In part A of *General Relativity from A to B* Geroch not only explains “the space-time viewpoint” but also offers compelling reasons why the natural home of Newtonian physics is Galilean spacetime rather than Newtonian spacetime (or neo-Newtonian rather than Newtonian spacetime). Aristotelian spacetime has geometric structure (straight vertical lines that indicate rest) that plays no role in the physics. The switch to Galilean spacetimes brings the underlying geometry in line with the physics (with the classical principle of relativity).

But Galilean spacetime has its problems. The simplest one is that in Galilean spacetime “the law of light”—the putative law that light travels at $3 \times 10^5$ km/sec—“makes no sense”. Since Einstein accepted the law of light as a basic law in the special theory of relativity, it would seem that one must—if one accepts the spacetime point of view—find a spacetime in which it does make sense. We will not try to justify this virtually incredible law

---

1 Or (approximately) one foot per billionth of a second—one foot per nanosecond.
here (by trying to indicate the theoretical developments and experimental results that led Einstein to accept it).\textsuperscript{2} We will just accept it and say \textit{adieu} to Galilean spacetime.

Ger	extsc{o}ch notes that it looked as if progress had been made when we took structure away from a spacetime. So he advocates the conservative (or radical?) approach that we take away as much structure as possible and then add in whatever seems to be required. So where do we start?

It is as though space-time were drawn on a rubber sheet, which can be stretched, pulled, and bent [but not torn!]. All that we retain, all that we care about for the moment, are the broad, qualitative features of the various points, lines, surfaces, and so on drawn on the sheet. (68)

I will use this paragraph as an excuse to introduce more large words.

\textsuperscript{2} The mu-meson experiment and the observation of binary stars systems that 

Ger	extsc{o}ch describes occurred much later than the acceptance of the special and general theories. Historically then, they had no role in the transition from Newtonian to relativistic physics. It is useful to note that logically, however, there are experiments that discriminate between the two views. It is especially useful to keep this in mind when the two views are said to be \textit{incommensurable}.
First, a few words about functions. A function takes you from its starting point, its \textit{argument}, to some end point, its \textit{value}. For example, the favorite \textit{movie} function will start with a persons as argument and wind up with some movie, their favorite movie, as the value, if they have a favorite movie. If not, then the favorite movie function is undefined for (or \textit{at}) that person.

We typically will want our functions to be reversible, like the Galileo transformations. If we change from un-primed to primed coordinates, we will want to be able to change back again. We probably can’t do this with the same movie function, since it’s likely that many people will consider the same movie--\textit{Casablanca}, perhaps--to be their favorite movie. If you then start with it as argument, there is no one uniquely defined value to go back to. For a function to be reversible (or invertible, or to have an \textit{inverse}), it must have no more than one argument leading to each value. The functions we deal with will typically have one value corresponding to a given argument and only one argument yielding that value. We call such functions one-to-one, or 1-1. (By the way, the inverse of a function $f$ is often indicated by writing $f^{-1}$, a notation that could be confused with raising to a power.)
A 1-1 function can sometimes preserve structure. Suppose, to be a bit abstract for moment, we have two elements in a set related by some operation ‘∘’. Further suppose that for some function $f$,

$$f(a \circ b) = f(a) \circ f(b).$$

Then the function $f$ preserves the structure of that operation. We call a function like this an isomorphism. (The equation above says that you can do the operation to the arguments and then apply the function to get a value or that you can first apply the function to the arguments and then do the operation on its values. The structure in the arguments is mirrored in the structure of the values.)

If the (1-1) function begins and ends in the same set of objects, we call it an automorphism. The Galileo transformations, viewed as active transformations, are automorphisms. The same movie function is not.

Much of what I have said before can be boiled down to these two observations:

(1) geometry is the study of invariants under different classes of automorphisms, and
in general the larger the class of automorphisms, the smaller the class of invariants.

An automorphism that is continuous and has a continuous inverse is called a **homeomorphism**. (Intuitively, a function is continuous if its graph is a solid line.) A homeomorphism can, for example, take some shape as argument and stretch, pull or bend it, since these are continuous deformations. It can’t tear as shape, however, since that would be to create a gap or discontinuity.

The study of structures invariant under homeomorphisms is called topology, the most general branch of geometry. If you now look back at the Geroch quote from page 68, you will see that he is proposing to start out by assuming that our spacetime manifold \( M \) has only topological structure.

There is one more term that it will be convenient to introduce here. A continuous function can have sharp corners. For example, consider the absolute value function at 0. Sometimes we’d like not to have these corners; we would like our transformations to be “smooth”. In that case we can require that an isomorphism and its inverse be differentiable (at each point). We call such an
isomorphism a \textit{diffeomorphism}. As we shall see, when the Leibniz shift is resurrected in the hole argument, the shift is implemented by a diffeomorphism.

Here we end our digression on vocabulary and return to \textit{GRAB}.

Geroch poses two questions. First, what instruments shall we use to probe the structure of spacetime and, second, what shall we assume about these instruments? In the case of Aristotelian spacetime, we plunged ahead in our construction of coordinate systems without asking such questions. The results were intuitively but not physically satisfying.

What we need to build out of our topological base are \textit{metric} concepts like length and duration. For the latter, we need clocks.

We will assume that clocks

1) embody some periodic process--like the swing of a pendulum or the oscillation of a quartz crystal,

2) have a counter (to keep track of the number of periods),
3) can be idealized as a point, so the history of a clock can be represented as a worldline, 4) can be replicated, so we can have as many of these idealized clocks as we need.

Given these conditions, we can represent a clock in our diagrams in the following way:

Representation of a clock in spacetime.
A glance at this picture should convince one that there is another idealization that must be made.

The history of this clock is presented by a curved worldline. That is, the clock may be accelerated during its lifetime. Acceleration can cause clocks to run badly. (Just drop your watch on a hard floor!) We can’t let contingencies like that disrupt or deflect our investigations. So we must suppose that the functioning of our clock is not altered by acceleration. (When this assumption is noticed, it is usually called the clock hypothesis.)

One feature of this diagram is striking if you contrast it with something that Newton assumed:

This very moment when I write this is the same in Rome and in London, on the earth and in the stars, and throughout all the heavens, because that’s the way times relate to space. (De Grav., §10)

Clocks read durations along their worldlines. They tell you nothing about time (or space) anywhere else. The picture above is not a picture of a moment that is the same “throughout all the heavens.” Geroch puts it very straightforwardly:
Note that clocks do not assign times to events not on their world-lines. If we wish to discover “the time” at some other event, we shall either have to run a clock through that event, or else we shall have to somehow carry information from some other clock over to the event in question. (73)

How to we discover what space is like away from our world-line? How do we carry information from one part of space to another? We use light pulses.

What can we assume about our tools? With regard to light, Geroch says this: “We shall suppose that two pulses of light, emitted in the same direction from the same event, move together, no matter what the emitters are doing.” (74) That is, neither the velocity nor the acceleration of the emitter affects the velocity of the light pulse. As a visual reminder, here is figure 26:
What figure 26 shows is the world lines of two persons, A and B, moving with different velocities intersecting at a point p. At p the both emit a light pulse in the same direction. The two pulses then travel together on exactly the same path in spacetime.

It is worth mentioning here that in discussions of relativity it is convenient to assume units such that light travels at one unit of space per one unit of time. In pictures using such units light pulses will travel on lines that make a 45 degree angle to the horizontal (or vertical, for that matter).

Now we turn to our (idealized) clocks. The first question that Geroch asks it this. If two synchronized clocks, A and B, are together at a point p and then travel via different paths to point q, where they meet, should we assume that the two clocks will read the same time when they meet at q? Two put this into readily understood notation, should we assume that

\[ T_A(q) - T_A(p) = T_B(q) - T_B(p) \]

Geroch says that we should not assume this.

I will by-pass the two arguments that Geroch gives. I believe that the first is incorrect and the
second is rather indirect. Fortunately, the conclusion he suggests is directly supported by a famous experiment.³ I will simply quote the abstracts to the two papers reporting the experiment for their results.

**Abstract.** During October 1971, four cesium beam atomic clocks were flown on regularly scheduled commercial jet flights around the world twice, once eastward and once westward, to test Einstein's theory of relativity with macroscopic clocks. From the actual flight paths of each trip, the theory predicts that the flying clocks, compared with reference clocks at the U.S. Naval Observatory, should have lost 40 ± 23 nanoseconds during the eastward trip, and should have gained 275 ± 21 nanoseconds during the westward trip. The observed time differences are presented in the report that follows this one.

**Abstract.** Four cesium beam clocks flown around the world on commercial jet flights during October 1971, once eastward and once westward, recorded directionally

---

dependent time differences which are in good agreement with predictions of conventional relativity theory. Relative to the atomic time scale of the U.S. Naval Observatory, the flying clocks lost $59 \pm 10$ nanoseconds during the eastward trip and gained $273 \pm 7$ nanoseconds during the westward trip, where the errors are the corresponding standard deviations. These results provide an unambiguous empirical resolution of the famous clock "paradox" with macroscopic clocks.

Leaving aside the clock “paradox” (or the “twin paradox”, which is not a paradox at all but just a surprising result), these results provide unambiguous support for the view that the time measured by clocks in moving from one event A to another event B (since all the clocks either remain at rest at or travel around the world from and then back to the U. S. Naval Observatory) depends (in more than one way, in fact) on the spacetime path that they traverse.

This result is by no means obvious or intuitively plausible. Nevertheless, it is a result, and we shall accept it as such. The time indicated by an ideal clock along its worldline is a path-dependent quantity.
On the other hand, we also make the following assumption regarding clocks:

Two clocks [meaning two of our identical, idealized clocks] having identical world-lines between two fixed events measure the same elapsed times between these events. (79)

This assumption ensures that the past history of a clock does not affect its behaviour.

Now we turn to the primary goal of this chapter, the discovery of relations between spacetime points. Let’s do that, starting with the situation illustrated in Figure 38:

The construction which lets us express numerically the relationship between events p and q.
In this case, nothing can travel from p to q (or from q to p) unless it can move faster than light. We will say that p and q are *spacelike separated*.

Now move q a bit to the future.

The construction which lets us express numerically the relationship between events p and q.

\[ s \]

\[ p = r \]

\[ t_1 \]

\[ t_2 = 0, \text{ since } r = p. \]

\[ T(s) - T(p) = t_1 > 0 \]

\[ t_1 t_2 = 0 \]

Clock

light pulses

In this case we say that p and q are *lightlike or null separated*, with q to the future of p because the light pulse travels from p to q.
And we can move q even more to the future.

In this case we say that q is *timelike separated from* p and to the future of p. Since q is to the future of the path of a light pulse emitted from p, we can see than an object, like a clock, could travel from p to q at a speed slower that the speed of a light pulse.

\[
t_1 = T(s) - T(p) > 0
\]
\[
t_2 = T(p) - T(r) < 0
\]
\[
t_1 \cdot t_2 < 0
\]
We could start from the initial drawing and move $q$ to the past to generate two more situations.

In this case $p$ coincides with $s$, which means that a light pulse can be sent from $q$ to $p$. We say that $p$ and $q$ are **lightlike or null separated**.

And finally

\[
\begin{align*}
  t_1 &= T(s) - T(p) = 0 \\
  t_2 &= T(p) - T(r) > 0 \\
  t_1 \cdot t_2 &= 0
\end{align*}
\]
In this case, event $p$ is later than $s$. A clock could travel from $q$ to $p$. We say that $p$ and $q$ are *timelike separated*.

A good way to review (and to think about) these five cases is to start with the case in which $p$ is earlier than $r$, then $p$ coincides with $r$, then $p$ is later than $r$ but earlier than $s$, then $p$ coincides with $s$, then $p$ is later than $s$. 

\[ t_1 = T(s) - T(p) < 0 \]
\[ t_2 = T(p) - T(r) > 0 \]
\[ t_1 \cdot t_2 < 0 \]
At this point, Geroch makes a somewhat opaque remark that is not easy to explain. In fact, a full explanation would lead into many technical details of GTR that we cannot hope to deal with here, but perhaps I can add a little to what he says.

First the remark: “... we shall regard our construction (for \( t_1 \) and \( t_2 \)) as being carried out only in the limit when the events \( p \) and \( q \) are ‘nearby.’” (86-7) Then he explains a bit:

As was the case with points on the earth, this statement by itself will not eliminate the problem. We therefore choose to interpret “nearby” in the same sense as we did for points on the earth [See pp. 85-6.], as follows: The event \( p \) and the clock world-line are to be chosen first, and only thereafter do we have to say how “nearby” \( q \) must be to \( p \). This convention, then, essentially prevents the clock from “wiggling in spacetime between \( r \) and \( p \), or between \( p \) and \( s \), distorting in one way or another the measured elapsed times between the two events.” That is, with this convention certain details of the motion of the clock (namely, that illustrated in fig. 40) are not relevant to the resulting values of \( t_1 \) and \( t_2 \). (87)
To see what he is saying, let’s look at figure 40.

We do not assume--and indeed it is very unlikely that--clock A and clock B will measure the same
duration for $t_2$ (or $t_1$) when they travel from $r$ to $p$ (or $p$ to $s$). But once we are told the curvature of lines $A$ and $B$, we can ensure that the differences falls below any standard we might set by moving $q$ closer to $p$ (and so also moving $r$ and $s$ closer to $p$, leaving less room for the clocks to diverge).

But there is another reason that $p$ and $q$ must be “nearby”. In the special theory of relativity (STR) we assume that $p$ and $q$ can be as far apart as we like and that the interval between them is a frame-independent quantity. But in the general theory of relativity we assume that spacetime looks as if STR is true only locally, only in the neighborhood of any given point. The best analogy is that of a tangent line to a curve. A straight tangent line can approximate a curve for some distance from the tangent point, the distance depending on just how curved the line is.
The spacetime of STR (Minkowski spacetime) *in GTR* is like the tangent line. It is a “flat” 4-dimensional space that approximates the underlying curved 4-dimensional spacetime at each point—but how well it approximates depends upon how curved the spacetime is at that point. So that is another reason why p and q must be nearby. GTR is only *locally* like the spacetime of the special theory. *Globally* GTR can have many different, and sometimes very strange, topological structures.

Geroch doesn’t make clear the differences between STR and GTR. Here we have the basic difference. We shall now go on to develop some insights into spacetime for “nearby” points. That is, we will look at some of the basic features of the STR.⁴

---

⁴ By the way, a popular book that develops the main features of STR using the “the radar method”—reflecting light pulses from objects, noting round-trip times, etc.—is Hermann Bondi’s *Relativity and Common Sense: a New Approach to Einstein*. I believe it was Bondi who invented this way of looking at the STR.