1

What is Spatiotemporal Locality?

1 The Big Picture

Here's an old Russian/Jewish joke:

> The peasant Piotyr lives in the town of Kishinev early in the twentieth century. He has just witnessed the first demonstration in Kishinev of a shortwave radio receiver and transmitter. Someone has used the radio to converse with another person in Odessa, hundreds of miles away. Piotyr is astonished and asks his better-educated friend Ivan how this amazing invention works.
>
> "Very simple," Ivan says. "Imagine a dog so large it stretches from Kishinev to Odessa. You step on the dog's tail in Kishinev and it barks in Odessa. Do you follow that?"
>
> "I think so," Piotyr says, hesitantly.
>
> "Good," Ivan says. "Now just remove the dog." (Telushkin 1992: 60)

Of course, the joke is that Ivan doesn't succeed in making the radio less mysterious to Piotyr. The radio used to be called the "wireless" precisely because unlike a telegraph or telephone, there is no wire (or apparently anything else) connecting the sender to the receiver. Understandably, Piotyr fails to grasp how the sender in Odessa manages to cause his voice to be heard in Kishinev without anything like wires to carry it there. Piotyr understands vaguely how your stepping on one end of a dog causes a bark from the other end: something inside the dog connects the tail to the mouth. But without the giant dog between Odessa and Kishinev, there is no connection between the sender and receiver, and so the original mystery remains: how does the cause in

Odessa have an effect in Kishinev without anything to carry its influence across the space in between?

Like Piotyr, we expect an effect to be located at the site of its cause. For example, suppose that a window shatters. The cause might have been a certain baseball's colliding with it. But suppose the ball was never anywhere near the window. Then the ball could not have had anything to do with the shattering. A cause must be *local* to its effect.

Of course, your stepping on a dog's tail causes the dog to bark despite the dog's mouth being some distance away from its tail. But this is understandable: a *chain* of causes and effects connects your stepping on the dog's tail to the dog's mouth barking. Your stepping on the tail has the *direct* effect of compressing certain skin cells. This compression stimulates nerve cells, which influence cells in the dog's brain, which excite nerve cells projecting to the dog's mouth, throat, and chest, which cause certain muscles to move, which ultimately produces a bark. The *direct* effect of each cause in the chain occurs locally to that cause. Through this chain, your stepping on the dog's tail has an *indirect* effect some distance away.

We expect an effect to be local to its cause both in space and time. For example, the cause of the window's shattering must have been happening when the window began to shatter. Of course, we should again distinguish direct from indirect causes. For example, suppose that lightning strikes a tree, ultimately causing a fire. The fire may not ignite until some minutes after the lightning strikes. But the fire is not a *direct* effect of the lightning strike. Rather, the lightning directly caused some chemical changes in the wood that it struck. These changes ultimately produced the fire. If, a moment after it was struck by lightning, the tree had been no different from the way it was a moment before, then (we think) the lightning could not have caused the fire that ignited later. A cause cannot make itself felt a few minutes after it occurs without having effects at every moment in between, just as (Piotyr presumed) a cause cannot have an influence at some distant location without having effects at all of the locations in between. I will call this important idea *spatiotemporal locality*.

"Action by contact," as when two billiard balls collide, is an obvious illustration of spatiotemporal locality. (However, even "action by contact" may raise serious puzzles, as we will see shortly.) "Action at a distance," as when two bodies separated by a gap in space affect each other gravitationally, electrically, or magnetically, seems to violate spatiotemporal locality. But these interactions are local if *fields* are real;

the gravitational, electric, and magnetic forces felt by a given body would then be *directly* caused by the corresponding fields *at that body's location* rather than by some *distant* body.

What *reasons* are there for believing fields to be real? (We cannot justify this belief on the grounds that gravity and electromagnetism would then obey locality – unless we *already* had a good reason for believing in locality!) As we will see, many philosophers and physicists worked to ascertain whether fields are real, leading to classical electromagnetic theory and the special theory of relativity. Furthermore, the reality of fields is bound up with many other fascinating issues in the interpretation of physical theories, including the character of energy and matter as real kinds of "stuff," the sense in which electricity has been *unified* with magnetism (and energy with mass), the apparent need for fields to be real in order for there to be explanations of certain facts, and the distinction between a scientific theory and its philosophical interpretation. By exploring issues arising from spatiotemporal locality, we will run into many (though not all!) of the most fundamental and historically significant questions in the philosophy of physics. This is the trail we will follow throughout this book.

Our first step must be to clarify "spatiotemporal locality." That's the task of this chapter.

2 Causal Relations between Events

Can there be space or time separating a cause from its direct effects, or must a cause be local to its effects? I will presume that this question makes sense. But it makes sense only if a cause and its effects have locations in space and time. Otherwise, we can't ask whether they must be near each other.

For causes to have locations in space and time, a cause needn't be confined to a single *point* in space and *instant* in time. A cause may occur throughout some time period or over an entire volume in space. For example, suppose that the shattering of a certain window was caused by a baseball's colliding with it. If the window was three feet wide, two feet high, and a quarter-inch thick, then its shattering took place in the entire region of space that the window was filling.

I just referred to the shattering as "taking place." Its cause, the collision between a baseball and the window, is also something that "occurs." Accordingly, it seems natural to call the collision and shattering *events*.

An event is a particular thing that happens somewhere and somewhen. Since it makes sense to ask whether a cause must be local in space and time to its effects, causes and effects must be events.

Here's another example: my pressing steadily but gently on my car's brake pedal causes the car to slow down gradually to a stop. The cause and its effect occur throughout a certain temporal *interval* a few seconds long. Since the car is moving while it is slowing down to a stop, the cause (my pressing on the brake) has a different location at different moments. The same applies to the effect: at a given moment, the car's slowing down to a stop is occurring wherever the car is. Since the car is moving, the event of the car's slowing down to a stop *moves* while it occurs.

Certain events would not ordinarily be described as "occurring" or, indeed, as "causes." For example, the lightning strike is *only one* of the events that conspire to cause the fire's igniting. It is not the *complete* cause. One of the other causes that combines with the lightning strike to form *a complete set of the fire's causes* is the tree's being dry. Another is the tree's being surrounded by oxygen. Ordinarily, if someone asked us why the tree caught fire, we might mention only the lightning strike; the person who asked presumably already knew that the tree was dry and surrounded by oxygen. Perhaps she mainly wanted to know why it caught fire *just when it did* rather than at some earlier moment or never. Only the lightning strike explains why; its occurrence is the only thing making the moment at which the tree caught fire relevantly different from other recent moments. Nevertheless, the tree's being dry and surrounded by oxygen helped to cause the fire to ignite. They are *causally relevant*, and that is all I mean by referring to them as "causes" of the fire.[1] We would not usually refer to them as "events" that "occur." Yet that is what they are.[2]

To help clarify what an event is, let's contrast events with facts. An event occurs somewhere somewhen; there are times when and places where it is not occurring. On the other hand, a fact holds everywhere and everywhen. For example, if it is now (at 1:45 p.m. on October 22, 1999) a fact that I am in my office at 1:45 p.m. on October 22, 1999, then in 1543 in Florence it was a fact that I am in my office at 1:45 p.m. on October 22, 1999, and this fact will still be a fact on Mars in the twenty-second century (see box 1.1). Of course, in Florence in 1543, no one yet *knew* this fact about me!

Perhaps a fact is just a truth.[3] It is true that I am in my office at 1:45 p.m. on October 22, 1999. But the window's shattering, an event, is

Box 1.1
The openness of the future

Some philosophers reject this view of facts as conflicting with the future's "openness": If it is already true *now* that humanity will colonize Mars next century, then the future is settled, so our current actions cannot affect it (Prior 1953). This worry is misplaced. If humanity colonizes Mars next century, our current efforts will have made humanity's later success possible. Success was not inevitable; were we now not exploring space, it would be true neither now nor later that humanity colonizes Mars in the twenty-second century.

not *true*. It's not false either. It is not the *sort* of thing that *could* be true or false. "The window's shattering" is like "My house" in that it would be a "category mistake" to call it true (or false). The window's shattering *happens*; it *takes place*. Of course, it is true *that* the window shatters. That the window shatters is a fact. A fact *involves* various particular things, such as a window. An event *is* a particular thing, and there are facts about an event just as there are facts about a window. Here is a fact about an event, the window's shattering: it was caused by the baseball's colliding with the window.

Of course, an event and a house can both be particular things and nevertheless differ in important ways. A house is made of matter, has mass, and is a certain color, whereas none of these properties can be possessed by any event. Furthermore, two events can occur at the same place and time (Lewis 1986: 245). For example, in Newtonian physics, when a body feels a certain force and is thereby accelerated, the body's feeling that force is an event occurring at the same place and time as the body's accelerating in response. On the other hand, two objects made of matter cannot occupy the same place at the same time. (Or so we are inclined to think – but see the next section!)

Although an event is less tangible than a house, they are alike in certain basic respects simply because they are both particular things. For instance, many different expressions refer to the same house: "the house next door," "the Jones's house," and "the building that I am thinking about right now." Likewise, many different expressions pick out the same event: "the shattering of the window," "the cause of that sound of glass breaking," and "the first event I told my wife about when she came home."

Facts do not work this way. Even though Marc Lange is the author of this book, the fact that Marc Lange is in his office is not the same as the fact that the author of this book is in his office. The second fact requires the cooperation of this book; the first fact does not. The second fact could have held without the first fact's holding. (Try to imagine how!) One fact (that Marc Lange is in his office) can *entail* another (that Marc Lange has an office) in the sense that it would be a *contradiction* for the latter to be false and the former true. The fact that Marc Lange is in his office stands in different inferential relations from the fact that the author of this book is in his office. (For example, only the former implies that Marc Lange has an office.) Hence, they are different facts. Events cannot stand in inferential relations at all. No sentence's truth is guaranteed by the truth of "The window's shattering" because "The window's shattering" is not the sort of thing that *can* be true.

A causal relation sometimes seems like a relation between *facts*, as when we say "The window shattered because the baseball collided with it." Of course, we could instead have said "The shattering of the window was caused by the baseball's colliding with it," portraying *events* standing in a causal relation. Perhaps causal relations hold among facts and also among events (Bennett 1988). Even if this is so, our interest lies in spatiotemporal locality, and therefore we must look at causal relations among events, since facts have no spatiotemporal locations. On the other hand, perhaps events *alone* stand in causal relations (Davidson 1980). Then the "because" relation in "The window shattered because the baseball collided with it" cannot be causal, even though it is *explanatory*: the fact that the baseball collided with it explains why the window shattered (rather than remaining intact).

Suppose the same event is picked out by "the lightning's striking the tree" and "the event that I am thinking about right now." So the event that I am thinking about right now caused the fire's igniting. Of course, my thinking about it right now did not enable it to cause the fire! This just goes to show that the properties an event possesses by which we identify it on some occasion need not be the properties that make it another event's cause. The lightning strike's electrical properties are the only features it possesses intrinsically that are *causally relevant* to the fire's igniting. (I will say more about "intrinsic" properties in section 5.)

3 Action by Contact

The idea behind spatiotemporal locality is that there can be no gap in space or time between a cause and its direct effects. It would strike many of us as mysterious – magical, even – for two billiard balls to interact without touching. If a person in Odessa talks to a person in Kishinev – if, as we say, they are "in touch" or "in contact" with each other – then we expect them to be touching or for there to be another object between them that they are both touching.

"Action by contact" has often been considered unproblematic. Here, for instance, is William Thomson (Lord Kelvin, the Scots mathematician and physicist) in 1884:

> It seems to me that the test of "Do we or do we not understand a particular point in physics?" is "Can we make a mechanical model of it?" (1987: 111, also 206)

But as the following discussion shows, causation by mechanical impact – touching – is not as straightforward as it might appear.

Consider points A and B in space (or instants A and B in time). They are separated by some distance (or span of time), which must contain other points (or instants). Suppose we select one: C. Now we can repeat the procedure: between A and C, there is some separation, within which lies another point (or instant): D, and so on. Hence, between A and B, there are infinitely many points (or instants). In short, classical physics assumes space and time to be "dense": between any two points in space (or moments in time), there is another. Closely related is the idea that classical physics treats trajectories as "continuous." For example, if between times $t = 0$ and $t = 1$, a body moves from the spot with coordinates (0,0,0) to the spot with coordinates (1,1,1), then for any real number n between 0 and 1, there is an instant of time $t = n$ and at that instant, the body has a definite location (coordinates (a,b,c), where a, b, and c are real numbers).

Imagine, then, that the basic constituents of matter are *point* particles. They have no length or width; each has zero volume. The only way they can touch is for them to occupy *the same point* in space. (They cannot touch by occupying *neighboring* points since there are no neighboring points; between any two points, there are infinitely many others.) But if two particles can occupy the same point at the same instant, then they cannot be hard, solid, impenetrable bodies. Indeed, since they can get inside each other, it is difficult to understand why they bounce off each other; when they collide, they might just as well pass right through each other! This seems bizarre: the whole point of matter is its solidity – that one bit of matter excludes every other from the space it occupies.

To avoid this puzzling result, imagine the two bodies not to be geometric points. Take two billiard balls colliding. The causal relation between them should definitely qualify as spatiotemporally local – as the *opposite* of action at a distance!

Let each billiard ball be a sphere of radius r. Each ball occupies all of the points at a distance from its center less than or equal to r. What happens when two balls touch? Their centers are separated by a distance of $2r$. A point located exactly between the two balls, at a distance r from the center of each, is where the balls touch when they collide: the "point of osculation" (literally: the kissing point). *Both* balls occupy that point, since (to repeat) each ball occupies all of the points at a distance from its center less than *or equal to* r. We have re-created the problem we faced with the point particles: two bodies occupying the same location at the same moment. This *seems* problematic, though the Irish physicist George Francis FitzGerald (1851–1901, perhaps best remembered for the "Lorentz–FitzGerald contraction" later incorporated by Einstein into relativity theory) wrote nonchalantly:

> When rigid bodies act on one another by non-slipping contact, certainly the coordinates of the point of contact are common to the two systems. (1895: 285)

Perhaps FitzGerald is correct. The northern and southern hemispheres of Earth are in contact and together constitute the entire Earth, leaving nothing out. There is no reason for the Equator to belong to one hemisphere rather than the other. So the Equator must belong to *both*. On the other hand, the Equator is merely an artificial, conventional boundary drawn *within* a single body. Unlike two billiard balls, the two hemispheres are not distinct bodies in contact.

The problem of overlapping bodies would not arise if a billiard ball occupied only the points at a distance *less than* r from its center. This conception might be inspired by a question famously asked in an 1893 paper by the American philosopher, physicist, and mathematician Charles Sanders Peirce (1839–1914). Regarding a solid black circle on a white piece of paper, Peirce (1933: 98) asked: Does the boundary line belong to the circle or the background? Are the points on this line black or white? Remember, a point has no width so it cannot be half black and half white. Moreover, a line in geometry is only one point wide, so it cannot be that each point on the outer half of the line is entirely white and each point on the inner half is entirely black.

Apparently, no view is so weird that it has never been defended in an interesting way by some philosopher. (That, to me, is one of the joys and strengths of serious philosophizing.) This case is no exception: Franz Brentano (1838–1917, Austrian philosopher and psychologist) suggested (1988: 41) that at the boundary line, a black line and a white line coincide; each point on the line is entirely black *and* entirely white!

At least, there is no obvious reason to award the boundary line to the circle rather than to the background; each seems to have as good a claim on it as the other. We might, then, be inclined to say that a point on the line is neither black nor white. On this view, neither the solid circle nor the background includes the boundary, though each region extends right up to the line. Analogously, perhaps the points at a distance of exactly *r* from the billiard ball's center do not belong to the ball, though the ball extends right up to them (Kant 1985: 63).

In that case, however, two billiard balls in contact must have a gap at least a point wide separating them. The two balls would be touching and yet any continuous path in space from one to the other must cross a point occupied by neither. This doesn't seem like contact at all![4] Furthermore, if the balls could interact across a gap of one point, then why not across a gap of two points, or even two miles? Action "by contact" would then seem as "magical" as action at a distance!

This problem does not arise if *one* of the colliding balls occupies all of the points at a distance from its center *less than or equal to r*, whereas *the other* occupies only those points *less than r* from its center. When the balls touch, they are like two parts of a single body: there is continuity without overlap. In defending this view, Bernard Bolzano (Czech philosopher and mathematician, 1781–1848) defined

> the contact of two bodies as taking place when the extreme atoms of the one, . . . together with certain atoms of the other, form a continuous extension. (1950: 168)

On this view, the answer to Peirce's question is that the boundary line belongs either to the black circle or to the white background (but not to both). Perhaps if we put black paint on white paper, then the boundary line belongs to the circle, whereas if we paint a white surround on a paper that was already black, then the line belongs to the background.

But applied to billiard balls, this view is unattractive. Brentano (1988: 146) called it "monstrous"! Surely all billiard balls of radius *r* are alike in either including or excluding the points at radius *r*. If some are "inclusive" and others "exclusive," then (on Bolzano's view of what contact requires) an "inclusive" ball cannot collide with another "inclusive" ball (since they cannot occupy the same point at the same time) and an "exclusive" ball cannot collide with another "exclusive" ball (since there is at least a point in between). Billiards would be a very difficult game.

As I warned you, action by contact is not straightforward. I have just surveyed three ways of understanding two colliding billiard balls: as sharing a point, as separated by a point, and as momentarily forming a continuous body. Each view encountered severe difficulties – perhaps not severe enough to prove it false, but enough to be worrisome. Accordingly, let's rethink our picture of collisions.

To gain some perspective on our problem, let's examine instead a single billiard ball moving in empty space. Suppose its speed at time $t = 0$ was 5 meters per second, and that from $t = 0$ to $t = 5$, it feels no forces. By Newton's laws of motion, it undergoes no acceleration and so continues moving at 5 m/s through $t = 5$. What, then, caused the body at $t = 3$ to be moving at 5 m/s? By spatiotemporal locality, the *direct* causes cannot be separated by any gap in space and time from the effect. Given Newton's laws, the direct causes must be the body's already moving at 5 m/s at some moment and its feeling no forces between that moment and $t = 3$.[5] But it's already moving at 5 m/s *when*?

The body's moving at 5 m/s at $t = 3$ cannot be a cause of the body's moving at 5 m/s at $t = 3$, since an event cannot be one of its own causes. And according to locality, the body's moving at 5 m/s at some moment *earlier* than $t = 3$ (say, at $t = 2.5$) cannot be a *direct* cause of its moving at 5 m/s at $t = 3$, since this would create a gap in time (of 0.5 second) between the effect and its direct cause. This is like the problem we encountered in trying to apply locality to two colliding billiard balls. There we faced the unwholesome choice between the balls overlapping at a point or not overlapping and so being separated by at least a point. With the isolated billiard ball, we face an analogous unwholesome choice: its moving at 5 m/s at $t = 3$ is caused either by itself or by its speed at a moment separated by a gap from $t = 3$.

These problems all arise from space and time being "dense." Between any two points in space (or time), there are infinitely many other points: a gap. Apparently, for a cause to be local to its direct effect – for there to be no gap between them – the cause must share a point in space (and moment in time) with its direct effect. But this seems implausible: it is problematic for two bodies to overlap, and an event cannot be its own cause.

A possible solution is to rethink the distinction between direct and indirect causes. Since time is dense, there is no *unique* "direct" cause of the isolated ball's moving at 5 m/s at $t = 3$. There is no moment where $t = 3$ is *the next moment*. So there is no particular moment where the body's moving at 5 m/s at that moment (and the body's feeling no force from then to $t = 3$) is *the direct cause* of the body's moving at 5 m/s at $t = 3$.

However, within *any* interval prior to $t = 3$, *no matter how short*, there is a complete set of causes of the body's moving at 5 m/s at $t = 3$. For example, within the interval between $t = 2.9$ and $t = 3$, there is the body's moving at 5 m/s at $t = 2.99$ and its feeling no forces between $t = 2.99$ and $t = 3$ inclusive. Within the narrower interval between $t = 2.999$ and $t = 3$, there is another complete set of causes: the body's moving at 5 m/s at $t = 2.9999$ and its feeling no forces between $t = 2.9999$ and $t = 3$ inclusive (see box 1.2). There is a complete set of causes *arbitrarily near* in space and time to the effect. Therefore, I suggest, there is no "gap" between cause and effect, satisfying spatiotemporal locality.

It might be objected: an arbitrarily small gap between cause and effect is *still a gap*. There *is* a separation between $t = 2.99$ and $t = 3$, and

Box 1.2
A complete set of causes

A "complete" set of causes of an event is like *one* link in the causal chain leading to that event. It is "complete" in the sense that it constitutes a complete link in that chain, not in the sense that it includes all of the event's causes (that is, all of the other links). The same event may therefore have *many* complete sets of causes. An analogy: My four grandparents might (loosely) be termed a "complete" set of my causes, even though there are other links in the chain (such as the link consisting of my parents). Three of my grandparents, without the fourth, would not be a complete set of my causes.

there is a smaller separation between $t = 2.9999$ and $t = 3$, and as we extend the sequence further, there will always be some finite separation. Accordingly, Kline and Matheson (1987) insist that contact requires no separation at all.

But then spatiotemporal locality is violated by exactly the cases that we would initially have considered prime examples. Rather than argue about what "contact" really means, I suggest that we define "spatiotemporal locality" in whatever manner best fits our initial ideas about the sorts of cases that should count as exemplifying it. We can then investigate whether this sort of locality holds in every actual case.

What does this approach say about billiard balls colliding? If each ball occupies only the points less than r from its center, then on this approach, the balls are "touching" when they are separated by only one point. It may appear contradictory for contact to involve a gap, even of one point. But if a "gap" must be a *region* of space – a finite non-zero volume – then there is *no gap* between the balls. If a "distance" must be some *finite* quantity, then two balls separated by a single point have no distance between them and so fail to involve action at a distance. That two bodies can be "aware" of each other across an unoccupied point does *not* suggest that they can "know" about each other across a gap of two points – which is a gap of infinitely many points, because space is dense. A "gap" of one point is *qualitatively different* from anything larger: it is not a *region* of space. Again, I recommend that "spatiotemporal locality" be defined to capture the cases to which we think it ought to apply most easily. Otherwise, the issue of whether "spatiotemporal locality" holds becomes quite different from the issue we set out to examine.

Of course, I have given no reason to believe that a billiard ball occupies only the points at distances less than r from its center. If instead each ball occupies all of the points at distances less than *or equal to r*, then perhaps we should join FitzGerald in allowing two bodies to share a point, but prohibit them from sharing a region. This allows spatiotemporal locality to apply to "action by contact." Does the "impenetrability" of matter preclude bodies from overlapping in a finite non-zero volume but permit them to overlap at a single point?

After we have introduced fields of force, we may be able to picture the "collision" between two balls as involving repulsive fields driving the balls apart. Each ball interacts only with the field "surrounding" the other ball. A "collision," then, is just a close approach between the two balls – close enough that each ball penetrates deeply into the other's field, to where that field becomes very strong. Spatiotemporal locality is guaranteed by each ball's interacting only with the field *at its location*. (These are "coming attractions.")

4 Spatial, Temporal, and Spatiotemporal Locality Defined

Our discussion of locality began with the thought that two things directly "in touch" with each other must be touching, so that nothing can be squeezed between them. Given the density of space and time, this became the idea that no *region* of space or *interval* of time can separate them – that nothing *extended* in space or time can fit between them. When no *unique* set of *direct* causes exists, locality says that a cause can have an influence at some distant location (or at some later moment) only by having effects at every location (or during every moment) in between. In other words, no extended region of space (or interval of time) exists around the effect that does not contain a complete set of its causes.

To sharpen this notion of spatiotemporal locality, let's distinguish "temporal locality" and "spatial locality."

Here's what "temporal locality" means. Take an event E. Choose a finite, non-zero amount of time τ: perhaps let $\tau = 1$ sec, or even let $\tau = 0.0000001$ sec. Temporal locality says that no matter how short τ is, the moments within τ of E contain a complete set of E's causes.[6] In other words, E is not surrounded by some finite non-zero gap (time τ in length) during which *no* complete set of E's causes occurs. More precisely:

> *Temporal locality*: For any event E and for any finite temporal interval $\tau > 0$, no matter how short, there is a complete set of E's causes such that for each event C (a cause) in this set, there is a moment at which it occurs that is separated by an interval no greater than τ from a moment at which E occurs.[7]

C and E may each last for more than a single moment. Temporal locality requires that *at least one* of the perhaps many moments at which C occurs be within τ of *at least one* of the perhaps many moments at which E occurs. For simplicity, however, consider the isolated billiard ball case, where E (the ball's moving at 5 m/s at $t = 3$) lasts for only a moment. If $\tau = 0.1$ second, then here is a complete set of causes where each event C in the set occurs at a moment no more than τ away from E: the set consisting of the ball's moving at 5 m/s at $t = 2.99$ and the ball's feeling no forces between $t = 2.99$ and $t = 3$. No matter how small τ becomes (so long as τ remains non-zero), there remains such a set.

Likewise, "spatial locality" means that E is not surrounded by some finite gap in space, extending out to a non-zero distance δ, in which *no* complete set of E's causes occurs. More precisely:

> *Spatial locality*: For any event E and for any finite distance $\delta > 0$, no matter how small, there is a complete set of causes of E such that for each event C in this set, there is a location at which it occurs that is separated by a distance no greater than δ from a location at which E occurs.

C and E may occupy extended regions of space. Spatial locality requires that *at least one* of the perhaps many points at which C occurs be within δ of *at least one* of the perhaps many points at which E occurs. (In fact, given two events, C_1 and C_2, that form a complete set of E's causes, spatial locality is satisfied if there is a point where E occurs that is within δ of C_1 and *another* point where E occurs that is within δ of C_2, even if there is no *single* E-point that is within δ of *both* C_1 and C_2.)

Spatiotemporal locality is not simply spatial locality plus temporal locality. That is because spatial locality requires only that there be a point in space at which C occurs *sometime* that is separated by a distance no greater than δ from a point in space at which E occurs *sometime*. This imposes no restriction at all on how close those times have to be. It could be that we cannot bring these times close together without disrupting the spatial closeness. Then *spatiotemporal* locality would be violated.

In other words: Suppose that spatial locality holds of E. Consider a complete set of E's causes where each event C in that set occurs *sometime* within a distance $\delta > 0$ of some point at which E occurs *sometime*. Suppose that every C in this set is separated in time from E by at least the interval $\tau > 0$. That is, every C in this set occurs *outside* the small interval τ around E. Suppose further that all this is also true for *every* complete set of E's causes inside δ. So to get causes within τ of the effect, as temporal locality demands, we cannot use causes within δ. Nevertheless, it may be the case that no matter how small $\tau > 0$ is, there is a complete set of E's causes within τ of E. But once τ falls beneath a certain value, every C in any complete set of E's causes occurring within τ of E is never within δ of a point where E occurs sometime. Then although *both* temporal locality *and* spatial locality hold of E, the sets of E's causes that make spatial locality hold are not the same as the sets of E's causes that make temporal locality hold. Intuitively, *spatiotemporal locality* then fails to hold, since it requires that we be able to diminish arbitrarily the spatial *and* temporal gaps *together*:

> *Spatiotemporal locality*: For any event E, any finite temporal interval $\tau > 0$, and any finite distance $\delta > 0$, there is a complete set of causes of E such that for each event C in this set, there is a location at which it occurs that is separated by a distance no greater than δ from a location at which E occurs, *and* there is a moment at which C occurs *at the former location* that is separated by an interval no greater than τ from a moment at which E occurs *at the latter location*.

Spatiotemporal locality requires spatial locality and temporal locality, but spatial locality and temporal locality could both hold even if spatiotemporal locality is violated.

Let us look at some examples. Take gravity according to Newton's law: A point of mass M exerts on a point of mass m, at a distance r, an attractive force of magnitude GMm/r^2. *Temporal* locality is satisfied since the effect (the gravitational force at time t impressed by one body on another) has a complete set of causes *simultaneous* to it (namely, the masses and separation of the bodies *at t*). However, *spatial* locality is violated; this is action at a distance in its purest form. For example, the point masses composing the Sun cause forces on the point masses composing the Earth across a gap of some 93 million miles. On this interpretation of the physics, there are no causally relevant events between the Sun and the Earth. (This will change when we introduce fields in the next chapter.)

Alternatively, imagine that the Sun's gravitational influence is *not* felt *instantly* on Earth. Suppose there would be a 500-second delay between a change in the Sun's mass and the corresponding change in the force on Earth. That is how long sunlight takes to reach Earth. Sunlight arrives at Earth only after first being one mile from the Sun, shortly later two miles, then three, and so on. But imagine that there are no events between the Sun and Earth that are causally relevant to the gravitational force exerted by the Sun on the Earth. Rather, the

Box 1.3
Other senses of locality

In the philosophy of physics, "locality" is used in many different senses besides the senses given here (Earman 1987). It is sometimes used to mean that C can be a cause of E only if the separation in space between C and E, divided by their separation in time, is finite (Reichenbach 1958: 131–2) – in other words, only if the speed at which you would have to travel to get from C to E is finite (even if high). The 500-second delay between the cause on the Sun and the effect on the Earth would not constitute a violation of locality in this sense. Alternatively, "locality" sometimes means that causal influences propagate no faster than the speed of light (in a vacuum), so that C cannot be causally relevant to E if they are "spacelike separated" – that is, if one would have to exceed the speed of light in order to reach E from C. These notions are different from "spatiotemporal locality." For example, a field version of Newtonian gravitation satisfies spatiotemporal locality (because of the field, there are no gaps) but causal influences propagate infinitely fast (that is, instantaneously throughout the universe). Conversely, if C and E are located at the same place at different times, C preceding E, then they are guaranteed to be "local" in the sense that an influence could get from C to E without traveling faster than the speed of light in a vacuum (it would not have to travel at all). But spatiotemporal locality could be violated (as in the next example in the main text). Other senses of "locality" have nothing directly to do with causal relations, such as the idea (discussed in chapter 9) that a system's intrinsic properties are determined by the properties of its parts, and their intrinsic properties are determined by the properties of their parts, and so on all the way down to the properties at points, the most "localized" properties.

latest cause of a certain gravitational force on Earth is an event occurring on the Sun 500 seconds earlier. Then not only spatial locality but also temporal locality would be violated (but see box 1.3).

The next example is entirely fictional. Suppose a body passes through a given location, leaving no trace behind it. One hour later, another body passes through there and is affected by the first body's having been there one hour earlier. Remember, nothing was left behind by the first body to be encountered later by the second. So spatial locality applies whereas temporal locality is violated.

Of course, this example may strike you as wacky: How could the first body's passing through affect the second body one hour later unless the first body left a residue behind? In that case, however, the residue would be a *local* cause of the effect on the second body. If the original example seems bizarre to you, that's perfectly fine; your reaction merely reflects how plausible locality seems to you. Nevertheless, locality appears to be violated by gravitational and electromagnetic interactions. We will look into this in subsequent chapters.

5 Intrinsic Properties and Noncausal Connections

Return to the case we just discussed, which was meant to demonstrate that temporal locality can be violated while spatial locality is satisfied. Here's a strategy you might consider for making this case satisfy temporal locality too. Notice that just as the second body passes through the given location, the following event occurs there: its becoming one hour since the first body passed through. This is a peculiar way to identify an event; I have mentioned nothing happening except for a certain amount of time having passed since the first body came through. This event occurs at the same moment as the effect on the second body. So if this event is causally relevant to the effect, then temporal locality is upheld! But this event must not qualify as causally relevant, else temporal locality will hold *no matter what*! It will hold *trivially*, no longer allowing us to draw the kind of interesting distinction that we wanted to draw between action by contact and action at a spatiotemporal distance.

A similar tactic would trivialize spatial locality. It is supposed to be violated by gravity operating according to Newton's law. Some gravitational forces on Earth are caused by the matter in a certain region of the Sun having mass *M*. This cause occurs in that region of the Sun, far away from Earth. But at the same moment, a certain event occurs *on Earth* (spatially *local* to the effect). Here it is: Earth at that moment belonging to a universe where a certain region of the Sun has mass *M* at that moment! (Compare Shoemaker 1969: 378–80.)

Obviously, if our locality principles are going to have any bite, we must somehow conclude that these events cannot be causally relevant. Let's try to find a good reason for doing so. Take the property of being red. Suppose it is possessed in a certain place at a certain time. For example, a shirt there and then is red. So redness is "instantiated" there and then. An event involves certain properties being instantiated in a certain region (or point) of space at a certain interval (or instant) of time, which constitute the event's "spatiotemporal location."

Take the property of containing 5 kilograms of matter. For it to be instantiated at a certain spatiotemporal location, something must be going on *at that location* – namely, 5 kg of matter must be there. That this property is instantiated at a given spatiotemporal location imposes some constraints on *that location*, but seems to impose no logical constraints on what properties may be instantiated at *other* locations. For example, no *contradiction* would ensue if 5 kg of matter occupied a given point in space at a given moment, but at every other moment in the universe's past and future history, there was less than 5 kg of matter in the entire universe. Of course, this scenario may violate some *law of nature* – for instance, a law that the total quantity of matter in the universe must remain constant (the conservation of matter). But presumably, a violation of this law would involve no *logical contradiction* (unlike "There exists a round square," which is logically contradictory). There being 5 kg of matter here now seems to impose no *logical* restrictions on what properties can be instantiated elsewhere elsewhen.

In other words, containing 5 kg of matter is an *intrinsic* property of a region of space. Roughly speaking, an intrinsic property of a thing requires something only *of that thing*; its intrinsic properties depend only on itself, not on its relations to other things. (Of course, a body's intrinsic properties can be *changed* by other things' effects on it, but the properties are possessed by the body entirely in virtue of what the body is like.) A body's mass, electric charge, and shape are among its intrinsic properties (at least in classical physics).

In contrast, a typical non-intrinsic ("extrinsic") property is the property of standing in a certain *relation* to something else. For example, the property of being 5 meters away from 5 kg of matter is *not* an intrinsic property: its instantiation here now requires the cooperation of other spatiotemporal regions. Obviously, for this property to be instantiated here now, there must be 5 kg of matter 5 meters away.

A body's intrinsic properties do not depend on other bodies (or anything else external to the body). For example, suppose I now possess the property of being the tallest person within 100 meters. This is not one of my intrinsic properties, since my possessing it depends on certain other people (who are taller than I am) not now being within 100 m. In other words, my now being the tallest person within 100 m depends not just on what is going on *at my location*, but also on what is happening in the rest of the region 100 m around me. In contrast, my being 148 cm tall is one of my intrinsic properties; it depends only on me. If different people come within 100 m of me but I am left alone, I will still be 148 cm tall. However, I may no longer be the tallest person within 100 m.

Having said this, I must now take some of it back. Although the property of containing 5 kg of matter is *supposed* to qualify as intrinsic, its instantiation here now does *sort of* logically require the cooperation of other spatiotemporal locations: for this property to be instantiated here now, there is a property that any point 5 meters away must instantiate now. Can you think of it? Of course: the property of being 5 m away from 5 kg of matter! However, this property is not intrinsic. So we have arrived at this idea: A property is "intrinsic" if and only if its instantiation at one spatiotemporal location (a point or a region) puts no logical restrictions on the *intrinsic* properties instantiated at locations that do not overlap the first location.

Of course, this idea cannot be used to *define* "intrinsic property" since "intrinsic" appears on *both* sides of the "if and only if." Nevertheless, it is a useful idea; let's refine it. Consider the property of containing 5 kg of matter *or* being 5 m away from 5 kg of matter. This property is *not* intrinsic, yet its instantiation here now imposes no logical restrictions on the intrinsic properties instantiated elsewhere elsewhen. (For example, that this property is instantiated here now does *not* logically require that there be 5 kg of matter 5 m away; the property is instantiated here now if there is 5 kg of matter here now, regardless of what is going on elsewhere.) This example suggests that for a property to be intrinsic, it is necessary *but not sufficient* that its instantiation here now impose no logical restrictions on the intrinsic properties instantiated elsewhere elsewhen (Lewis and Langton 1998).

Now take the property of having a brother (one who is currently living, I mean). Jones's here now having a brother logically restricts the intrinsic properties instantiated in other regions of space-time: for example, there must now somewhere be another human being (namely, Jones's brother). Consider, then, the property of *not* having a brother (alive right now) – being "brotherless." Since *having* a brother is a non-intrinsic property, being *brotherless* must be non-intrinsic. Yet Smith's being brotherless here now imposes no logical restrictions on the intrinsic properties instantiated elsewhere elsewhen. (Smith can be brotherless whether or not there are other human beings around.) So let's amend our proposal to read: if property P is intrinsic, then no logical restrictions on the intrinsic properties instantiated at other places or times are imposed *either* by P's being instantiated in a given spatiotemporal region *or* by not-P's being instantiated there. So if containing 5 kg of matter is an intrinsic property, then being 5 m away from 5 kg of matter is *not* an intrinsic property, since the latter's instantiation here now logically requires the former's instantiation now at the circle formed by the points 5 meters away.

We can use the notion of an intrinsic property to explain why the events threatening to trivialize our definition of "spatiotemporal locality" cannot be causes. Consider again a gravitational force at time t on a chunk of the Earth having mass M. One of its causes (according to Newton's law) is a region of the Sun (a distance r away) possessing mass m at t. At the place on Earth where this gravitational force occurs at t, there also occur at t various events possessing this *non-intrinsic* property: occurring a distance r away from a region containing matter having mass m. If such an event, together with the Earth chunk's having mass M, formed a complete cause of the gravitational force's occurring on Earth, then, weirdly, spatial locality would apply: the causes and effect would all be at the same place. But we can rule out this cheap way of satisfying spatial locality by noticing that in order for such an event to be a *cause* of the gravitational force on Earth, one of this event's *non-intrinsic* properties would have to be *causally relevant* to the gravitational force.

Let's take this more slowly. At the end of section 2, I mentioned that if C causes E, then only some of C's properties are "causally relevant" – that is, enable C to cause E. If C is lightning's striking the tree and E is fire's igniting, then C's being an event I am thinking about right now, one of C's *non-intrinsic* properties, is *not* causally relevant to E. Here's another example: If E is the window's shattering, then C's causally relevant features include its involving a hard object colliding with the window. C's occurring at the same instant as Jones's screaming "Not my window!" is one of C's *relations* to other events – one of C's *non-intrinsic* properties. It is *not* causally relevant to E. Likewise, only some of E's properties are causally relevant to C (that is, enable E to be caused by C). E's occurring exactly where C occurs is one of E's causally relevant, non-intrinsic properties.

Considering examples like these, I will assume that if C is a cause of E, then in connection with this causal relation, C's and E's causally relevant features must be some of C's *intrinsic* properties, some of E's *intrinsic* properties, and some of the spatiotemporal relations between C and E (such as their being separated by a distance r and occurring simultaneously) – and that is all! None of C's or E's *non-intrinsic* properties can be causally relevant, with the sole exception of C's and E's spatial and temporal relations to each other.

Suppose a complete cause of the gravitational force at t on a chunk of Earth were that chunk's possessing mass M at t and some other event C taking place *at that region of the Earth* at t. Then given Newton's law of gravity, C's causally relevant properties must include its occurring at a distance r from matter with mass m. But this is not an *intrinsic* property of C, since it has nothing to do with what is happening where and when C occurs (on Earth at t). No matter what is happening on Earth at t, it would be happening at a distance r from matter with mass m. Furthermore, C's occurring at a distance r from matter with mass m is not C's standing in some spatiotemporal relation *to E* (the gravitational force's occurring). So C, the event on Earth, cannot be a cause of E (Kim 1974; Lewis 1986: 262–4).

The same argument applies to the other example I gave earlier, in which a body passes through some location leaving no trace behind, but its having passed through there affects a second body passing through one hour later. Any event C occurring at the given location when the second body passes through automatically has the property of occurring one hour after the first body passed through. But this is not one of C's *intrinsic* properties (or C's spatiotemporal relation to the effect on the second body). So it cannot make C a cause of that effect. The events occurring at the given location when the second body passes through could (without any logical contradiction) have been *intrinsically* just as they actually were even without the first body's having passed through an hour before.

I have assumed that an effect's causally relevant properties must be intrinsic to it or involve its spatiotemporal relations to its cause. For example, suppose Jones lives in Seattle but his only brother lives in Chicago. One day, Jones's brother is killed in an automobile crash. Instantly, Jones becomes brotherless. If the crash *causes* Jones's suddenly being brotherless, then spatial locality has been violated: no chain of intermediate causes crosses the spatial gap between cause and effect. But this is not supposed to be a case of action at a distance. How can we keep spatial locality from being violated so easily?

We can do so by recognizing that Jones's being brotherless is not a property that Jones possesses *intrinsically*, since the property of having a brother is non-intrinsic (as I mentioned earlier). So Jones's becoming brotherless is not something that occurs at the time of the crash entirely in virtue of what is going on then *where Jones is*. Jones's becoming brotherless is not an *intrinsic* feature of any event occurring at Jones's location at the moment of the crash. Suppose Jones was eating a sandwich at that moment. Imagine him: there he is, becoming brotherless before your eyes. Did you see anything happen to him that was *intrinsically* his becoming brotherless? No! Everything happening at Jones's location could have happened just as it did (as far as its intrinsic properties are concerned) without Jones becoming brotherless. So Jones's becoming brotherless cannot be an event's causally relevant feature. But it would have to be, for the crash to cause Jones's becoming brotherless. So the crash does not *cause* Jones's becoming brotherless, though of course, the crash caused the death of Jones's brother. The event of Jones's turning brotherless consists of whatever is happening to Jones at that moment (say, his biting his sandwich, his heart beating, his moving at 3 m/s). That event is not intrinsically a becoming brotherless; it has causes, but among its causally relevant features is not its involving Jones's turning brotherless.

Thus, spatial locality cannot be undermined so easily. When a 5 kg body moves to a location 5 m away, I instantly acquire the property of being 5 m away from a 5 kg body. Is my acquiring this property *caused* by the body's moving to a certain location 5 m away? If so, spatial locality would be violated. But this is surely not a case of action at a distance! To see why, suppose for the sake of argument that this *is* a case of cause and effect, and then watch how this supposition leads to an impossible result. Under this supposition, the causally relevant features of the effect must be its involving my being 5 m away from a 5 kg body. But this is not an *intrinsic* property of anything happening where I am. Every event (E) happening here (such as my taking a breath) could have happened just as it did, as far as its intrinsic properties are concerned, without the 5 kg body being 5 m away (C). So for any event E happening where I am, *none* of its intrinsic properties is causally relevant to C. But this, we said earlier, is impossible: if C is a cause of E, then C's and E's causally relevant features must include some of C's intrinsic properties and some of E's intrinsic properties, as well as some of C's and E's spatiotemporal relations to each other. For if C causes E, then some of E's intrinsic properties must tie E to C. It cannot be that C would have caused E no matter what E was like intrinsically!

Strictly speaking, nothing *causes* me to acquire the property of being 5 m away from a 5 kg body. But the property's instantiation is not spooky or inexplicable. The explanation is humdrum: that a certain 5 kg body moved to a certain location, while I occupied a certain other location, explains why I acquired the property of being 5 m away from a 5 kg body. If you insist that the 5 kg body's arriving 5 m away from me *causes* me to be 5 m away from a 5 kg body, then I ask you: via what fundamental force – gravity, electromagnetism, the "strong" and "weak" nuclear forces – does the distant body affect me?

The Sun's causing gravitational forces on Earth stands in violation of spatial locality unless the forces have a local cause: the Sun's gravitational field on Earth. In the next chapter, we will begin to look more carefully at fields.[8]

Discussion Questions

You might think about . . .

1 Can you think of some examples – perhaps actual, perhaps imaginary, but in some way different from those given above – in which spatial locality holds but temporal locality is violated? What about where temporal locality holds but spatial locality is violated? Extra credit for cases where spatial locality *and* temporal locality hold but spatiotemporal locality is violated.

2 *In Sartor Resartus*, Thomas Carlyle (Scottish essayist, 1795–1881) wrote:

> It is a mathematical fact that the casting of this pebble from my hand alters the center of gravity of the universe.

Is this a violation of spatial locality? (Interpret "universe" as "solar system.")

3 Consider:

> [A]ny continuous medium . . . , if it consists of material particles, cannot provide an ontological basis for an alternative mode of transmitting actions to the mode of action at a distance. The reason . . . is that every particle with a sharp boundary always acts [in a collision] upon something outside itself, and thus acts where it is not. (Cao 1997: 28)

Is this "acting where it is not" a violation of "spatial locality" as defined in this chapter?

4 Return to the single billiard ball in empty space feeling no forces from $t = 0$ to $t = 5$. Its trajectory from $t > 0$ to $t = 5$ is caused by its feeling no forces during that interval along with its position and velocity at $t = 0$. Its velocity at $t = 0$ is defined as the speed and direction of its motion at that instant. Since speed = (distance covered)/(time taken in covering it), a body's speed at $t = 0$ is usually defined as the "limit," as t approaches 0, of $[r(t) - r(0)]/t$. In other words, take the body's position $r(t)$ at some time t and its position $r(0)$ at time 0. Take the distance between them (which is how far the body moves in a period lasting for t), divide this by t (how long that period lasts), and see what value this quotient converges upon as the period around $t = 0$ shortens. So the body's speed at time 0 is nothing over and above a certain mathematical function of the body's positions at various moments surrounding time 0. Some of these moments occur *after* time 0. So the body's speed at time 0 is *constituted* partly by the body's having various positions at various moments shortly after time 0. But the body's having those positions at those moments are among the events that the body's speed at time 0 was supposed to help cause. The body's having those positions at those moments cannot be one of its own causes! How can we avoid this result?

Notes

1 This broad sense of "cause" is common in philosophy (Lewis 1986: 162). Conversational context influences what counts as *the* cause of (say) a police officer's death: the gunshot, failure to wear a bulletproof vest, or blood loss. Each is "causally relevant."
2 The broad sense of "event" is common in philosophy (Lewis 1986: 216).
3 Similar remarks apply if a fact is a state of affairs – a situation the obtaining of which *makes* a sentence true.
4 See Kline and Matheson (1987), from which I stole my next remark. Another puzzle concerning "action by contact" that I'm not going to explore is that a collision between perfectly hard bodies would result in an instantaneous, finite change in velocity – an infinite acceleration.
5 The body's moving at 5 m/s at $t = 2.5$, and the body's feeling no force between $t = 2.5$ and $t = 3$ (including those boundary moments), presumably constitute a complete cause of the body's moving at 5 m/s at $t = 3$. Of course, we would ordinarily not refer to the body's feeling *no* force as a "cause." We might prefer to say that a body continues to move at a constant speed and in a constant direction because *nothing* causes it to accelerate.

But the body's feeling no force is surely causally relevant, just like the tree's being dry and surrounded by oxygen. What is intuitively a *lack* (the body's feeling no force) *can* be a cause, as when a cause of the barn's burning down was the lack of a prompt response by the fire department.

6 Of course, we would expect a cause of E *not* to begin *after* E has already ended. But I see no reason to build this notion into the definition of "temporal locality." Some philosophers have seriously entertained the possibility of "backward causation" (Price 1996).

7 This is similar (but not identical) to Shoemaker (1969: 376).

8 We will not discuss a host of other interesting questions about causal relations. See Sosa and Tooley (1993).