicist hawking his popular-science books with the attitude, "Only I can understand this stuff, because I'm smarter than you are." These poses are very popular, because we'd all like to believe that someone, somewhere, has this strange stuff under control. And if it's stuff that only mystics or geniuses can understand, all the better — that relieves the rest of us from the challenge of even trying.

But obviously, that sort of sterile elitism is false on two counts: anyone can grasp the basic concepts to at least some degree; and nobody can ever grasp them fully.

Finally, there's the "dualist" attitude made popular by Niels Bohr and his followers. This has dominated a lot of the discussion of photon and atomic physics. Bohr and his school insisted that the photon was both a particle and a wave — a "wavicle," as some put it — which exhibits one kind of behavior or another, depending on how you poke at it with your experimental apparatus.

This dualist approach has a certain deceptive appeal to those of us brought up on a theology of Jesus as True God and True Man. Indeed, there are some points of similarity. For instance, in both cases, it is important to remember that neither the photon nor the Christ can be thought of as some sort of hybrid, half one thing and half the other, like a centaur or a flying car.

But in fact, what is remarkable are the ways that the photon's dualism is exactly the *opposite* of the two natures of Christ. Jesus Christ, according to traditional Christianity and as specifically defined in some of the earliest and most basic writings of the Church, really was a human being, not just some divinity dressed up in a man-suit. (That's the point of the Christmas season.) And he really was God, not just some very nice, very holy teacher. (That's the point of the Easter season.)

By contrast, it is of fundamental importance to recall what we said at the end of the last chapter: the photon may obey the equations that describe waves, but it isn't really a wave at all; and it may obey the equations that describe "bullets," but it really isn't a bullet, either. It is something fundamentally different from both.

This is a point that even physicists have had a hard time getting.[42] How is it, they would ask, that a universe so tantalizingly understandable

---

[42] The late Richard Feynman was probably one of the first to really grasp and understand this. His book *QED,* a discussion of Quantum Electrodynamics written for the advanced lay person, shows how he approached the problem.

Feynman, Nobel Prize physicist and bongo-player, has become a modern legend; and, for the most part, he deserves his reputation. If you're interested in a fascinating, but ultimately sad, picture of Feynman the man, read his book *Surely You're Joking, Mr. Feynman*.

can be made up of stuff so alien to our experience that we will never understand it?

But here is where our experience as religious believers gives us an advantage. We are used to a universe that is bigger than our understanding, and yet still understandable. We know you can't fit God in a box. If God were not bigger than our understanding, He would not be God.

But, at the same time, if He were not understandable, we could not deal with Him — salvation history, our four thousand years of experience with God from Abraham to last Sunday's sermon, would have no meaning. The God we know in our prayers and in our lives is utterly beyond us, and yet still intimately approachable. And so it is not at all surprising to us that a universe made by such a God has the "look and feel" of its Creator.

Part 2: What's the Matter With the Matter?

No doubt about it, light is truly strange stuff. Much different from ordinary, run of the mill matter… rocks, or baseballs, or bullets. Matter, at least, is made up of mass. It has "heft." It has a physical size. It has a measurable speed. It follows certain very simple rules outlined by Isaac Newton back in the 17th century.

Still, the oddities seen in photons and decaying atomic nuclei began to raise some doubts. In the mid 1920's, a young French nobleman with a keen imagination, Louis de Broglie, suggested a bizarre idea.

Twenty years earlier, Einstein had convinced the world that light came in photons; what was once thought to be a wave, actually was better described as a little lump of a wave. Electrons, on the other hand, were thought of as particles… as if they were tiny specks of charged-up dust, "orbiting like planets" (once a false picture has taken root, it's awfully hard to dislodge) around the center of the atom.

But if light, a wave, also can have a particle nature… could it be that electrons, which are particles, could also have a wave nature? In that case, recall how the wavelength of light was related by Planck's Constant to the energy carried by a photon. de Broglie worked out an elaborate mathematical theory that suggested that the "wavelength" of the electron should be related to its *momentum* — essentially, how fast it is going.

A silly idea, on the face of it. Still, it should be easy enough to test. We know that light is a wave because when you shine a beam of light through two narrow slits onto a screen, you see a diffraction pattern on that screen. What would happen if you could shoot a stream of electrons through two narrowly-spaced slits?

They couldn't actually do that kind of experiment back in the 1920s; there's no way they could make the slits small enough, or the detectors fine enough, to see an effect. But they could get the same effect by bouncing electrons off the face of a crystal. At just the right angle, the hypothetical "wavelength" of electrons bouncing off one layer of atoms ought to "interfere" with the "wavelength" of electrons bouncing off the next layer. Clinton Davisson and Lester Germer (in the US) and George Thompson (in Britain) did the necessary work. *And the interference pattern was indeed seen*. Just as de Broglie predicted.[43]

The electron does indeed have a wavelength. And, turns out, this wavelength is related to the momentum of the electrons, also just as de Broglie predicted. Same holds for protons and neutrons, the guys that make up the nucleus of the atom. Time to pass out another Nobel Prize.

But what does this really mean?

Notice, this is a bit of a different case from the photon. Yes, it does mean that the pieces of "stuff" that make up atoms — subatomic particles — are fundamentally different from our ordinary commonsense idea of

---

[43] Actually, Davisson and Germer were working on a totally different project, trying to understand the structure of atoms by bouncing electrons off them. But they got strange peaks in their results, which they couldn't understand.

Needing a break, Davisson took his wife to England for a second honeymoon and, while there, decided to drop in on a meeting of British physicists in Oxford. At that meeting he heard Thompson use his "strange peaks" as proof of de Broglie's theory!

"particles;" sometimes they behave like solid objects, sometimes they behave like waves. In this, they are similar to photons.

But in other ways they're different. They have mass, which photons don't have. And they can travel at any speed they want, whereas photons can only travel at the speed of light. That's why it's important to say their *wavelength* is related to their *momentum,* whereas for photons we said their *frequency* was related to their *energy*. Momentum is different from energy, and wavelength is different from frequency.[44]

It's not just that the electron has a wavelength; it's this connection between wavelength and momentum that turns out to be the key bit of craziness. So let's back off, and remind ourselves precisely what it is when a physicist talks about "momentum."

We have a good commonsense idea of the meaning of momentum, when used in ordinary speech. But for a physicist, that word has a precise definition. The momentum of an object is its mass, times its velocity. A one-ton Toyota at sixty miles an hour has the same momentum as a two-ton Caddy at thirty miles an hour; if they crash head on, neither wins. (And a Mack Truck at either speed creams them both). It makes intuitive sense. But, what exactly is mass? And what is velocity?

Velocity is how fast an object is moving, *and* the direction that it is moving in. If two guys start out at the same place, and travel at the same velocity, they should always travel together. On the other hand, ten miles an hour north is a very different velocity from ten miles an hour east. So direction matters.

(Energy, on the other hand, doesn't care about direction. It takes as much energy to go 10 miles an hour east as it does west, assuming the road is flat, but you'll wind up at different places.)

---

[44] The relationship between frequency and wavelength is simple enough — the frequency times the wavelength gives you the speed of the wave. That's especially easy for light, where the speed is constant. But for electrons, the speed can be anything; so knowing just the frequency isn't enough to tell you the wavelength.

And mass? Actually, mass is a far trickier thing to define. You can define mass in terms of how heavy an object feels in a gravity field; or you can define it in terms of the amount of momentum an object has at a given speed.

In fact, when Isaac Newton came up with his laws of motion and gravity he made the huge assumption that both definitions are defining the same thing. It's called the Principle of Equivalence. It seems to work, though no one is sure why. But notice, this is circular… we defined momentum in terms of mass, and now we're defining mass in terms of momentum. Like I say, mass is tricky to define.

The fact is, Newton's classical laws of motion are properly understood as being about momentum, not mass. Momentum is the fundamental quantity. And ultimately you can't define it; you have to assume, axiomatically, that you know what it means already.

Newton came up with three laws about momentum that are the building-block assumptions of classical physics:

(1) An object with a certain amount of momentum keeps that momentum, unless you apply a force.

(2) The amount of force you need to apply is directly related to how fast you want to change the momentum. (And since momentum involves direction as well as speed, it takes force to turn something from its path just as it takes force to speed it up or slow it down.)

(3) Whenever one body acts on another to give it a new momentum, its loses just as much momentum as the other guy gains.

Note that the first two define force in terms of momentum, and the third says momentum itself is conserved. These laws can be defined mathematically; they are rigorous, and precise; *and they work*.

The point is, if you know exactly where an object is, and exactly what its momentum is, and exactly what forces are acting on it, you can determine with equations exactly where that object is going to go. And I mean, *exactly*. Sure, we may not be clever enough to make such exact measurements; but presumably the object itself has exact values of those things — position and momentum. Nature should know, exactly, even if we don't.

Thus everything that happens in the physical world ought to be 100% completely and exactly determined by the physical forces acting on the individual particles of the world. That's the whole idea of scientific determinism in a nutshell. The whole nineteenth century was built on this basic premise. Darwinism, Progress, Manifest Destiny were all just ways that the Inevitability of Nature was playing itself out on the human stage.

Pushed to its extreme, obviously it led to all sorts of absurdities and abuses — like Social Darwinism. (If I am richer than you, it's not your fault or mine, it's merely an inevitability of nature… so inevitable that, clearly, there's no sense in me helping you out, or you trying to overturn it!) But, frankly, there didn't seem to be any obvious way around determinism. Because, darn it all, it worked.

Cause and effect explained the motions of the planets, and the workings of the steam engine. It conquered magic and superstition. It made people look for the causes of diseases, and thus try to find their prevention or their cure. Though Social Darwinism was a great excuse for the lazy, seeing society in terms of cause and effect also inspired social reformers to try to fix abuses like poverty, child labor, political injustice. It wasn't all bad. And it isn't all wrong.

But here's where the de Broglie wavelength of the electron plays its havoc. Remember what we said, back in the photoelectric effect chapter, about photons and their wavelengths? A photon can have a finite size only if it is a mixture of many different wavelengths.

The same is true of the electron. If a particle starts existing at one point, and stops existing at another point, it is a mixture of different wavelengths. Which means, *it is a mixture of different momenta*. You have to know the momentum exactly, in order to apply Newton's laws exactly. But no well defined subatomic particle has a single, exact value for its momentum.

Oh, you could try to have a single-momentum wave… but then, it would have to be infinite in extent. That means, you wouldn't know its exact position; it wouldn't *have* an exact position. So you still couldn't apply Newton's laws, exactly.[45]

Werner Heisenberg was the man who realized this conundrum, known today as the Heisenberg Uncertainty Principle. It is important to recognize exactly what this principle is, and is not. It most decidedly does NOT say that "everything is uncertain."

Nor is it a statement of merely human inadequacies, that somehow we just aren't clever enough to be able to make these measurements precisely. Instead, it is a fundamental description of the way nature is put together.

But it is noticeably important only at the tiniest levels of nature, on the scale of the atom. Newton's laws worked fine as far as anyone could tell in the 17th century, and they work just as well today. The macroscopic world we live in *is* a world of cause and effect. Modern popularizers to the contrary, the rise of quantum physics did not overturn Newtonian physics. Instead, it perfected it.

All rules have exceptions (including this one, I am sure); or at the very least, they all have a limited realm of applicability. Quantum physics completes Newtonian physics by outlining very precisely just where Newton's rules fail, and how: on the scale of very minutely defined positions and momenta. (Relativity does the same thing, on the scale of very large distances and very large momenta.)

And yet, at its foundations, at the basis of that from which every object in the universe is formed, our common sense ideas of cause and effect do not hold. Our common sense ideas of position and momentum no longer have their familiar meaning. We live in a universe that we are used

---

[45] Which leads to the famous excuse of the physics student, "I'm sorry, professor, but last night I accidentally determined the momentum of my homework to a very high precision, and now I can't find it!"

to, and can understand; but its underpinnings are things we do not understand, and can never get used to.

It suggests a very familiar Hand at work…

Part 3: Thought Experiments

Richard Feynman, famous physicist mentioned in a footnote above, wrote a series of *Lecture Notes on Physics* back in the early 1960's that are still a standby for graduate students studying to pass their General Exams on the way to their PhD's. (Of course, he thought he was writing for freshmen.) His third volume covers quantum mechanics. As you would expect from Feynman, his description is unconventional but remarkably clear.

What I'd like to do now is borrow an illustration from the first chapter of those *Notes*. Since we're not physics graduate students here, I won't reproduce all the glorious detail of his picture. Still, the essential meaning, and strangeness, of quantum mechanics really does come out if we examine a certain "thought experiment" that he describes in that chapter.

But first, a word about thought experiments.

Having taught freshman physics for many years, indeed having suffered through it myself, I have learned to recognize a fundamental problem that my students (and I) had. We'd sit in the class, hear the lecture, and it would all make perfect sense; but then, when we'd go home and try to do the homework problems, nothing that we heard in class seemed to help. In fact, nothing that we heard in the lecture seemed to connect at all with the problem before us (which of course was due the next morning).

The trouble was that we thought we understood the physics; but in fact, we didn't. We really hadn't gotten used to the concepts well enough to be able to think them through, use them, and feel comfortable with them.

The same thing happens to scientists at all stages of their careers, not just freshmen. Whenever you're faced with a new way of looking at the universe, you can only really get comfortable with it by trying it out in some simple, hypothetical cases. After all, that's how Jesus taught. He didn't just set down a whole lot of moral rules; he told parables.[46]

A thought experiment, then, is just a "parable" for a physicist. It's a simplified "what-if" story. Its purpose is not to prove anything — that's what equations and experiments can do. Instead, it's a device to get you to start thinking of the implications of what your experiments and equations have proved; a way to play with the ideas, until they become familiar. Sometimes it's even a way to think up new *real* experiments, to see if the conclusions of your thought experiment really work the way you think they should.

And, like a parable, it's important to accept it for what it is. The Good Samaritan was a fictional character, but there's nothing fictional about loving your neighbor. The thought experiments of modern physics, likewise, are made-up stories to illustrate a point. They are not proofs; nor does modern physics stand or fall on their being true. But they're a wonderful device for getting used to a universe that, we've learned to see, is a lot stranger than our common sense would lead us to believe.

So, with that in mind, let's walk through Richard Feynman's thought experiment:

What would happen if we *could* shoot a stream of electrons through a two-slit experiment? What precisely would we see?

If the electron really does have a wavelength, you might think we should see interference patterns, just as we saw with light back in Chapter One. But what does an "interference pattern" for an electron mean?

Well, of course, finding interference patterns in light was easy. We know how to create light; we can scratch tiny slits through an opaque bar-

---

[46] Likewise, law students really learn law not merely through legal theory, but by seeing how the law is actually applied in real cases. The same approach is also found in business schools that use the "case study" approach.

rier; and we can use our eyes to see the light falling on the far wall, making those funny dots that we illustrated in Chapter One.

But how do we make a beam of electrons? Or a set of electron-scale slits? And since we can't *see* electrons how do we find our electrons once they've gone through those slits?

In practice, all of those difficulties are very hard to overcome. But in our thought experiment, we can just pretend to trundle off to the Acme Scientific Store (no doubt the same place where Wile E. Coyote buys his Roadrunner-hunting supplies) and pick up a few bits of imaginary apparatus.

First, we'll buy a gun that shoots electrons. (Not so hard to imagine; actually, there's an electron gun in the back end of every cathode ray tube, like a TV screen or a computer monitor.)

Next, we'll get a metal screen with two incredibly tiny slits, very-very close together. (Harder to imagine, and actually not possible to construct until fairly recently).

And finally, we'll purchase a tiny little electron detector, a device that has a tiny bit of area to catch electrons, that goes "click" every time an electron hits it… sort of like an ultra-miniature Geiger counter. (Totally fictitious.)

The idea is that we're going to shoot the electrons at the screen with the slits. In the light case, we had a wall on the other side of the slits-screen, and we looked at light patterns on the wall. Now we'll still have a wall; but in order to see the electrons, we'll need to put our detector on rollers, up against this wall, and move it back and forth, keeping track of how many electrons per minute hit the detector at each location along the wall.

Since electrons are, we're sure, individual bullet-like particles, we have to sit our detector at each point behind our metal sheet, wait some time, count up the "clicks" to find out how many individual electrons hit our detector. Then we move the detector over a bit, and count the electrons hitting at that point; and keep counting, and moving over, and counting some more, until we've counted electrons at every spot on the entire wall.

Better yet, let's get a whole bunch of detectors and line them up, side by side, along the wall. Every time an electron hits a detector, it'll go "click"; and by listening very carefully we'll be able to tell which detector went off. We'll make sure the detectors sit cheek by jowl next to each other, and hope that no electron confuses the issue by hitting *between* two detectors; in any event, we'll devise some sort of scheme so that in the end we'll be sure that each electron sets off one and only one detector click.

We've designed our experiment. Now, our first question is this: Where do we find the most electrons getting through the slits?

Let's start with just one slit, to get a feel for how the apparatus works.

Do electrons only pass through the slit traveling in straight lines, like bullets? Or do the electrons spread out, concentrated mostly straight through the slit but dying away, gradually, in a bell-shaped curve, as you move out of the line of sight? Or do we find electrons hitting the wall in some other pattern?

Your common sense reaction, the classical picture, would be the first picture. You'd expect that when we are sitting in the shadow of the barrier (rather than directly behind the slit) we'd expect to see nothing, but as soon as we emerge into a point exactly behind the slit we'd get a full dose of electrons — either you're in the electron stream, or you're not.

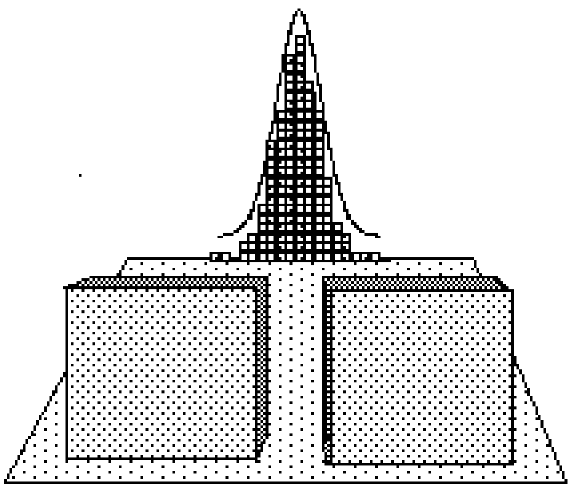
On the other hand, as Feynman points out, if the electron bullets are a lot smaller than the thickness of the screen, you could have these little electrons bouncing off the edges of the slits. That'd spread them out a bit; and so, really, you might rather expect to see a bell-shaped distribution.

In fact, the laws of quantum mechanics also say you'd see the bell-shaped curve.

Its reasoning is that you can never know precisely the left-right momentum that the electron has as it passes through the slit — it's the Heisenberg problem. The instant the electron is in the slit, we have determined its position (at least at that instant) to a precision equal to the width of the slit, and so that means its left or right momentum is a little uncertain. Thus any given electron in the slit might have a tendency to go off a bit to the left or the right.

If we look at lots of electrons, eventually we'll see both left-leaning and right-leaning ones; but of course, most will still wind up going more or less straight through. Hence we wind up with the bell shaped curve again. And, notice, the narrower the slit, the more precisely we've determined the position of the electron; so the wider the range of momenta we'll have. The narrow slit should, ironically, have the effect of spreading the electrons out into a wider bell-shaped distribution once they get through the wall.

So we run our experiment, and listen for the detectors. Click, click… pause… click-click-click… Even though our gun is using a constant amount of electricity, the clicks come at an irregular rate — the uncertainty in their momentum, their speed, makes itself felt. However, every electron that is shot off the gun gets collected; and we never have more than one detector clicking at a time, though some electrons can come one right after another in rapid succession. One electron, one click.



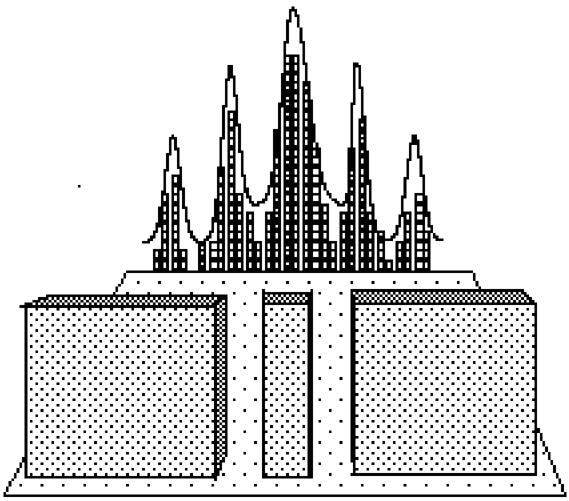
At first it looks like a pretty scattered collection of shots. But after a few hundred electrons have been counted, the rough shape of a bell-shaped curve can be seen.

Now, put in the second slit.

If it is close to the first, you'd expect (given the classical picture) that the two bell-shaped curves will overlap. If they are really close, you can see that the overlap might work out so that the spot exactly halfway between the two slits, presumably in the shadow of the screen between the slits, might actually get more electrons than the spots directly in the line of sight of the slits themselves, since that center area can get fed by electrons drifting in from either slit.

But that's not what you see at all. Instead, if you count up electrons, you still hear them clicking, one by one, irregularly timed as before; but there are certain areas that don't get any electrons at all, and others that get a lot more than you'd expect. These regions alternate, high and low…

Just like an interference pattern.



In other words, *where* it arrives makes the electrons look like waves.

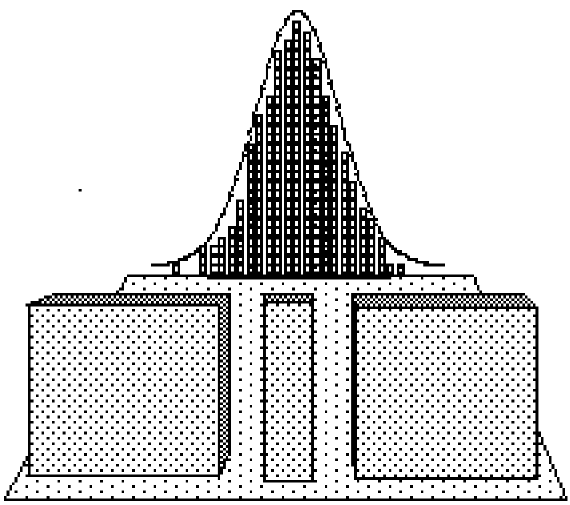
This is truly bizarre. The electrons are still arriving, click by click, as individual particles. *But how does any individual particle, as it passes* *through a slit, even know — much less care — about the existence of the other slit?*

Say we squeeze off just one electron from our electron gun, and we hear just one click at the detector end. Is there any way we can determine which slit the electron went through?

At Mr. Feynman's suggestion, we rush back to the Acme Store to pick up a couple of different electron detectors. These are (mythical) little lights we can set up right next to our slits; as the electron goes by, it will reflect a glint of light and so we can see where the electron is, and thus which slit it went through.

So, we set up our lights and run the experiment again. We shoot an electron from our gun; we see one, and only one, flash of light as it goes through one slit or the other; and we hear one, and only one, click as it hits the wall. The electron is definitely passing through only one slit. And we can see which slit. So why are we getting that strange interference…

Um. Oops. We're not getting the interference pattern anymore. When we install our lamps to see which slit is being used, the electron no longer acts like a wave. The pattern we get is exactly the sum of two bell-shaped curves, just like we expected in the first place.



Quick, we turn off the lamps. *The interference pattern returns.*

Lamps go back on; the interference pattern goes away.

Well, obviously, the light from the lamps is changing the behavior of the electrons. Electrons are very delicate particles, I guess. Maybe if we use less light…

Well, if we turn down the intensity of the lamps, slowly the interference pattern starts to return. But we notice a funny thing happening now; some electrons manage to get through the slits, and make a "click," without being detected by either lamp. Are they traveling through some back door that we don't know about?

No, it's just a problem with the lamps. The light, after all, is a stream of photons; by turning down the intensity, we're emitting fewer and fewer photons, and so increasing the chances that an electron might go by when there's no photon around to glint off it. The fewer photons emitted by the lamps, the fewer electrons seen as they pass through the slit; and the more the pattern at the wall resembles an interference pattern.

Well, turning down the intensity doesn't help; those electrons that do encounter a photon, still wind up being messing up. How about if, instead, we use lots of photons but give each photon less energy? If they're wimpy photons, maybe they won't interfere so much with the electrons.

So, we change the color of the lamps; now they emit a dull red light. We get a flash for every electron that goes past (since there are lots of these photons); and we get the interference pattern that we're looking for, since we've turned down the energy so much that the photons hardly change the electrons at all. So, which slit does each electron go through?

We look carefully; but, alas, the light is so red-colored, so unenergetic, that it has a very long wavelength; longer, in fact, than the distance between the two slits. And as a result, we can't tell with any kind of precision just exactly where the light is coming from. Instead of two fine pinpoints of light, one at each slit, we just see a dull red blob coming from the general direction of the slits. So there's still no way of telling which slit any given electron is going through.

It all boils down to the same principle: an electron is an entity without a precise location or a precise speed. Any time you attempt to determine its location, you spread out its possible range of speeds; any time you attempt to determine its speed (more specifically, its momentum) you spread out its possible locations.

An electron is not a “thing.” Left to its own devices, it doesn’t simply exist in one and only one place, and it doesn’t simply travel in one precise, unique path as it goes from region to region. And any attempt to see what it is doing, changes its behavior.

An electron is not a real “thing;” but it is real. Just touch a live wire if you don’t believe me. It obeys its rules, very strictly and very precisely.

They just happen to be rules that are utterly foreign to our experience of bullets and baseballs and falling rocks.

Chapter 5: Relativity

Part 1: Relative to What?

All the rules are changed; because modern science has proved that everything is uncertain, and everything is relative.

All right. How many falsehoods can you find in that statement?

For starters, think for a minute about the idea that "all the rules are changed." What rules? Certainly the universe itself has not changed since the 17th century, even if our scientific description of the way the universe works (the "rules") has become somewhat more sophisticated since then.

Next, the idea that science "proves" anything is a gross misreading of science. We saw that back in Chapter One. Science describes; it does not "prove" in the popular (or philosophical) sense.

Further, as we saw in the last chapter, the Heisenberg Uncertainty Principle says nothing even close to "everything is uncertain." Yes, there is a fundamental "fuzziness" at the tiniest realm of physics. But Heisenberg puts a strict limit on the size of that fuzziness.

Actually, the rules of quantum mechanics define just how tiny our uncertainty really is. And once we know where the limits are, we can work safely up to those limits to build lasers and miniscule microprocessors that take advantage of the degree of precision inherent in the universe and increase our confidence in its predictability.

The last idea, "everything is relative," is the craziest of them all. That phrase is often used to justify some behavior that has nothing to do with physics, usually something that we have a sneaking hunch we really shouldn't be doing. But why should a principle of physics be blindly applied to morality? Yet even worse, those who use Einstein to defend moral relativism have got his principle of physics exactly backwards.

When you speak of relativity, the real question is, "relative to what?" In fact, it was the old classical Newtonian universe that had a hard time finding a common reference point. Einstein's theory, rather than saying there is no universal yardstick, says exactly the opposite. There is a standard, one that remains the same in every frame of reference.

That standard is the speed of light.

But before we get into describing why Einstein hit upon this universal standard, let's take a step back and recall why it is we need yardsticks in the first place.

In the days of Aristotelian science, where the goal was to describe nature in words, the need for a "yardstick" was not so obvious. Where's the hidden assumption in a simple, clear description of a flower or a rock?

And yet, like Humpty-Dumpty, we all use words to mean exactly what we want them to mean, and rather arrogantly assume that everyone else will read into them the same meanings we intended. Just play this game with any dictionary: Look up a word. Then look up the words used to define that word. See how long it takes before you come back to the first word you started with.

All definitions, all descriptions, ultimately are circular. And thus, every description of nature that uses words has, built into it, the assumption of a common intuitive understanding of what those words mean. If you don't know what he meant by his language, reading Aristotle is useless; it'll be all Greek to you.

When Galileo and Kepler started measuring the universe with numbers, and using the power of mathematics to help them out with their descriptions, the need for a common yardstick became more obvious. For instance, to measure the rate of a falling object you've got to have a yardstick to see how far it falls, and you've got to have a clock to time its descent.

It was the need for a constant time-keeper that led Galileo to discover the regularity of a pendulum's swing; but even then, he could only determine that it was regular by comparing it to his pulse. (He was watching a swinging lamp while standing in a church; I guess even in those days it was hard to find a liturgy that would stir one's pulse to a faster beat.)

One power of a mathematical world description is that numbers have the same meaning everywhere. One man's fish may be another man's *poisson,* but everywhere you go you'll find that 2 + 2 = 4. Yet even math must be based on intuitive assumptions.

Euclid, after all, could only prove his theorems by starting out with obvious but unproved, and unprovable, axioms. It turns out, you can assume different axioms and come up with some really interesting new mathematics (which, we'll see, seem to describe the real world better than Euclid's axioms did). Indeed, a modern mathematician named Gödel has proved, mathematically, that any mathematical system *must* start with at least one unprovable axiom, someplace. (Of course, he had to make some assumptions for his proof to work…)

In fact, the way that logic and reason works in science is often misunderstood. It is a rare advance in science, or philosophy, that is arrived at in a purely rational way. Rather, the scientist has an intuition, an insight, arising out of her instinctive familiarity with a problem. It is only after an intuitive leap has carried her over the river, that she can look around for smaller stepping stones to secure the path and let others follow behind. But all that these stepping stones can do is make the leaps smaller and easier. For, indeed, it takes at least a tiny bit of intuition even just to recognize and admit the truth of any link in the chain of a logical proof.

The danger to physics in its dependence upon logic is that a physicist can sometimes lose sight of the essential nature of insight.

Intuition is central to science. While it's true to insist that science is a description of nature, one must remember that it is a *deep* description of nature. A videotape machine can record the path of a pendulum, but that machine is not a scientist; that machine has no understanding.

Aristotle, Galileo, and Newton, looking at the identically same pendulum, would see three utterly different phenomena. Aristotle would note, quite correctly, that every pendulum eventually slows down and stops; the nature of things to come to rest in their "natural position" was a fundamental point of his physics. For Galileo, the exciting discovery was that the pendulum kept a constant period, even as the swings grew smaller and smaller. Newton carried this insight farther, and insisted that in the absence of frictional forces in the wire and the air, the pendulum would in fact swing with not only the same period but also the same amplitude, forever.

Same pendulum; totally different messages. All of them are true descriptions of the physics, but from Aristotle to Galileo to Newton they became deeper and more subtle.

Other sciences, like geology, are much more aware of the subjective yet repeatable nature of their descriptions. A geologist once told me the story of his experience as a student… having been taken out into the field by his professor, he listened as the prof used the complex bedding and fracturing of rocks in an outcrop to expound on the history of the terrain beneath their feet. My friend followed the explanation with with wonder and amazement.

"When you first looked at the rocks, they just looked like rocks," he told me. "It was hard to see what the professor was talking about. Once I understood the theory behind it, though, then I could see all the subtle cracks and things that he was using to prove his point. Still, if I hadn't believed it, I would never have seen it."

Religion isn't the only human activity that depends on faith.

So what are the articles of faith for a scientist? They're much harder to pin down than they used to be.

For instance, in the 19th century it was assumed without even thinking about it that things happen for a reason; every effect had a cause. At least at the quantum level, that article has been seriously shaken.

We once assumed that if everyone used identical yardsticks, they'd make identical measurements. Relativity will show us when and where that assumption goes wrong.

A principle of science that Kepler stated explicitly was that one cannot come to correct conclusions based on false premises. He then proceeded to start with some truly ridiculous assumptions about the nature of the solar system, and used them to deduce some profoundly true laws of planetary motion — laws that are still the starting point of modern celestial dynamics. Indeed, just the opposite seems to be the case: truth will out, regardless of our starting assumptions, if we approach nature honestly.

But still, there are some principles that continue to hold. The first, and most basic, is called *realism*: the assumption that the universe does exist. You and I are part of the same universe, even if our assumptions and beliefs lead us to see very different aspects within that universe. Every scientist believes in Truth. Thus, for that reason alone, all scientists are believers in some sort of God, whether they would use that word or not.

Next, and more profoundly, the scientist insists that somehow, in some way, the universe makes sense. In fact, as we mentioned back at the start of Chapter One, that assumption derives directly from the belief in a loving Creator God (though some scientists may have lost sight of that connection).

A still more explicit assumption made by the modern scientist — one we're willing to admit might be wrong, but which seems to work up to now — is something called *The Cosmological Principle*. This principle asserts that, in whatever way the universe makes sense, it makes the same sense in every part, at every time. The rules are the same in my lab as they are in yours. And the same gravity that pulls an apple from a tree, to use Newton's famous example, also holds the Moon in its orbit — and, indeed, the orbits of other planets around other stars, and stars around the centers of their galaxies. This principle is, in fact, the key to understanding Relativity.

There are several levels to this Cosmological Principle.

First, our laws of physics shouldn't care about what kind of yardstick we use. If I work out a physics problem in inches, and you use centimeters, we should still come up with equivalent answers, answers that we can convert from one set of units to another.

Second, our laws of physics shouldn't care about where our starting points are located. If I measure degrees of longitude from the geodesic marker at the foot of the Washington Monument, and you measure them from the Greenwich Observatory in London, we should still both get the same answer for how far you have to fly from San Francisco to Honolulu.

Third, our laws of physics should not depend on where or when I did my measurement. What worked in Cambridge, England, last century should work in Cambridge, Massachusetts, tomorrow.

But there's a more subtle result that comes out of this principle.

If the laws of physics are the same everywhere; if an experiment done in my lab gives the same result as the same experiment done in your lab; then there's no way we could tell whose lab we were in just by doing an experiment. If I *can* tell, if there is some experiment I can do such that, even if you put me in a box with no windows, I could tell from the results that I was in Canada and not in New Zealand, then one (or both) of those labs must be subject to some external force that's different from one to the other.

Maybe it's the force of magnetism (the Earth's magnetic field, an external force, is stronger in Canada); maybe it's the "Coriolis force" of the Earth's spin that causes water to go down the plug spinning in the opposite direction in the southern hemisphere, compared to the northern.[47] Maybe I can tell that it's summertime outside, even though it's January. But all these are, indeed, external effects. If I am truly isolated from external forces, I should not be able to tell the two places apart.

---

[47] All things being equal, the Coriolis force would indeed cause water to spin down a drain counterclockwise in the northern hemisphere, clockwise in the southern hemisphere. In effect, what's going on is not so much the water spinning around the tub, but the Earth (with bathtub attached) spinning underneath the water.

However, in real life, all things are never equal and the Coriolis force in a bathtub is so tiny as to be negligible. You need weather patterns hundreds of miles across, or winds speeds of a hundred miles an hour, before this cyclonic force becomes important.

Now, an object traveling at a constant speed and a constant direction has no net forces acting on it. That's Newton's first law. A jumbo jet cruising at a constant speed, at a constant altitude, at a constant direction, would be just such a case. If you put my windowless box inside that jet (and keep me unconscious while we're accelerating through the take off and the climb up to 30,000 feet) then there's no way I could tell, from my experiments, that I was traveling at 600 miles per hour rather than standing still — unless I had a compass or some other such device for measuring external forces.

When you are sitting in a jet airliner, a two-mile-per-hour jolt from a gust of wind may cause the stewardess to pour coffee in your lap but the coffee doesn't go flying backwards at 600 miles per hour; it falls straight down, just like on the ground. Only the external acceleration, the jolt, is evident. The 600 miles per hour of the jet causes her no problems at all.

Galileo had worked this much out. (Of course, he didn't use a jumbo jet; in those days, people talked about boats in a calm sea.) And he generalized it into a very commonsense statement of relativity, which in modern terminology can be expressed: in the absence of external forces, every law of physics should remain unchanged when you travel at a constant velocity.

It gets slightly tricky when you have two different people, traveling at two different constant velocities, measuring the same phenomenon. The guy sitting next to you in the plane sees the coffee pouring straight down into your lap. An eagle-eyed observer on a nearby mountain sees you, plane, and coffee moving past at 600 mph, along with the downward motion of the coffee. We say the two observers have two different frames of reference.

Yet the physics is the same. The Cosmological Principle says that neither frame of reference is privileged. Experiments on the plane are as good as experiments on the ground. Gravity is working on the coffee the same way, regardless of who's looking.

(If you wanted to apply this to moral principles, it would come out just the *opposite* of moral relativism. It would say that, what's right in first century Palestine, is also right today in a jumbo jet at 30,000 feet. Instead of cursing out the stewardess, give her a break; forgive and forget. She's only human, and we all have bad days.)

What would happen if our two observers wanted to compare notes? What if they wanted to calculate the rate at which the coffee lands in your lap? All they need to do is take whatever speeds the first observer measured, and subtract away the relative speed between the two frames of reference, to come up with the value of the speeds that the other guy measured. That's Galileo's Relativity.

What could be simpler? So, as you probably suspected by now, it turns out reality is not exactly this easy after all. But before we even get to Einstein's Relativity, which operates by a somewhat different rule, there's a deeper problem with the Galileo-Newton version of commonsense relativity that we need to address.

Remember our bathtub in New Zealand. Our scientist-in-the-box really wants to know where he is, so he goes out of his way to make a very careful measurement of water flowing out a laboratory-grade bathtub.[48] As it swirls clockwise, he sees with a grin what hemisphere he's in.

How come? Why, because the Earth is spinning.

Everything that travels in a circle experiences a constant change in the direction of its motion; a change in the direction of motion means a change in velocity (if not speed) since "velocity" includes direction as well as speed. A change in velocity is an acceleration. An acceleration occurs only if there's a force involved. There's an external force pushing his laboratory around.

---

[48] Actually, he'd probably build a Foucault Pendulum, which really can demonstrate the spin of the Earth… as long as you don't sneeze on it, or otherwise disturb it. A few years after Foucault, a second experiment also showed the Coriolis force on delicately balanced weights, proving the Earth spins just like Copernicus said. It was done by Fr. Hagen at the Vatican Observatory.

In this case, the force of gravity is keeping him and his bathtub glued to the surface of the Earth, instead of being flung off the spinning Earth and into space. The corner of the bathtub closer to the Earth's axis (the South Pole, in New Zealand) experiences slightly less force trying to fling it away, and the molecular forces inside the bathtub are holding it into one shape while it spins as it is tied to the Earth, whereas the water…. well, you get the idea. So what's the problem?

In the 14th century, a scientist (and bishop) named Nicholas Oresme wrote a brilliant physics text, *On the Heavens and the Earth,* which probably would have launched the scientific revolution right then and there if only that century's wars and plagues hadn't set back civilization in Europe by several hundred years. One thing he discussed at length was the question of whether or not the Earth was spinning. He showed, through a series of elegant (and correct) arguments, that logical reasoning alone could never settle the issue either way. (And Foucault's Pendulum wasn't invented until the mid 1800's.)

But then, at the end of his argument, he states (I'm translating slightly loosely here) "Of course, everyone agrees that the Earth is standing still; and so do I." His point was on the limits of logic, which could not even demonstrate something as obvious as the fact that the Earth was standing still.

We chuckle today, but without much cause, because his argument can be stood on its head. We understand that the Sun and stars rise and set because of the Earth's spin; but, according to Galileo and Newton's relativity, and Bishop Oresme's arguments hundreds of years before them, there is no inherent reason why we should insist that it's the Earth spinning and not the rest of the universe. It is only subtle forces like the Coriolis force, operating on our pendula and in our bathtubs, that "proves" the Earth is spinning.

The Earth is spinning. Relative to what?

Relative to the other planets? They're all going around the Sun, too. The Sun? It's spinning, as well… relative to something. But what? Newton tried to get out of it by referring to a universal “inertial reference frame” or “the fixed stars.”

But we know today that the stars are not fixed, either; our Sun and the other stars in our Galaxy are orbiting its center, the galaxies in our Local Group are orbiting each other, the Local Group itself is moving relative to other groups of galaxies… Where does it all end?

The whole point of the Cosmological Principle was that there should be no privileged reference frame, no one standpoint for looking at the universe that’s any better than any other frame of reference. But Newtonian physics demands just such a framework. Pushed to its extremes, the whole logic of Newtonian physics collapses under the weight of this inconsistency.

Which doesn’t stop it from working just fine most of the time, mind you. But now we’re beginning to see where to look for the boundary beyond which it no longer works.

Part 2: Surfing the Ether

The problem — and, indeed, the solution — really came with Maxwell’s equations.

The problem arose when Maxwell showed how electricity and magnetism ought to travel like a wave. That’s all fine and good, but there’s a trouble with waves. Unlike Newton’s laws, and all the other laws of physics based on them, the equations for waves are changed completely when you add in a constant velocity. Maxwell’s equations are not “invariant” when you change your frame of reference.

The guy on the beach sees the ocean going up and down as waves come into shore. A surfer riding the wave may see the shore approaching, but to her the wave is a solid wall of water that her surfboard is sliding down. And if she starts to travel at some constant speed *other* than the speed of the wave, she knows she'll wipe out; her surfing "experiment" will have a very different outcome.

The surfer is riding the wave; she's also riding the water. Someone just bobbing up and down in the water knows he's riding the water but not the wave; a slight change in his speed hardly affects him at all, whereas the same slight speed change on the surfer can be devastating.

Every wave (at least in Maxwell's experience) travels through a medium. And the speed of the wave is measured relative to the speed of the medium the wave is traveling through. (After all, our surfer doesn't care about the fact that the ocean she's riding on is part of a planet traveling through space at thousands of miles an hour; she's got enough of a challenge as it is.)

But if there's a medium that electric and magnetic waves travel through, then we should be able to detect that medium. When I'm in my windowless box in the airplane, I should be able to make some measurement of how light travels from one end of my box to the other, and from that determine both the speed and the direction of my motion… relative to the medium that's carrying the light wave.

It just so happens that Michaelson and Morley, back in 1887, tried to do exactly those sorts of measurements to determine the velocity of the Earth as it moved through the "ether" that presumably was carrying the light of the Sun and stars. And they couldn't find any such motion. From that alone, you can see that there's a real problem.

And, in fact, that's how most modern textbooks start their description of Relativity, detailing the Michaelson-Morley Experiment and then Einstein's solution, his theory of Relativity. It makes a clear, simple story… even if it didn't really happen that way. (There's no evidence that Einstein had ever heard of Michaelson or Morley at that time.) Heaven knows, in this book I've taken some pretty severe liberties with history to tell my story, so I can't complain about other books on those grounds.

But, partly out of contrariness — and the fact that there are dozens of other books out there that explain Relativity the traditional way, probably better than I could — but even more because I want to tell a slightly different lesson, I am going to stick a little bit closer to Einstein's reasoning itself.

While Einstein is universally credited with Relativity, all of its basic principles had already been proposed during the ten years before his landmark 1905 paper. In fact, it was Henri Poincaré who first explicitly stated, in 1899, the principle we discussed in Part 1, that "absolute motion is undetectable." In 1904 he went farther, and outlined what *he* called "The Principle of Relativity." In this paper he proposed the starting point of Einstein's work, the idea that no velocity could exceed the velocity of light.

His point was quite simple, if rather bizarre. His philosophical principle demanded that no experiment should be able to determine a constant velocity, or "inertial," motion. Obviously, this could only be true if Maxwell's pesky light waves kept the same speed in any frame of reference, so that if they travel at 300,000 km/s on a mountaintop, relative to the mountaintop, they also travel at 300,000 km/s inside a rocket going past the mountain at 200,000 km/s.

But is this speed measured relative to the mountain, or to the rocket?

Now this is a tricky point. Since it is predicated on a point of philosophy, let me try to illustrate it with another version of the same philosophical principle.

Say I own a fruit stand. People are always stealing my apples. I've just caught two thieves red handed, and the police are carting them away even as I speak. One is a sweet little ten-year-old girl, the girl next door; the other is a hardened criminal, a member of the notorious Beagle Boys gang from the Bad Neighborhood on the Other Side of the Tracks.

So who has committed the greater crime? Maybe you think it's worse for the little girl, her first step on the road to perdition. But then, shouldn't the judge be lenient? After all, it's just her first offense. On the other hand, you might think that one more apple should mean nothing to the hardened criminal. But if it's his third offense, stealing that apple might mean a mandatory jail sentence in some states. So what really is fair?

The physicist's answer is to note that, in both cases, I'm only out one apple; no more, no less. And so, to this sort of philosopher, both crimes are equal. The Biblical principle of "an eye for an eye," rather than being a recipe for revenge, is actually an exhortation to let the punishment fit the crime, not the criminal. You take an apple, you pay for an apple; no more, no less.

To put it yet another way, once an action has taken place its consequences are the same, regardless of the previous (or future) history of whatever it was that acted. A smashed up car is still a wreck, whether the driver was a drunk out joy-riding or a concerned grandson racing to help his aged grandma. A dropped fly ball still lets the runner score, even if the outfielder had never made an error before. Crying doesn't mop up the spilled milk.

In the same way, once a lamp has launched a beam of light, that light should travel out from the lamp in the same path, at the same speed, regardless of what zig-zags and motions the lantern went through before (or after) it emitted the light. That was Poincaré's idea.[49]

Notice, this is philosophy, not a scientific proof. If we had a bullet theory for light, then you'd expect that the motion of the lamp ought to matter. Say I'm in the rocket, traveling straight at the mountain, at 200,000 km/s, and I shine some light bullets at you sitting on that mountaintop. The photons come shooting out of my flashlight at 300,000 km/s. Shouldn't the speed of my photons, as seen from the mountaintop, equal of the sum of their speed out of my lamp, plus the speed of the rocket itself? Won't the speed of the photons, relative to your mountain, be 500,000 km/s?

No, says Poincaré.

---

[49] This kind of process, one that does not depend on the previous history of the system, comes up all the time in physics and chemistry. It's given various names in various contexts, like "conservative" or "local" or "equilibrium" or "adiabatic" or "reversible."

Poincaré believed in the wave theory of light; and Poincaré believed in the ether. Once the lantern started a wave in the ether, then the light should travel at the same speed in all directions through that ether. But because Poincaré also believed in his version of Galileo's relativity, that uniform motion should not be detectable, he also thought that the light should travel at the same speed in all directions whether you're measuring it from a moving rocket or a stationary mountaintop.

But how could that be true? It violates common sense.

Picture it this way. Think of the "ether" as if it were a big sheet of rubber. Now smack a hammer into the sheet. Waves run away in every direction from the place where you struck. The same wave would also result if your little sister smacked the sheet with a hammer while she was riding by on roller skates. In either case the waves all make nice circles, moving in every direction at the same speed away from the place where the hammer hit.

However, if your sister happens to be skating at exactly the same speed as the wave's speed in the rubber sheet, from her frame of reference the part of the wave traveling along in the same direction that she's traveling would look like it was standing still to her… wouldn't it?

But then your sister would know how fast she was moving, relative to the rubber sheet. And that's precisely what Poincaré was proposing did *not* happen, at least with light waves. Instead, he insisted, the person on roller skates would measure a light wave's propagation to be racing on ahead at exactly the same speed that a person standing still would measure.

On the one hand, the rubber sheet, the ether, the hypothetical stuff that does the waving when a light wave passes by, which Poincaré was certain must exist (he was wrong), must have its own inertial frame of reference, and light waves must travel at a certain speed relative to that medium, that frame of reference. But Poincaré insisted that no physical experiment, like measuring the speed of light, should be able to pick out that frame of reference from any other frame. So the speeds would have to be the same in every other frame of reference, too.

It was a paradox. On the face of it, it's impossible. Here we have two rock-solid principles of physics: once the light wave is started, it doesn't care about the motion of whatever started it; and given a light wave, its speed should be the same in all inertial frames of reference. These two principles make completely contradictory predictions. And yet we insist that they both must be true.

Well, either one of them is wrong; or else there's something else going on that we haven't thought of before.[50]

Of course, this paradox is all predicated on the existence of an ether, the medium that carried the waves of light. Einstein, in coming up with the photon theory of light, had an easy out. If light were bullets, then there was no need for an ether, and Poincaré's paradox could be thrown away.

But Einstein didn't take that route. Instead, he asked (in effect), what if Poincaré's paradox were true *even if there was no ether?* How would that change the way we do physics? What would the consequences really be?

Because, you see, Poincaré had come up with a solution to his paradox that seemed to work, even though no one knew why it ought to work, or why it ought to be true.

It was, to him, a purely mathematical trick. The mathematical way of stating that "you can't tell the speed of your inertial frame of reference" is to say that Newton's equations of motion should give you the same mathematical formula for motion, regardless of which frame of reference you use to measure your velocities.

For old-fashioned Galilean relativity, you take the relative velocity between the two frames of reference, which must be a constant. Wherever you measure velocity in the first frame of reference, you subtract away the

---

[50] We come across paradoxes all the time in life, in religion and in science. Rather than using them to try to disprove one side or the other, they almost always turn out to be exciting hints of something (or Someone) More Complicated than we had imagined…

value of that constant to find the velocity in the second frame of reference. And if you remember to include it everywhere you talk about velocities, then that constant term winds up canceling out or dropping out when you turn the crank and do all the mathematical operations involved in solving Newton's equations for the future behavior of whatever motion you're following.

This constant velocity factor works just fine for Newton's equations; the only trouble was, it doesn't drop out of Maxwell's equations for light waves. But a "constant velocity factor" is not the only mathematical transformation that will give you equivalent answers if you switch from frame to frame.

About 1895, a Dutch physicist named H. A. Lorentz realized that other more complicated formulae might also work. He cooked up a formula that satisfied both Newton's equations of motion *and* Maxwell's equations for light waves. It kept the speed of light the same in any frame of reference.

You can find Lorentz's formulae in any physics book, but formulae aren't really what we are interested in here. So, instead, let me just describe them briefly in words.

I measure a speed by measuring how far something travels, and how long it takes to travel that distance. But if my yardstick for how far it travels, and my clock for how long it takes, change as I change my speed, then the speeds that I measure will change as well.

What Lorentz proposed was that, as an object (like a yardstick) moves, its length as measured by a stationary yardstick actually shrinks in the direction of motion, in such a way that this contraction is impossibly small to measure at ordinary speeds but gets quite noticeable as you approach the speed of light. When it reaches the speed of light, the length of the moving yardstick has shrunk to zero.

Furthermore, he proposed, the time measured by a moving clock appears to slow down, compared to time measured by a stationary clock. At the speed of light, the clock doesn't move.

So, say I juice my rocket up until it's traveling at two-thirds the speed of light, or 200,000 km/s. I shoot a photon out of a lamp, and time it as it zips past a meter-stick.

I see that after one billionth of a second (according to my super-accurate wrist watch) it has traveled 30 cm. And, yes, 30 cm in one billionth of a second is the same as 300,000 kilometers per second: the speed of light.

But you're observing from the ground. Measuring my 100 cm stick as it zips past you at 200,000 km/s, it appears to you to be only 74.5 cm long.

When the end of the stick is even with you, I launch the photon. You wait for one second to pass, according to your clock. At that point, you see that the photon is sitting at the 13.4 cm hashmark of my ruler. And you see my clock, riding along with me and my meter stick at that point, registering 0.447 billionths of a second.

Of course, since my stick looks "compressed," to you my 13.4 cm really only measures 10 cm long. On the other hand, the whole ruler (including that hash mark) have been moving at 200,000,000 m/s for a billionth of a second, in your frame of reference, so the photon is actually 20 + 10 or 30 cm away from the point where the photon was at time zero.

So you calculate that the speed of light is 30 cm per 1 billionth of a second, or 300,000,000 km/s, in your frame of reference.

And if, as you see, the photon has traveled 13.4 cm in 0.447 billionths of a second, you realize that I probably think the speed of light in my frame is also 13.4 cm divided by 0.447 billionths of a second, or 300,000,000 m/s in my frame as well. Thus, with this shrinking yardstick and slowed-down clock, you and I both measure the exact same speed of light, relative to each of our frames of reference.

This sort of transformation makes Poincaré's relativity possible, but it doesn't prove that it really happens. What Einstein did was to take this suggestion and build a new idea of how physics would work, given this "Lorentz contraction."

Other scientists then took Einstein's ideas, systematized the mathematics, and began to work out some of the more subtle implications — looking for ways that you could actually see and measure relativistic effects.

For instance, nowadays it is possible to accelerate radioactive particles to high speed and watch them decay. You find that their decay rate slows down, just exactly as relativity would predict; from our frame of reference, the "internal clock" of the decaying particle seems to have slowed down.

A few years ago, somebody put a super-accurate "atomic clock" in the cargo hold of a jet airliner, a plane that spent a lot of its time traveling at 600 mph. Sure enough, after a few months, the clock appeared to be running noticeably slow. Nowadays they're talking about putting such clocks inside the International Space Station as it whizzes around the Earth, to see if the accuracy of the clock depends on which way the clock is facing.

But the point is this: the bizarre consequences of relativity are not just something dreamt up by mathematicians with too much time on their hands. They can actually be seen and measured in the lab.

In fact, there is one clear and beautiful demonstration of relativity that you and I take advantage of, all the time. It involves electricity.

Remember when we talked about electricity, back in Chapter 2, we found that a static electric force could be a pretty powerful force. If you take a Coulomb of charge — the amount of electricity that runs through a one Amp wire in a second — and place it one meter away from another Coulomb of charge, you'll have a force equal to the weight of a million tons.

Now consider a metal wire carrying a 1 Amp current. The current is just a line of electrons, each carrying a bit of negative charge, bouncing down a string of positively-charged metal atoms that make up the wire. In fact, there are just as many electrons in the wire as there are positively charged atoms. The net electrical charge of a wire is zero.

If I took another electron and sat it outside the wire, it would feel no force from the wire. The positive charges of the metal atoms would each be canceled by the negative charges of the electrons traveling down the wire.

But say I start to move that little electron in the same direction that the current is flowing in the wire, and (to make things simpler) at the same speed as the electrons are flowing in the wire. In the frame of reference of the electron, it thinks it is standing still; it thinks the electrons in the wire are standing still, too; and it thinks the positively charged nuclei in the wire are moving backwards.

Well, if they are moving (relative to my electron), according to the theory of relativity they must appear to be slightly contracted — packed together more tightly than they appeared to be when I was standing still. Likewise, the electrons that now appear to me to be stationary will appear to be slightly more spread out than they seemed to be when they seemed to be moving.

That means my electron thinks there are more positive charges, and fewer negative charges, in any given length of wire. (Don't worry that this seems to be creating positive charges out of nothing. The wire is eventually part of a continuous circuit, and the far end of the loop where the current is running in the opposite direction back into the battery will experience the opposite apparent effect. It'll all balance out.)

The excess positive charge should attract the electron, and pull it towards the wire. (The pull of the rest of the loop, where the current runs back into the battery, is weaker because it's farther away.) If my electron is in fact just one of a Coulomb of electrons running through a second wire, one that also has a one-Amp current, then the whole wire will feel a force pulling it towards the first wire.

Now, the it turns out that electrons in wires don't move all that fast, certainly nothing close to the speed of light, and so the effect of relativity ought to be pretty darn small. Still, the electric force is so potent that even a tiny fraction of an imbalance can make itself felt.

If you actually rig up two wires, a meter apart, and run an Amp of current through each of them, you can see them pull slightly towards each other. The force is pretty small… 2 x 10^{-7} Newtons, or about the same force as the weight of a speck of dust less than a millimeter in diameter. Of course, if you wrap the wires into a coil you can get this tiny force from each loop, and with a few thousand loops you begin to get an effect that's easier to notice.

Sound familiar? If not, go back and read Chapter 2 again. Because what we've just described is Magnetism. The pull of an electromagnet is a direct consequence of the principles Einstein described in his Theory of Relativity.

Part 3: A Pretty Taste for Paradox

The "Modern Major General" in the H. M. S. Pinafore bragged that he had a pretty taste for paradox. It would serve him in good stead if he were to study relativity. G. K. Chesterton, too, was an admirer of paradoxes. As he knew, paradoxes are wonderful ways of revealing subtle truths, by forcing you to see things you might otherwise ignore.

Relativity has given rise to all sorts of paradoxes, most of them fairly easy to resolve with a little thought. But inspiring that little thought is, of course, the joy of the paradox.

One paradox might be called the *light race:*

Recall my meter stick in the last section, moving past you at 200,000,000 meters per second. When I, riding on the meter stick, flashed a light forward, I showed that you and I both would measure the same speed of light. But when one billionth of a second passed in your frame of reference, the light was only at the 13.4 cm mark in my frame of reference, and my clock measured only 0.447 billionth of a second's passage of time in my frame.

Let's repeat the experiment, only this time I'll shine the photon backwards, away from the direction I am traveling. Once again, in my frame I see it reach the -30 cm hashmark (the minus sign means looking backwards) in a billionth of a second.

But what do you see? Well, just like the last time, in one billionth of your seconds you see the photon appear to travel 30 of your centimeters, albeit now in the other direction. Now remember, 30 of your centimeters looks like 40.25 centimeters on my meter stick. And since the stick has moved forwards by 20 of your centimeters during that time, which is 26.8 of my centimeters, after one of your seconds you see the light passing the -67.1 cm hash mark of my meter stick.

And the clock there tells me that when the flash of light gets there, 2.24 billionths of my seconds have elapsed.

But wait a minute. How can that be? Is the passage of one second in your frame of reference the same as the passage of 0.447 seconds, or 2.24 seconds, in my frame of reference? Which is it?

To put it another way... in my moving space ship, I sit exactly in the middle of the room and set off a flash of light, lots and lots of photons, which travel — at the same speed — out in all directions. The walls before me and behind me are equally far away from me. So both walls should experience a flash of light at the same time.

Now, remember, my spaceship is traveling forward at two thirds of the speed of light. If you are sitting on the mountaintop watching my ship go by, you will also see the light flashing and traveling out, in the same speed in all directions, relative to your frame of reference.

But in your frame, the front wall of my room is running away from the spot where the light was emitted, and the rear wall is rushing towards that spot. The rear wall should run into the light wave first; only later will the light catch up to the front wall.

So which is true? Do the walls get lit up at the same time, or does the back wall get lit first? Relativity's answer may leave you very uncomfortable. But its answer is: *both are true*. It depends on your frame of reference. I, in my spaceship, will see the walls light up together; you on the mountaintop will see the back wall lit first.

Does this mean that realism is dead? That two contradictory things can be true at the same time? No… what is dead, instead, is the idea of "the same time." Time works funny in a relativity universe, and one of the funny ways it works is that our sense of "simultaneous events," two things happening at the same time, is not as common-sensible as we thought.

But, at first blush, this just makes things worse. It makes us think in a whole new way about causality.

The whole drive of the scientific, Western view of the universe is the principle that things don't "just happen;" they happen for a reason. Now we already saw that, in the case of polarized photons and decaying atoms, this principle has been badly shaken. Some things appear to happen *without* a cause. But when you lose the idea of time being the same in all frames of reference, could you ever get yourself into a situation where there is cause and effect, but the effect occurs *before* the cause?

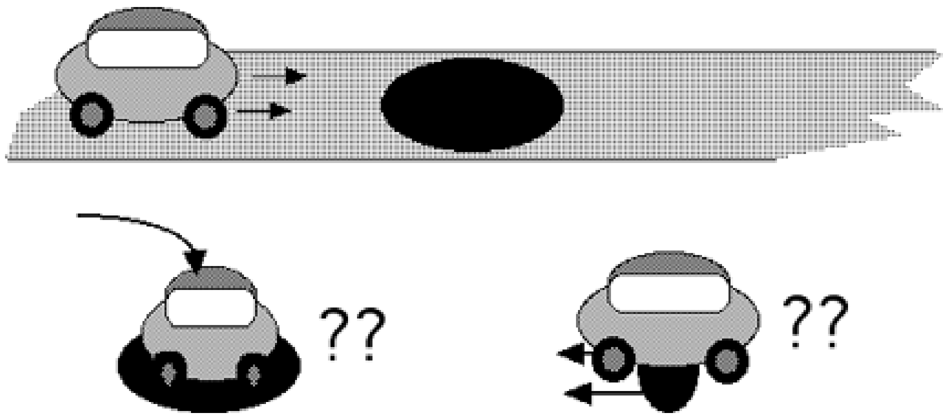
Common sense says no, and in this case at least there's no reason to doubt common sense. But this now means that causes have to occur in a time *and place* that will be earlier than the time of the effect, as viewed in *every* possible frame of reference.

Here's another one, the *pothole paradox*:

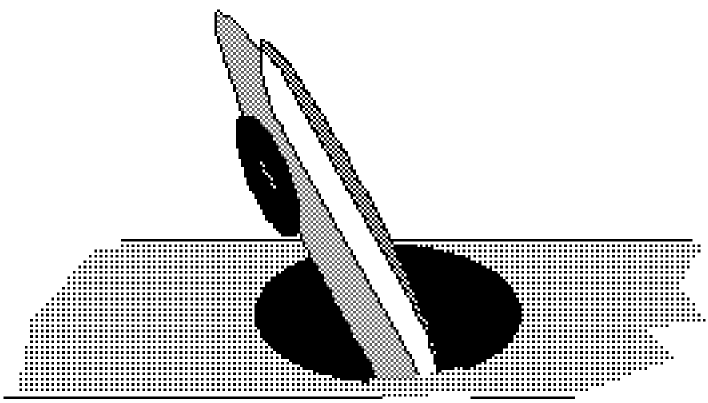
Say I have a car that is ten feet long, and I am approaching a pothole that is nine feet long. A physicist standing next to the pothole is warning me, don't go too fast! If my car approaches near the speed of light, it will appear to shrink relative to the frame of reference of the ground; if it Lorentz-contracts to less than nine feet long, it'll fall in the hole and never get out.

But from my frame of reference, the hole (and the ground) is rushing past my car, while I think that I am standing still. So, in my frame of reference, it's the hole that's getting smaller, not the car. The faster I go, the smaller the hole will appear. I should have nothing to worry about.

Who's right? In this case, you can't say it merely depends on your frame of reference. The car won't glide over the hole in one frame of reference, and fall into it in the other. No way… either it falls, or it doesn't.



Turns out, the way to squirm out of this paradox is to invoke the problem of "simultaneity" again. From the point of view of my physicist friend, the car can plummet, front and back simultaneously, into the hole; from my point of view, the front falls first, then the back, such that the car fits down the hole by being twisted.



Finally, there's the famous *twin paradox*:

One twin takes off for Alpha Centauri in his speed-of-light rocket ship, and 10 years later (by Earth reckoning) he returns. Having spent most of his time traveling close to the speed of light, he appears to have aged only a few months, while his twin sister is ten years older.

But from the traveler's point of view, doesn't it appear that he's standing still, while the Earth has receded at near the speed of light, then turned around and come back to his space ship? By his frame of reference, shouldn't *he* be ten years older, while the stay-at-home sister hasn't aged a whit? How do we get out of this one?

The answer is that there is one visible difference between the space ship and Earth: the space ship *knows* it has accelerated to near the speed of light, then slowed down, stopped, turned around, speeded up, and come back to Earth. So obviously these *accelerations* play a funny role in Relativity.

It's a role that isn't obvious from the simple theory Einstein worked out in 1905. Worrying about accelerations kept Einstein busy for the next ten years. And what he came up with was something called General Relativity. (Relativity without accelerations is now called Special Relativity.)

Remember, Einstein's goal was not just to come up with "equations that worked" but rather to rethink all of physics, in a way consistent with the equations that Lorentz and the others had already come up with.

For instance, he realized that tacking on the Lorentz correction factor to the velocity of a moving object did more than just get around the apparent paradox of having the same speed of light in all frames of reference. This substitution involves the ratio of an object's speed, squared, to the speed of light squared. But this means that in the equation for energy — mass times the object's speed, squared — you wind up with a term of mass times the speed of light squared that hangs around even when the velocity of an object drops to zero.

Einstein then interpreted this mathematical curiosity with a physical fact: every object that has mass must have a certain energy content even when it's at rest. That energy content is $E = mc^2$.

Remember, radioactivity was being explored at the same time that Einstein's equation was announced. Soon it was discovered that the products of radioactive decay had, in sum, less mass than the atom that decayed in the first place.

Mass was not being conserved. Instead, through radioactive decay of materials, mass was being converted into pure energy. How much energy? Take the mass loss, multiply by the speed of light squared, and you wound up with exactly the amount of energy that was seen coming out of these decays. Once more, Einstein's theory worked.

Could this energy be harnessed and controlled? The era of atomic energy was born…

And that was just one example of turning equations into physics. Rather than being satisfied with clever ways of doing math to get around logical paradoxes, Einstein insisted in understanding of how those clever math tricks worked in the physical world. Recall, to a physicist "understanding" doesn't mean a deep philosophical exposition, but rather getting to see the same phenomenon in all sorts of physical settings, as a way of getting used to it.

Einstein himself was not much of a mathematician… at least, not compared to some of the top mathematicians around during his time. One of them, the German genius Hermann Minkowski, expressed Einstein's theory of relativity in an elegant (if rather obscure) mathematical form of "operators" and "contravariant vectors." He took the three dimensions of space and added time, treated as a fourth dimension.

It all sounded like making a mountain of a molehill to Einstein; the math was truly obscure at first.

But remember the paradox of the car falling into the pothole. The orientation of the car, in three dimensions, changed as one changed the way one calculated the time in which the front and back fell. In other words, the shape of space can appear twisted in one frame, untwisted in another; and the time that one end experiences can be different from the time the other end experiences, as viewed from differing frames of reference. This confusion of times and spaces was exactly what Minkowski was expressing in his equations.

No longer can time and space be measured independently of one another. They have to always be considered together: *space-time*.

Those equations, Einstein realized, were “obscure” only because he wasn’t used to them. So he set about getting used to them, playing with them, asking how they could be generalized to the case where frames of reference were not merely moving at different speeds, but also speeding up or slowing down, turning about or turning around.

According to Newton’s 17th century physics, these changes in velocity, these accelerations, had to be caused by some external force. But, pondering this problem as a professor in Berlin (while World War I raged on) Einstein thought of another way of looking at this. The car falling down the hole, being pulled from a straight-ahead motion to a downward motion, must experience an acceleration just as much as the twin coming home from his light-speed rocket trip. And as a result, its space-time looks twisted as one moves from frame to frame. Maybe, realized Einstein, anything that we see as an acceleration is precisely the result of such a twisting of the space-time continuum.

Imagine an ant walking on an apple.[51] The bug walks in a straight line, each foot placed before the other (and the other and the other…). At least, the bug thinks it is walking in a straight line. Next to it, a second ant is walking in a straight line, too. If these ants took high school geometry, they would probably think that because they’re walking in straight, parallel paths, their paths should never cross.

But you and I, looking at the curved shape of the apple, can see that as they walk up toward the stem their paths will get closer and closer together. The curve of the apple looks to the ants like a force pulling them together.

In fact, an even more interesting thing happens at the stem. There’s a dimple in the apple there. As it happens, one ant walks straight into the

---

[51] I’m swiping this image from the classic text, *Gravitation,* by Charles Misner, Kip Thorne, and John Archibald Wheeler.

stem; the other ant's path carries it down into the dimple but misses the stem itself.

Now remember, each ant thinks it's walking in a straight path. But if you peeled the apple, pounded the skin flat, and looked at the footprints the ants left behind, it would look as if the two ants were walking closer and closer together, with one falling straight into the stem and the other doing a funny dip towards the stem, before starting to dip back as it left the stem behind.

Add to this a slight variation on Poincaré's idea, that the laws of physics shouldn't depend on past (or future) events. Another way of saying this is that only things immediately touching me can cause me to react.

You can push the first domino in a string, but the last one won't fall until the one next to it has been knocked over. The light can turn red, but until its photons touch my eye I won't know it's time to stop. The peace treaty between the US and Britain can be signed in December of 1814, but until Andrew Jackson has the piece of paper in his hand telling him about it, he'll think the war is still on and in January of 1815 he'll be fighting the Battle of New Orleans.

Cause and effect depends on the cause communicating itself to the thing being affected. Once the cause arrives, it's a local phenomenon. There's no such thing as "action at a distance." Like politics, all physics is local.

Now, finally, mix in Einstein's "Principle of Equivalence." Newton had said that the force of gravity accelerates me toward a massive object, like the Earth. But a fellow in a car stepping on the gas pedal is also being accelerated. And (if you're blindfolded, anyway) there's no way that you can tell the difference between the acceleration of gravity or the acceleration due to the motion of a car.[52]

---

[52] That's the trick they use in amusement parks, like the "Star Wars" ride at Disney World — by tipping your seat back and forth, they can use the acceleration of gravity to make it feel like you're in an accelerating rocket.

The ant tries to walk in a straight line, but notices that its path is getting bent towards the apple stem. What would be more reasonable than to assume that the apple stem is putting out some "force" drawing it closer? But all physics is local. There's no way that the ant can tell the difference between a magic force that seems to be emanating from the stem, way over there, or a change in the shape of the apple skin, right below its feet.

Indeed, we know that the true nature of the "force" put out by the apple stem is that the space that the ant lives on, the surface of the apple, is bent in a funny way around the stem of the apple. In the same way, Einstein realized, the force of gravity could simply be a bending of the shape of space.

Is there any way you could prove this? It turns out, for reasons we needn't go into here, that there are subtle differences between the way things ought to behave in a bent space versus the way they'd behave in the old Newtonian version of gravity.

One of the most clear cut is the way light behaves. If light has no mass — as we'd expect — then it should not be attracted by gravity. Light should always travel in straight lines. However, if in spite of having no mass it did fall due to the effects of gravity, it should be moved in a very specific and predictable way so that the light from a distant star, passing close by the Sun, should be bent away from a straight line by a tiny but predictable, and measurable, amount.[53]

But if light travels in a "straight line" through a curved space, it turns out that the amount of bending that the curved space theory predicts would be *half* the amount that the gravity theory suggested.

During a solar eclipse, you can actually see the faint light of a star passing close to the Sun, while the Sun's glare is blocked off by the

---

[53] The rate at which something falls does not depend on its own mass. Remember Galileo's famous experiment dropping two weights from the Tower of Pisa? The big one and the small one hit at the same time. Thus, it seems that you should be able to extrapolate to the case of a falling light-bullet, even if the light-bullet's mass goes to zero.

Moon. Thus, during an eclipse in 1919 (soon after Einstein published his theory of General Relativity), international expeditions were held to observe and measure the position of stars near the Sun.

The light indeed was bent, the apparent position of the stars ever so slightly shifted, but only half as much as the falling light-bullet theory suggested — and just as Einstein's theory of bent space predicted.[54] The space near the Sun was warped by the mass of the Sun. In general, one concludes that indeed what we call the "mass" of an object is precisely a measure of how much it can warp space-time.

Part 4: Finite and Infinite Both

Why doesn't the Moon fall out of the sky?

Why don't satellites fall down? A common, naive view, one which dates in some form all the way back to Aristotle, thought that gravity somehow "turns off" once you get outside the atmosphere of the Earth. But Newton's law of gravity, confirmed by every mission into space, says that isn't true. Earth's gravity does get weaker as you move away from Earth, but it never falls to zero.

And in fact, every body with mass in the universe, no matter how far away, attracts every other body in the universe — strongly or weakly, depending on its mass and how close it is. Or, in the General Relativity view, every body with mass warps space in such a way that every other body's path is bent, in a way that tends to bring those bodies together.

Since every planet in our solar system has a gravitational pull on every other planet, and all of them are attracted by the gravitational pull of the Sun, why don't they all fall together, and fall into the Sun? If every star in our galaxy has a gravitational force comparable to our Sun's, what

[54] Einstein, a German, had come up with the theory; Eddington, an Englishman, performed the observation. The collaboration of these two, within months of the end of World War One and amidst the intense bitterness between their two nations, was an impressive display of the best in the human spirit.

holds them apart and allows the galaxy to maintain its shape? And if every body in the universe has an attractive force of gravity then why don't all the stars in the universe fall into one another?

Good questions. Let's start with the first one. What holds satellites up in their orbits? It's the motion of those satellites, moving at a great rate of speed (something like 18,000 miles per hour). A satellite whipping around the Earth once every 90 minutes or so has a centrifugal force that exactly counters the force of gravity.

The Moon, farther out, can go slower (one orbit a month) since the gravity it feels is correspondingly weaker. Exactly the same principle works for the planets around the Sun; at their distances from the Sun, an orbit once every few years suffices to produce enough centrifugal force to counter the distant Sun's gravity.

The stars in our galaxy likewise orbit about the center of the galaxy, roughly once every two hundred million years. Indeed, there are clusters of galaxies out in the universe, and they too seem to be slowly orbiting around a common center, at a proportionately much slower period.

Does this go on forever? Do all the clusters orbit each other in superclusters? Do all the superclusters orbit about some one unique point, a place that we could name the Center of the Universe?

Apparently not. There are two strong arguments against this final view.

The philosophical argument, which dates from long before Einstein's time, is that such a unique and special point, the one place in the universe that everything else was spinning around, would violate the "cosmological principle" expounded in the previous chapter — the idea that no one spot, anywhere, should be somehow special or unique.

The second argument comes from our observations, and is much more recent: no such "super" rotation of galaxy clusters or superclusters about some magic "center of the universe" has ever been seen. Rather, the motions of clusters of galaxies follow a very different pattern.

How can we measure the motions of distant galaxies? Turns out, it's surprisingly easy. Most stars have trace amounts of elements like iron or sodium that emit light in certain precise wavelengths. When these characteristic "spectral lines" are emitted from a star moving towards us, the frequency of the light wave is increased, like frequency of a speedboat hitting wave crests is increases when it is driven straight into the waves.

Shifting the frequency of light means shifting its color; all the emitted light appears slightly "bluer" to our telescopes. Likewise, a receding star's light spectrum is shifted to the red. This "blue shift" and "red shift" can be measured precisely, and translated directly into a speed of the star or galaxy relative to us. It's the same principle, called a "Doppler shift," that causes a truck horn or ambulance siren to be high-pitched when the vehicle is coming at us, and lower pitched as it moves away. (It's also how a cop's radar gun measures the speed of passing cars on the highway.) Observing these kind of color shifts got underway soon after the end of World War I, especially in observations at Mt. Wilson, California, by Edward Hubble.

At the same time, Einstein's General Relativity theory had shown that gravity could be thought of as the warp of space, which still raised the question of why the universe didn't collapse on itself. Einstein could only suggest a fudge factor, some "cosmological term" that held things stable against gravity.

Willem de Sitter, about 1917, pointed out that there was no need to assume the cosmological term and the mass of the universe were exactly in balance; if there were a slightly larger cosmological term (or a little less mass) then the universe could actually be flying apart.

But then in the early 1920's Alexander Friedmann in Russia argued that if you had the expansion as a starting condition, you wouldn't even need the cosmological term! His work, unfortunately, was ignored in the West at that time. His basic idea was published in an important German physics journal, but its full exposition only appeared in a book published in Russian.

The same idea, however, occurred to the Belgian cosmologist (and Catholic priest) George Lemaître in 1927. He had the additional benefit of being in close touch with Hubble and the astronomers in California. Within two years, Hubble's observations of galactic clusters had confirmed his idea of an expanding universe.

Critics mocked this idea by calling it a theory of a "Big Bang." The name stuck. So did the theory. By 1931, Einstein himself had dropped the idea of a cosmological term, calling it "the biggest blunder of my life."

The Big Bang idea is simple: the reason that the universe isn't falling into itself is because it started out as a giant explosion, with every "piece" flying outwards away from every other piece. In the case of the universe, each "piece" is a cluster of galaxies, containing billions of stars. The mutual gravity of these pieces still slows them down, and — if there's enough mass, enough gravity — it could even eventually stop this expansion and cause everything ultimately to collapse again. But that hasn't happened yet. And it's not at all certain that it ever will happen.

The implications of this theory are both obvious and subtle. The most chilling, to the "cosmological principle" (which could be thought of as saying that "things are boring, everywhere and always") is that it means the universe has a beginning: a point in time that was most definitely *not* boring. For this reason alone, the critics of the theory worked hard to come up with alternate explanations. But by the mid 1960's, the observational evidence was hard to refute.

First of all, as Hubble first indicated and subsequent, improved observations have confirmed, every galactic cluster is indeed red-shifted — each one is moving away from us. Still, there were other possible ways the universe might be put together that could account for this motion. However, theorists in the 1940's trying to figure out what the early universe must have been like, when the explosion was just starting, predicted that if there was a Big Bang, all of space today should be filled with the "echo" of the Bang in the form of radio waves of about a centimeter in wavelength. And in 1965, that radiation was actually discovered.

There really was a Big Bang. It looks like the universe did have a beginning.

And there are even more surprising implications to this theory.

For instance, you might ask why every galaxy cluster is moving away from *us* — are we, or at least our galaxy cluster, the center of the universe? Or is there some other preferred center?

No. If you think about it, you'll see that every fragment of an explosion appears, from its frame of reference, to be the center of that explosion.

Imagine a tree stump about to be blasted out of the ground. Three ants happen to climb up on the stump just as the dynamite goes off; three bits of bark, each with an ant, go flying away from the explosion. (We'll have them all flying in the same direction, just to make things easy).

One bit is moving at 50 feet per second, the second at 100 feet per second, the third at 150 feet per second. Once the explosion has given these fragments their initial velocities, nothing else happens to speed them up or slow them down.

After a second, the first ant will have traveled 50 feet; after two seconds, it'll be 100 feet away from the original site of the stump. Likewise, two seconds after the explosion you'll find the other two ants located 200 and 300 feet from the original tree stump. The more distant fragments get to be more distant by traveling at proportionately higher speeds.

But the ant standing on our middle fragment thinks it is standing still (relativity again!) and looking back, the first fragment seems to be lagging behind; indeed, from the ant's point of view, it's going backwards at 50 feet per second. The other fragment, from the ant's viewpoint, is moving forward at 50 feet per second. More distant fragments appear to be moving away at ever greater speeds. The fastest fragments travel the farthest; the slowest fragments (from the point of view of the original tree stump) lag the most behind our ant.

That's exactly what Hubble saw in his galaxies. The farther a given galaxy cluster was from our cluster, the faster it appeared to be moving away from us. Their speeds were proportional to their distances.

That means you can use the constant of proportionality, now called the "Hubble Constant," to actually calculate the beginning of the universe — the time when, if you run the film backwards, all the galaxies had to be located at the same spot.

This time is still imprecisely known, given the enormous problems of finding a really accurate way to measure the distances to those far-off galactic clusters[55]; the precise number also depends on how much you think the galaxies have been slowed down, since the Big Bang, by their mutual gravitational attraction. (And further complicated by the apparent presence of other unseen masses and forces that we're just beginning to understand.) But best estimates put the age of the universe at somewhere between ten and fifteen billion years.

More surprisingly, you can also use this proportionality to work out the *size* of the universe.

Recall our rule for the speed of galactic clusters: the farther, the faster. But there is an ultimate speed beyond which no galaxy can travel: the speed of light.

Define the distance that light travels in one year (about six trillion miles) to be one "light-year;" a cluster of galaxies moving at the speed of light for 10 billion years will travel a distance of ten billion light-years (sixty billion trillion — 60,000,000,000,000,000,000,000 — miles, if that helps). So if the universe really is ten billion years old, it *can't* be any bigger than that.

Does this mean the Universe is finite? And if so, what's "outside" the Universe? And what's at the center, the place where the Big Bang

---

[55] How do we measure the distance to far off galactic clusters? There are lots of different ways — check out a good astronomy textbook for more details — but the trouble is that the answers they give can vary by 50% or more from technique to technique, and nobody is really sure which answers are closest to being correct.

happened, today? These seem like very sensible questions. But in fact, none of them make sense.

There is no such thing as a center, or an "inside" and "outside" to the universe. The Big Bang does not represent the spreading out of material into an otherwise empty universe. Rather, what is going on is that *space itself is expanding*… between the galactic clusters, at any rate. At the "beginning" when all matter was concentrated into a point, that point *was* the entire universe and nothing existed, not even nothingness, except for it.

The mass of the universe warps space. And the mass of the universe warped into a point, warps space into that point as well. That sounds like a lot of nonsense words, so let's approach it from another angle. Let's do a thought experiment.

Imagine we could see to the edge of the universe, where material appears to be receding from us at the speed of light. Our own cluster of galaxies (known, unimaginatively, as "The Local Group") is roughly ten million light years across. But as we look at a similar cluster receding from us at 0.999999999999 times the speed of light, it will be "contracted" such that it appears, to our eyes, to be only ten light years thick.

But what then happens if we could magically transport ourselves into that "ten light year thin" cluster of galaxies? Once we start traveling at its speed, we see it "uncontract" out to a cluster not much different from our own. And, looking back at our own galaxies, 10 billion light years away, we see that now it is our own "Local Group" that looks ten light years thin.

And looking in the other direction, away from our galaxy into the direction we always thought of as marking the edge of the universe, we now find that we can look another ten billion light years farther down the road! We will now be able to see, as galaxies, clusters that would have looked paper-thin back on Earth.

And there is nothing in the physics, at this level, to prevent us from thinking that we could not continue progressing in this direction indefinitely.

So the Universe could be a finite ten billion light years across, and yet still continue on to infinity. Finite and infinite, at the same time.

But there's another possibility.

Recall that the mass of the universe warps space. It is possible, therefore, that this infinite progress into "paper-thin" galaxy clusters may not lead us deeper and deeper into uncharted worlds. If there is enough mass in the universe to warp its space sufficiently, then we could instead be finding ourselves traveling in a circle; that after a sufficient number of "jumps" to the "edge" we find ourselves right back in our own galaxy cluster again (like the ant walking around the apple eventually crosses its original path).

Notice, in such a universe, we could keep traveling in that "direction" forever, but we would encounter only a finite number of galaxy clusters. This second kind of universe is called a *closed* universe. The first kind is *open*.

One of the other differences between open and closed universes is their ultimate fate. A closed universe has enough mass in it that eventually the pull of its gravity will overcome the initial velocity of the "Big Bang" and cause the galaxies to collapse… or, as you might intuitively see, since motion in the same direction eventually gets you back to your starting point, the warp of space means that the motion of all the galaxy clusters eventually will bring them together again. Imagine a dozen ants leaving the stem of the apple, all walking down a different path. At first the ants seem to be walking away from each other, but eventually they will all meet together again at the bottom of the apple.

So, is our universe open or closed? Fact is, we don't know. Within the data we have so far, on how much mass there is in the Universe and how the clusters of galaxies seem to be receding from each other, either possibility could be true (though, at the moment, the best data appears to support the open model). One hint: the simplest model for the Universe suggests that it has just enough mass to be exactly at the boundary between open and closed. Such a universe would expand forever, but at an ever slower rate, almost but not quite coming to a halt.

Let's just pause a minute to catch our breath here, and think about what we've just said. With a handful of observations — difficult, and tedious at times to take, but nothing particularly extraordinary — and simple logic based on common sense and a few not-unreasonable assumptions, we have painted a broad-brush picture of the beginning and the end of the universe.

It's not just a dream-picture; it's something we have every reason in the world to believe might be true. We have seen, with telescopes and with our mind's eye, to the farthest reaches of the space. We have been able to engage in reasonable speculation as to the possible situation at the end of time.

It's possible we've got it all wrong, of course. After all, haven't old theories and ancient cosmologies been overturned in scientific revolutions before? So why should we take this theory seriously?

As the so-called "creationists" tell us, and rightly so, there's no way science can prove that the universe wasn't created 6000 years ago, with all its elements, the isotopes in the rocks beneath our feet and the stars over our heads, set up so as to imitate a much older universe. Maybe our prime assumption of the "cosmological principle" is wrong, and the laws of physics were different at different times or in different locations.

But notice what we would have to give up, to maintain the static unevolving universe that the fundamentalists want. To have an unchanging universe, one would need to postulate a God who at different times and places changes the rules. By contrast, the scientist's universe — the universe of evolving stars (and evolving life) — is what one deduces when one assumes that the laws of the universe do *not* change. Which picture is closer to the God whose "faithful love is everlasting, his constancy from age to age" (as we hear in *Psalm 100*)?

And before we dismiss any theory as “merely a theory,” something too strange to be true, we should be aware of what the scientific “revolutions” really accomplished. None of the “revolutions” cited by the philosophers of science[56] ever really disproved what had gone before. True observations remain true, even if new and improved theories are used to interpret them, or show the limits to what was thought before.

There is nothing false about the observations of Newton, or of Aristotle for that matter. Their physical theories, within their realms of applicability, remain as true as ever. Each succeeding “revolution” did not tear down the previous edifice but rather built upon, expanded, and saw in a new light what had gone before. Meanwhile, the principle of observing nature and the concept of general “laws” of nature which Aristotle used more than 2300 years ago are still the cornerstones of quantum physics today.

So don’t expect new developments of astronomy or physics to completely repudiate this picture. If we can learn anything from the history of science, it’s that our old ideas always turn out to be mere oversimplifications of an ever-stranger-looking universe.

And yet, it is still comprehensible. The more we understand, the more we realize we have yet to understand; but still, the more we understand, the more we do understand.

For we have followed our science now back on a trail of light to the foundations, where God laid the cornerstone of Everything. Not bad, for a small bit of protoplasm on an insignificant speck of a planet, circling a run-of-the-mill star in a very ordinary part of a not particularly unusual galaxy.

To really appreciate it, go outside at night sometime and look up at the stars. That’s all, just look.

---

[56] Nothing is more maddening to a scientist than listening to a philosopher of science. Most of them approach science with very strange ideas of how we work, and then blame us when they discover these ideas are wrong. Talking to a philosopher about science can be as frustrating as talking to a nonbeliever about religion. More often than not, they just don’t get it.

Those stars? That's what we're talking about. It's not some mystery trapped in a lab far from where we live. It is a universe that any of us can see, and marvel at, on any clear night. We need use nothing more than the equipment God gave us at birth: our eyes… and our imagination.

There's a blessing that comes from all this. If the concepts and distances, the wonders and the paradoxes all seem too much to comprehend completely, recognize that nonetheless you have comprehended them at least to some degree.

The rocks on the hills know nothing of space-time. Your cat is innocent of all knowledge of the quantum. Sometimes, nowadays, we're embarrassed to admit it, but we human beings really are something special. We have intellects. The fact that we did nothing to earn them only serves to emphasize that they are gifts; and so we have a responsibility to the Giver for their use, and for the love and care of the universe in which we use them.

There's one final lesson to be drawn. An atheist might look at the scale of the Universe and our own relative insignificance, and ask how any conceivable "god" could possible care about us. But we believers know from our own experience that there is indeed a personal God who does exactly that. Throughout salvation history, recorded in our scriptures, He has paid special attention to us, His creation.

There could even be other beings, cousins of ours in other parts of our universe, who also fall under His care. No matter how alien they might seem to us, we would both be children of the same universe, and children of the same Father.

The immensity of the universe tells us (and them, should they exist) a very different lesson than the atheist takes. The finite infinity of space and time lets us glimpse, ever so tentatively, the faintest shadow of what it really means when we talk about the unbounded infinity of God.

Afterword

"I am the Light of the World"

It's commonplace to talk about the endless fight between Science and Religion. It's a cliché to equate all religion with the strictest form of Creationism, one that says the world was created in exactly Seven Days and Genesis is our only science text; or to think all scientists treat their science as their religion, and preach that there is no God but Physics. The whole theme of this book, you probably suspect, has been to insist that both viewpoints are wrong.

But, having developed our taste for paradox, it might be more fruitful to try the opposite point of view. There's a reason why some good, sincere, and very smart people can espouse Creationism; while others, equally good, sincere, and smart, have abandoned religion for a materialistic version of science. The reason may be, both views are at least partly right.

In order for atheists to say they do not believe in God, they must have a pretty clear picture of the God that is being rejected. And the god that the atheists reject may well be a god worth rejecting: one who is indeed far from the God we believers embrace. A simple or arbitrary god who creates by whim or at random is inconsistent with the complex but rule-bound nature of the universe. Science rejects a god of chaos, one without laws, who makes no sense. But then, so does Christianity.

Science rejects a god who mutters "Let there be…" at random. But though God of Genesis does indeed create (literally!) by *fiat,* His rhyme and reason are also there. The story of Genesis tells us that creation was formed in stages, step by step, with the most subtle hints of an ultimate plan. So the Creationists are right, too, when they insist that discarding the Genesis story of creation and fall would mean throwing away the only clue we'll ever get of the why and wherefore of this universe.

And, as it turns out, most scientists are not strictly speaking atheists. The proportion you'd find in church on Sunday (or in a synagogue on Saturday, or a mosque on Friday) is not all that different from the general public at large. Even those scientists who don't belong to an organized religion still are, most often, theists — or at least agnostics, suspecting the existence of a God but never expecting to know Him. Only a few claim to be atheists; and even they still worship at the altar of Truth.

(And likewise, of course, very few believers are Fundamentalists. Furthermore, even Creationists happily live with technology and a world-view that's far removed from ancient Palestine's.)

The biggest issue to those scientists who are agnostics is the question of a personal God who acts in our daily lives. But even the most unreligious of scientists looks in nature for a key, a rhythm, a sense of a familiar characteristic pattern, one that has succeeded in the past and one that can give a clue to future research, opportunities for further understanding.

For lack of a better phrase, nature has a *personality*. And a successful scientist is one who is familiar enough with that personality to recognize when a theory gets it "right" or "wrong." Just as we know our favorite characters from a novel or a TV show, and will react badly if a new writer tries to take them over and gets them "wrong," so a badly worked out theory will set a good scientist's teeth on edge.

Recall how the God of *Genesis* remarks on creation, judging it good. Even the most atheistic scientists echo that sense of joy, that simple happiness, that sense of rightness, when they uncover the elegance in nature reflected in its laws of science.

Those religious believers whose mistrust has kept them distant from science may never learn the personality of nature that the scientist has come to know. On the other hand, they may well know the Person whose personality the scientists sense.

Taking *Genesis* as a science book may be poor theology, but it does not prevent one from being a good, loving child of God. (Nor does it guarantee it.) Science is not necessary for salvation. But if one loves the

Creator, it does seem that engaging one's self, mind and soul, with His creation is a logical response. A mother proudly displays the creations of her toddler child; how much more should that mother glory in the creation of her God?

There is a tension we must maintain, however.

Both science and religion are concerned with creation, with the nature of reality and the origin of things, and both are involved with issues of truth. To hold them separate, in watertight boxes, is a sterile solution that smacks of dishonesty. And yet, in a fundamental way, science and religion are very different.

Science consists of human-made theories. Because they are human-made, we can fully test and understand and know them. But because they are human-made, they will always be limited and inadequate. Nonetheless, they can lead us toward an ineffable truth beyond science's complete understanding. Thus Science starts with Understanding, to approach Truth.

Religion, by contrast, starts with that Truth, complete and beyond question. When God speaks, it is indeed God and not some pale substitute. But His truth is passed to us through the medium of human beings: the authors of Scripture, the teachers of tradition.

Jesus was also — by His choice — human; even if we had been there listening to Him ourselves, our concept of what He was saying would still be limited by the human language He used, and also *our* human limitations, our own frail human understanding. And day by day we must rely upon our all-too-fragile grasp of our own personal religious experience. Thus this Truth is at best only poorly understood. Religion starts with Truth, but only begins to approach Understanding.

That's the human experience. We spend our lives on the road linking Truth and Understanding. Scientists travel in one direction, believers the other; those of us who are both, get to experience both. It may have two directions, but it's the same road either way.

It is the Way to the Dwelling of Light.

Acknowledgments

It'd be impossible to list all people who have helped me write this book. The trouble is, the people I most need to acknowledge are the teachers I heard years ago: people like the young grad student who tutored me in electricity and magnetism when I was an undergrad at MIT; the visiting professor from Yale who gave a guest lecture about G. Willard Gibbs; the older physics researcher who was humble enough to finish out his career at MIT by teaching a homework-review session for my "vibrations and waves" class; and countless others. I remember what they taught me, and I've gleefully swiped their best stories for this book; but, sorry to say, I no longer remember their names.[57]

The fact is, there is nothing original in this book at all. No new physics, and no new theology either. If there is, then I've goofed; all I tried to do is tailor an old tale for a new audience.

But still, even for that I needed help, to insure that the tale did not get changed in the re-telling and to be sure that the words I used could be well understood by my audience.

Bob McTeigue SJ was my first reader and most stringent critic. The philosophical howlers that remain in this book are in all the places where I didn't take his advice. Paul Nienaber SJ, a particle physicist, passed judgment on my physics; other scientists may find those places where I ignored his comments, too.

My parents, Patricia and Joseph Consolmagno, served as my models of the intelligent lay persons for whom this book was written. Being themselves teachers and writers, they helped repair the worst of my

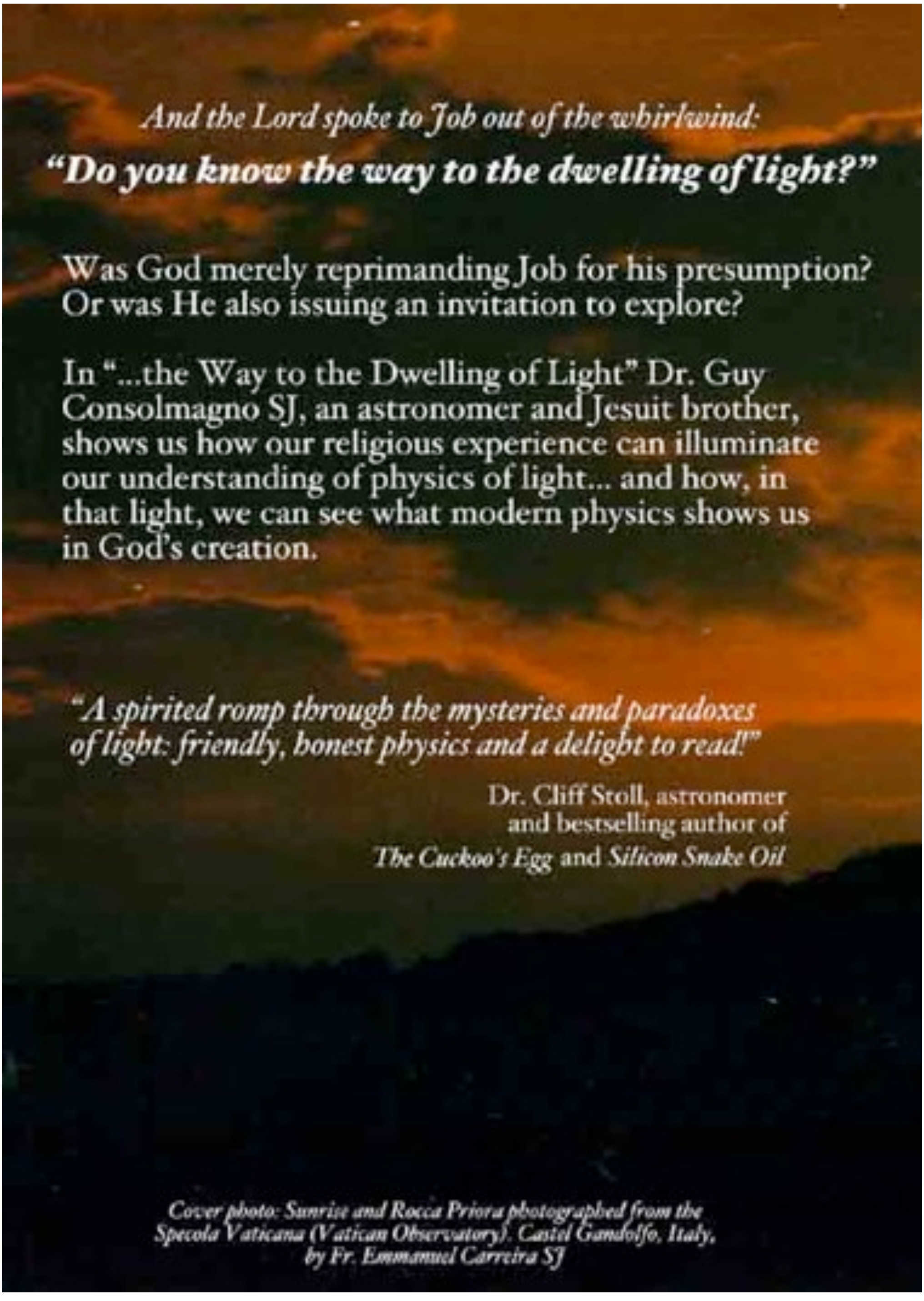
[57] Except for that humble physics teacher, Clifford Shull. He made the papers a couple of years ago. Retired for good and nearly 80, he was woken up at 5:30 a.m. one morning by a phone call from Sweden. Seems that a panel of distinguished physicists had decided his work on neutron scattering in the 1950's deserved a share of the 1994 Nobel Prize in Physics.

grammatical and syntactic lapses. (If you find that nonetheless I have abused the language, all I can say is that it was even worse before they took their red pencils to it!)

The members of my community at the Vatican Observatory all had a hand in the writing of this book, especially Fr. Chris Corbally SJ, Fr. George Coyne SJ, and Fr. Sabino Maffeo SJ. Chris, in addition, passed copies of it on to fellow scholars of the annual Star Island conferences on science and theology. Two of them sent me lengthy and extremely useful, but anonymous, commentaries. I thank you profusely for your time, effort, and advice.

Andy Whitman SJ, a mathematician at the Vatican Observatory, and Patrick Gavin who teaches science at the Jesuit school at Stonyhurst, England, uncovered a couple of scientific and historical howlers in the first edition, which I have tried to address here.

Cliff Stoll, astronomer/computer maven/raconteur and friend of many years, was both a general inspiration and also a close and careful reader of the text. (But I may never forgive him for pointing out that a passage I'd written about wave interference, one that I was especially fond of, had gotten the facts wrong and had to go…) And finally, I'd like to acknowledge Nick Nussbaum; an old friend from my MIT days, he first taught me the meaning of the word "ineffable."


And the Lord spoke to Job out of the whirlwind:
"Do you know the way to the dwelling of light?"
Was God merely reprimanding Job for his presumption? Or was He also issuing an invitation to explore?
In "...the Way to the Dwelling of Light" Dr. Guy Consolmagno SJ, an astronomer and Jesuit brother, shows us how our religious experience can illuminate our understanding of physics of light... and how, in that light, we can see what modern physics shows us in God's creation.
"A spirited romp through the mysteries and paradoxes of light: friendly, honest physics and a delight to read!"
Dr. Cliff Stoll, astronomer and bestselling author of *The Cuckoo's Egg* and *Silicon Snake Oil*
Cover photo: Sunrise and Rocca Priora photographed from the Specola Vaticana (Vatican Observatory). Castel Gandolfo, Italy, by Fr. Emmanuel Carreira SJ