Ten Proofs of Special Relativity

Eric Baird

Einstein's special theory is generally considered to be one of our most reliable pieces of foundation theory, due to the sheer quantity of supporting proofs, arguments and evidence. We consider some of the main pro-SR arguments and find problems with each one, suggesting that the prevailing quality of scientific analysis with respect to SR is lower than is generally realised. We present a more credible argument for special relativity (total energy conservation) and find that this too has problems.

doi: 10.13140/RG.2.2.30051.17441

Table of Contents

1. Introduction

Einstein's 1905 special theory of relativity ("special relativity", "SR" [1]) is one of the major scientific foundation theories of the Twentieth Century, the others arguably being Einstein's subsequent 1916 general theory of relativity ("general relativity", "GR") [2] and quantum mechanics ("QM").

Theoretical physicist and author **Clifford M. Will** has written (1986 [3]) that special relativity is "*correct without a shadow of a doubt*" and that the experimental case for special relativity is so strong that SR should no longer be referred to as a *theory*, but as *fact.*

> Will (1985) [3], page 246: " *Special relativity is so much a part not only of physics but of everyday life, that it is no longer appropriate to view it as the special 'theory' of relativity. It is a fact, as basic to the world as the existence of atoms or the quantum theory of matter.* "

Is this claim justified? We will start with the Will proofs, and then move on to other arguments generally presented as proving the validity of special relativity beyond reasonable doubt.

First, though, we will need to quickly review some history.

1.1. Historical background

Einstein's special theory ("SR") met a genuine need. **Isaac Newton** (~1643-~1727) had used Galileo's principle of relativity to argue that inertial physics operated identically in all systems moving simply in straight lines at constant velocity, and favoured a sort of duality between "particle" and "wave" descriptions of light [5], [6] ... however, an unfortunate error in his "particle-centric" implementation of how light was supposed to behave led to this part of the theory having to be rewritten in the early Nineteenth Century, after which wave theories of light became dominant. [6] Nineteenth century physicists were unable to get Newtonian optics to integrate politely with "wavelike" descriptions of light – wave theory implied the presence of some sort of "aetheric medium" in which the waves could propagate, and a proliferation of theories and hypotheses sprang up concerning how this light-medium might interact with moving matter, and whether the principle of relativity should or shouldn't be entirely correct for light.

The simplest relativistic implementation of these ideas was proposed independently in the C19th by **George Fitzgerald** (1851-1901) and **Hendrik Antoon Lorentz** (1853-1928), [7] and said that although we obtained different physical results if the speed of light was universally fixed for the observer or for the emitter (the predictions diverging by the ratio $1 - v^2/c^2$), if various physical properties scaled with velocity by the square root of this ratio, $\sqrt{1 - v^2/c^2}$, the resulting intermediate physics would be the same regardless of whether the aether was supposed to be stationary in the laboratory, or moving in some arbitrary direction with constant velocity. [8]

After Lorentz published his updated theory in 1904, [9] Einstein responded with a distillation of the relationships (1905), [1] and pointed out that the Lorentz ratio was *mathematically necessary* if we wanted to combine the principle of relativity with global lightspeed constancy – the resulting framework was then mathematically sufficient to define physical behaviour without having to introduce the existence of a physical aetheric medium, or hypothesise how it might interact with light and matter.

While aether models often needed extra variables or coefficients to describe specific assumed interactions of matter, light, and medium, part of the impressive minimalism of Einstein's special theory was that it assumed that the motion of matter had *no effect at all* on the background light-geometry, giving a fixed theory with definite predictions and no free parameters. All we had to do was choose an (arbitrary) simple state of motion as a reference for light-propagation (usually that of the observer), and we could map between different observer's experiences by applying suitable Lorentz redefinitions of apparent distances, times and velocities. **Hermann Minkowski** (1864-1909) then translated the relationships of special relativity into a more abstract description of a flat four-dimensional relativistic spacetime, **Minkowski spacetime** (1909 [10]), which could be analysed and proved to be internally consistent as a *geometry*.

If an inertial observer sees the path of a light-pulse as straight, a different observer who is accelerating or rotating will see the same lightpath as curved, as if the light is being deflected by a gravitational field. An extension of the principle of relativity to include relative acceleration and rotation therefore also needed to be a *theory of gravity*. After special relativity, Einstein started to work on curved-spacetime physics, rediscovering the principle that gravitational differentials shifted the energy of light (previously presented by Michell in 1784), [11] and pointing out that an unavoidable result was that time had to flow more slowly in a more intense gravitational field (gravitational time dilation, [12]).

Einstein's general theory of relativity (1916 [2]) then built on special relativity by saying that since the inhabitants of a laboratory in freefall feel no gravity, we knew that the new gravitational theory had to reduce to more conventional inertial physics over small freefall regions – and since we knew that inertial physics was correctly described by special relativity, the general theory had to contain the physics of the special theory as a limiting case.

By the mid-Twentieth Century, special relativity was solidly embedded in theoretical physics as foundation theory. Thanks to Minkowski spacetime, the SR relationships could be proved *geometrically* to be the only possible solution that combined the principle of relativity with flat spacetime, and for sceptics who argued that flat spacetime was an outdated concept, we had a general theory, that dealt with curvature but also reduced to SR physics as a limiting case.

By the 1960s is was being argued that special relativity *could not* be wrong (although some researchers still believed that they could find paradoxes in the system), and by the 1970s it was being taught that it was compulsory for any new theories to agree with SR. [12] If a system did *not* reduce exactly to special relativity, it was a failure.

We had now accumulated so many results that we *knew* could not be explained without special relativity (when supplemented as necessary by general relativity or quantum mechanics), that there was no point in trying to develop competing systems. Consequently, the body of scientific work that we might normally have expected to see, parameterising relativity theory to see what the results might be if relativity itself was correct but that SR was a *wrong implementation* of it, never seemed to appear. We had no survey of special relativity's place in a wider range of potential competing theories of relativity (demonstrating that SR *really was* superior to the competition), because we believed that no such competing theories were possible.

But belief is not science. The purpose of this paper is to try to review and analyse some of the generally-accepted proofs and more compelling arguments for special relativity, and to try to find which ones *really* demonstrate the unavoidability of Einstein's 1905 theory, and how strongly.

(*Because of the degree of overlap between different arguments supporting special relativity, this will involve a certain amount of unavoidable repetition between sections.*)

2. The Will proofs

2.1. Proof One: E=mc2

Albert Einstein, 1905

The famous equation $E=mc^2$ was one of Einstein's proudest moments. Published in 1905 [13] immediately after the "Electrodynamics" paper that laid out the basis of special relativity, Einstein essentially said that if the previous paper was right, the result of calculating the effective mass of a moving lightcomplex from its momentum would give the now-iconic result, $E=mc^2$. This result explained the known anomalous energy-output of radium (spontaneous nuclear fission), and later, the source of the energy fuelling stars (nuclear fusion).

> Will (1985): " *It is difficult to imagine life without special relativity. Just think of all the phenomena or features of our world with which in which special relativity plays a role.*
>
> *Atomic energy, both the explosive and the controlled kind. The famous equation E=mc^2 tells how mass can be converted into extraordinary amounts of energy.* "

It was natural to interpret the rapid sequential appearance of these two papers as suggesting a *dependency* between the new result and special relativity (Einstein: "*If the theory corresponds to the facts ...* ") [i] and to assume that we knew that SR was correct because otherwise we wouldn't have $E=mc^2$, nuclear reactors wouldn't work, and the sun wouldn't shine.

While Einstein's statements were technically correct, he neglected to mention that if his special theory was *wrong*, and the earlier Newtonian relationships held, the result was *still* $E=mc^2$. [14]

Isaac Newton, Opticks

The concept of mass-energy conversion dates back at least as far as Newton, [15] who suggested the interconvertibility of light and matter, and drew parallels with the transition between ice and water, and between water and steam – perhaps matter was a form of condensed light? This C17th idea has parallels with electromagnetic theories of matter that were current in Einstein's time.

> Isaac Newton, ***Opticks*** (1704): " ***Quest. 30.*** *Are not gross Bodies and Light convertible into one another, and may not Bodies receive much of their Activity from the Particles of Light which enter their Composition? … The changing of Bodies into Light, and Light into Bodies, is very conformable to the Course of Nature, which seems delighted with Transmutations. …* "

If the *idea* predates Einstein, then what about SR's precise conversion factor? If we take the usual derivations of special relativity, and swap out the SR-specific content for older Newtonian relationships, we obtain ... precisely the same result. [ii] Further, if we modify special relativity by a "Lorentzlike" factor,[iii] to produce a smooth range of intermediate theories *between* SR and

i This is "correlation *vs.* causation". Einstein did not try hard to avoid giving the impression of a dependency, $E=mc^2$ paper, 1905: " *The results of the previous investigation lead to a very interesting conclusion …*".

ii It's common to derive $E=mc^2$ by relating the calculated momentum of an energy-complex E to an equivalent amount of matter *m* that would have the same momentum when moving at the same speed. If we do this calculation using the NM relationships and then migrate to SR, SR assigns a greater momentum to the mass for a given nominal velocity (by the Lorentz factor) … but also assigns a greater light-energy (and therefore momentum) to the light-complex, by the same Lorentz factor. With the Newtonian calculation we have $m=p/v$ and $E'/E=(c-v)/c$ … with the SR version we have $p= gamma\ mv$, and $E= gamma\ (c-v)/c$ – the gamma factors for E and *m* cancel, and *E* still equals mc^2, regardless of whether we use NM, SR, or some intermediate relativistic theory.

iii We are defining a "Lorentzlike" factor as the equation $(1 - v^2/c^2)^{\mathbf{exp}}$, where the exponent "**exp**" is a variable. We can then define NM as being the result of a Lorentzlike deviation from SR, or SR as the result of a Lorentzlike

Newtonian mechanics ("NM"), *all* of these will generate $E=mc^2$ as an exact result. [14]

> $E=mc^2$ is a consequence of **any** consistent theory of relativity. Its validation does not show that SR is any better or worse than any other implementation of the principle of relativity, including Newtonian implementations.

Imagining life as it would be is the SR relationships were wrong and the earlier NM counterparts correct ... as far as $E=mc^2$ is concerned ... turns out not to be difficult at all. Life would look exactly as it does now.

2.2. SR Proof Two: Atmospheric muons

According to Will, more evidence is,

> Will (1985): " *Evolution of the species. One possible source of the genetic mutations that permit evolution of living species is cosmic rays. At sea level, the main component of the cosmic rays is the unstable particle known as the mu meson or muon. But the muon is so unstable that it would decay long before reaching sea level from the upper atmosphere where it is created … if it weren't for the time dilation of special relativity, which increases its lifetime as a consequence of its high velocity.* "

This argument again seems confident and convincing. The muon wouldn't reach the ground unless SR time dilation was real, and the muon *does* reach the ground, therefore ... we know that SR must be correct.

But is the statement true? Unfortunately, it's not – if we start with a muon with an agreed energy, momentum and rest mass, and agreed rest-frame decay time t' , we can assign it a nominal velocity according to Newtonian theory using the Newtonian momentum relationship $p=mv$. This gives us a Newtonian *definition* of the muon's velocity, $v_{NM}=p/m$, and a corresponding prediction for the distance d_{NM} travelled before decay.

Using "distance = velocity × time", we get $d_{NM}=v_{NM}t$, or $d_{NM}=tp/m$

In the corresponding SR calculation, the relationship is changed to $p=mv\,\gamma$, where γ (Greek lower-case letter "*gamma*") here expresses a "Lorentz factor" increase, of the ratio $1/\sqrt{1 - v^2/c^2}$. [i] The particle has more momentum under SR than it does under NM for a given nominal velocity, by the Lorentz factor – in times past, we might have chosen to include a Lorentz increase in the mass value and call the effective mass "mγ" ("relativistic mass"). [ii] [16]

Special relativity therefore assigns a smaller velocity value to the same particle than NM, so that v_{SR} is *smaller* than v_{NM}, by the Lorentz ratio (calculated from v_{SR}). Our initial expectation might be that this reduced velocity means that the muon decays *earlier* under SR than under NM, and does not reach the ground. However, thanks to SR time dilation, the distance travelled is *not* $d=v_{SR}t$, but the longer distance $d_{SR}=v_{SR}\,t\,gamma$. Cancelling, we then have $d_{SR}=d_{NM}$.

The expected shortening of the muon's path that we'd otherwise expect when we switch from the

deviation from NM, and by varying the exponent, produce a smooth "fade" between the two systems (and through a continuum of intermediate relativistic solutions). As long as this Lorentzlike correction factor is applied consistently throughout, all solutions will still generate $E=mc^2$.

i The Lorentz factor, "*gamma*" ("γ") is a very special ratio that appears throughout special relativity. It can be found in different contexts written either as $\sqrt{1 - v^2/c^2}$ or as the inverse, $1/\sqrt{1 - v^2/c^2}$ (depending on context: for instance, depending on whether one is talking about a modified signal's decreased frequency or its increased wavelength).

ii However, the idea of relativistic mass can complicate the math for some other situations – it's now generally considered "cleaner" and more proper to use just rest mass throughout, and to keep the Lorentz factor explicit.

NM description to special relativity is precisely compensated for by special relativity's time dilation effect. [i]

> For agreed theory-neutral inputs, a muon's the predicted decay point is exactly the same under both SR and NM.

It would appear that Will may have felt that the obvious superiority of special relativity was so self-evident that there was no point in trying the alternative calculation, without imposing SR-based assumptions. If he had, he'd have realised that for agreed inputs, the predicted decay points are identical under both theories.

Additional arguments for SR

An auxiliary argument might be "*but in the NM calculation, the velocity of the muon has to be greater than the speed of light, which is impossible! Therefore we can rule out the NM calculation and all that remains is the SR version*"

But the "traditional" lightspeed limit is only technically valid in flat spacetime – even under standard theory, if we let ourselves fall into a black hole, we'll normally expect to be travelling inward at $v=c$ when we pass the horizon, and faster beyond that. A more accurate statement is that we are allowed to move faster than the global, averaged, background speed of light, provided that we don't move faster than than our region's local velocity of light in the direction we're moving in – what is forbidden is overtaking or outrunning our own lightsignals, along the same path.

We also have the complication that in *both* the NM and SR calculations, the ultra-high-energy particles are travelling *faster than the speed of light in the air* whose molecules they are passing between. While we could create a mathematical argument that this is impossible, Nature seems to grudgingly allow it anyway, and protests by creating a shockwave (Cherenkov radiation), [17] with the associated energy-loss then slowing a particle, regardless of which of these two systems of physics is correct. This shared phenomenology makes it difficult to be sure that the particles are not – initially at least – travelling at more than background c.

A further auxiliary argument that has been raised is that the NM calculation is wrong, because it does not take into account deceleration effects due to Cherenkov braking, which should eventually slow the muon to less than the background speed of light in air. However, a similar objection applies to the SR calculation, which *also* omits the effect – even though the "SR muon" is always travelling at less than c_{VACUUM}, it's still travelling at more than c_{AIR}, and should therefore undergo a braking effect that slows it eventually to less than c_{AIR}.

This cascading series of further arguments defending the SR version of the scenario do not change the fact that the original claim made for special relativity was mathematically wrong, and should not have been made.

---

i … or, if we do the SR calculation from the point of view of the muon, the muon manages to penetrate the same distance before decay as NM (measured with "Earth rulers") even though its nominal relative speed is lower, because of the Lorentz length-contraction of the Earth's atmosphere.

2.3. SR Proof Three: Particle accelerator limits

The existence of the "lightspeed limit" in particle accelerators is commonly given as an argument for the correctness of SR: the early Lorentz/Einstein concept of "relativistic mass" suggested that a particle's resistance to acceleration tended towards infinity as its velocity tended towards lightspeed.

> *Will: " The US National Budget. In 1983, particle physicists proposed that the United States build a gigantic new particle accelerator. … It would cost around three billion dollars. One reason for the enormous size and cost is the special relativistic increase in the inertia of a particle moving near the speed of light that makes it harder and harder to accelerate it to higher velocities. "*

Since this description seemed to be novel, we assumed that the result was specific to SR (and Lorentzian equivalents), leading to statements that if NM was right rather than SR, our accelerator hardware could (with enough power) give a particle any velocity we liked.

In terms of theoretical physics, this is, unfortunately, quite wrong.

Sticking with the special theory's "raw" equations, and avoiding the concept of relativistic mass, the SR calculation lets us express the Doppler shift on signals from a receding body (Einstein, 1905, §7) [1] as:

$$E'/E = \sqrt{\frac{(c-v)}{(c+v)}}$$

and, for a "passing" body:

$$E'/E = \sqrt{1-v^2/c^2}$$

Both equations give E'/E=zero when v=c. As a result, when we use an accelerator's coils to beam energy at the rear or side of a moving particle, the coupling efficiency of the beam (the energy of the beam as seen by the particle) reduces towards zero as v approaches c. If v *equals* c, the effective energy-transfer is zero, making the standard SR prediction valid even without invoking changes in mass, simply as a result of the theory's Doppler equations. Surely, if we discarded the SR Lorentz time dilation effect and switched to NM, this behaviour would disappear?

It doesn't. If we try the equivalent Newtonian exercise, we get the *even redder* nominal relationships for recession, of:

$$E'/E = (c-v)/c$$

, and for transverse motion, an "**aberration redshift**" (*see* section 5.2) of:

$$E'/E = 1-v^2/c^2$$

Once again, the coupling efficiency drops to zero as v tends to c, and it takes an infinite amount of energy to get the particle up to the speed of light (at least, by direct acceleration).

> The SR particle accelerator lightspeed limit (for direct acceleration) exists regardless of whether we use the SR or NM equations.

Of these three reasons to believe that special relativity is correct "*without a shadow of a doubt*", *all three* are based on an apparent ignorance of previous theory's predictions and a lack of basic mathematical checks that would have shown that the arguments weren't true.

This is somewhat troubling.

2.4. Quantum electrodynamics?

Will also gives fourth proof, "*Chemistry, the basis of life itself*", which argues that atomic structure depends on Quantum Electrodynamics ("QED"), a "welding together" of SR and quantum mechanics.

We will not attempt to provide a full replacement theory of quantum electrodynamics in this paper. However, it is worth noting that in some cases where older classical theories predict an effect and an SR-based approach doesn't, quantum mechanics will have an analogue of the same non-SR effect, which can then be *retrofitted* to an SR-based classical physics as a separate QM *correction*. In other words, if SR was in some ways an *inferior* classical theory, and failed to predict some classical behaviours that QM insists must exist, we could *correct* SR by adding these missing behaviours as separate quantum effects to bring the predictions back into line with reality.

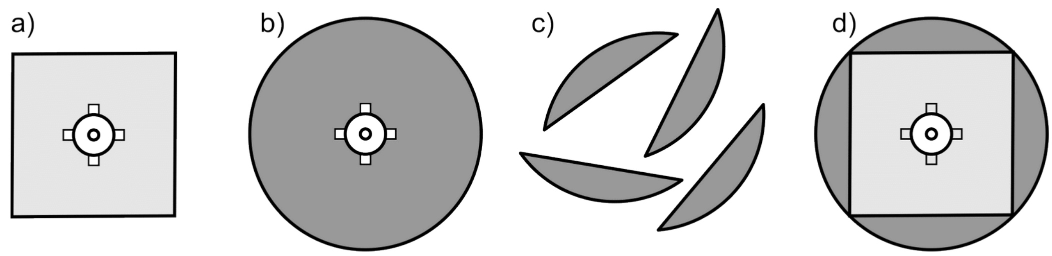

*Figure 1: **The Square Wheel.** The square wheel in (a) represents the predictions of special relativity, and (b) the predictions of quantum mechanics after smoothing (the desired end result). Subtracting one from the other gives (c), "missing" effects that we then define as being non-classical and QM-specific. Adding these "QM effects" to special relativity gives (d), a working wheel. The success of the final system does not depend on the fact that we started out with a square. If we had started with a hexagon, we would merely have a **different definition** of which effects we believed to be QM-specific, and the wheel would have a differently-shaped tyre, but with the same circular outline. If we had started with a circular wheel, with statistical outcomes **already equivalent** to the QM predictions (a **QM-dual** classical theory), then no additional QM corrections would have been necessary. The success of (d) is not a proof that all wheels must be square.*

We can see this QM "retrofitting" in action in the case of light-dragging (section 22) and also for Hawking radiation. [20], [21] In a Newtonian-based system of physics, gravitational horizons are "leaky", fluctuating and weakly radiating massenergy and information (Thorne 1994, [22] pages 252 and 443), in a manner that seems suspiciously similar to the result of QM statistics. [i]

This behaviour is impossible in systems of physics based on the SR equations, but can be "retrofitted" by overlaying the missing effect onto a classical background as a separate "QM" correction. But adding additional QM behaviours onto SR and achieving wonderful results does not necessarily mean that it is special relativity that deserves the credit – in an ideal system, classical and quantum effects would be dual, counterparts of the "additional" QM effects would already be present in our classical model, and no "welding" would be be necessary.

---

i Compare Thorne (1994) [22] figure 6.8, page 252 (visiting particles in the region around a Newtonian dark star), with figure 12.3, page 443 (Hawking radiation appearing to originate in the vicinity of the horizon, as seen by a suspended observer, according to QM). Both show particles appearing to be restricted to a region outside the horizon, but able to be knocked free by interactions with passing masses or each other (for a distant observer, physically accelerating a particle converts it from a virtual particle to a real particle).
The Newtonian indirect radiation effect is a statistical, *system behaviour* effect, in that it does not show up in analyses based on plotting the trajectory of a single test particle – in order for one particle to be accelerated free from the field, it requires two or more particles to interact and exchange momentum in the larger field. The existence of *physical* differences between the two descriptions are not obvious, and the Newtonian description now technically counts as a case of classical Hawking radiation. This behaviour cannot be replicated in a system of physics based on the equations of special relativity. [23]

3. What went wrong?

Of Will's four arguments, three are easily shown to be bad using simple “high-school” mathematics, and the fourth is somewhat “hand-wavy”, and resistive of easy analysis. Given that the three examples that we *can* check all turn out to be bad, we can be forgiven for being reluctant to “take someone’s word for it” on the fourth. This does not mean that simple Nineteenth-Century Newtonian theory wasn’t wrong and/or troublesome for many *other* reasons … it definitely was … but if three of these four proofs are destroyed by a simple comparison of special relativity against previous mainstream theory(/ies), then they cannot be taken seriously as scientific arguments.

This represents a departure from the normal scientific standards expected in other fields of physics. We would expect these arguments to have been checked.

It is difficult to explain this failure rate. If special relativity *really is* as good as we think it is, and there are many *good* and entirely *genuine* reasons to believe that the theory is right, then it would seem that Will was desperately unlucky to have picked three arguments to “showcase” that failed a basic inspection.

Although Will’s book is aimed at a “popular” audience and is not peer-reviewed, its author is one of the world’s most eminent authorities on the more complex subject of testing *general* relativity, and one might also have expected that somewhere along the line, that he might have encountered *somebody* who knew special relativity well enough to be able to point out these issues (perhaps even a publisher’s proofreader). [i]

Analysis

It would seem that the prevailing community standards when it comes to analysing the case for special relativity might be a little more “relaxed” and not quite as scientific as we have been led to believe.

Perhaps we don’t really *care* whether an individual argument is valid, because we think that it doesn’t matter … since we *already know* that the theory is correct, mistakes in our supporting arguments are – in terms of baseball, a case of “*no harm, no foul*”. Even if, unbelievably, *three major proofs are all wrong*, well, we can reassure ourselves that this changes nothing, as we have many more independent mathematical and experimental results at our disposal that cannot be explained without special relativity, and they can’t *all* be wrong.

Or can they? If we were to try to compile a listing of the “top ten” reasons why we knew that SR was correct, and the first three that we investigated turned out to be bad, then should we bet that these three failures are all vanishingly-unlikely exceptions, and that all the others others are still valid physics? Or should we wonder if the three “promoted” failures are symptomatic of the community’s more general tolerance of “careless” science as long as it supports current orthodoxy? Is the failure suggestive of deeper problems with the relativity community’s attitudes to fact-checking and rigorous analysis, forming part of some larger pattern of systemic failure?

That would be an outrageous suggestion, and quite unthinkable. Let us now do some quick additional checks to reassure ourselves.

---

i The book, published in 1985, has since had a second edition in 1993, [4] and is still in current and in print after a third of a century. An updated sequel to the book is apparently due to be published in 2020.

4. SR Proof Four: Relativistic aberration

It is sometimes argued that the SR relativistic aberration formula (as presented in §7 of Einstein's 1905 paper) must be unique to special relativity, because **(a)** it's in the 1905 paper and therefore "belongs" to special relativity, **(b)** because it's referred to as "relativistic", and **(c)** because one would normally (quite understandably) expect different theories, with different physical predictions, to express themselves as different geometries.

4.1. The special case of 90°

Suppose that we stand in a room, and throw a ball (or aim a laser) at a mirror fixed to the wall directly in front of us. If we aim the ball or laser-pulse directly at the mirror, it should return exactly to its starting-point. [i]

As seen from an overhead watcher, looking down, we can define the mirror surface to be parallel to the *x*-axis, and the path of the ball or pulse to be parallel to the *y*-axis.

If we now switch to the viewpoint of a different observer moving parallel to the mirror (*i.e.* along the *x*-axis), both we and the ball or pulse appear to be moving *x*-wards, at precisely the same rate. Since the ball or pulse now has a motion along the *x*-axis, it cannot now be aligned with the *y* axis – the path has to now be described as being deflected, forwards, in the same direction that the room is seen to be moving.

1

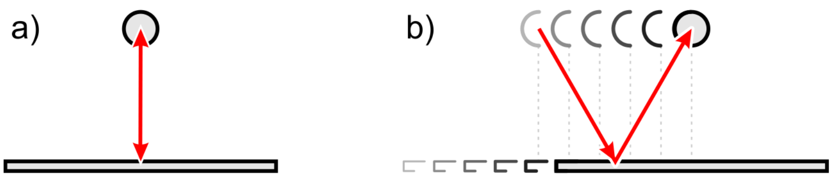

Figure 2: ***Bouncing a ball against a wall***

This "forward deflection" prediction is not unique to special relativity – the prediction is precisely the same, for instance, under Newtonian theory.

4.2. The full relativistic transverse angle-change formula

Now imagine that we replace the entire room with a single enclosing spherical mirror. We set up an apparatus in the centre to emit a single unidirectional outward pulse of light, which is mostly blocked by a surrounding shroud. Holes in the shroud allow rays to pass through at specific angles, they bounce off the spherical mirror, return through the same holes and reconverge perfectly at the centre, where the reconvergence triggers our machine to emit a further pulse. The machine continues like this, emitting pulse after pulse, indefinitely.

From the point of view of the external otherly-moving onlooker, the rays must still be reconverging at a single event in spacetime (otherwise the machine wouldn't continue firing), but a pulse's emission event $\boldsymbol{e}_1$ and its reconvergence event $\boldsymbol{e}_2$ now occur at different locations in space, the apparatus having moved while signal was in flight.

This gives us all we need to know to calculate the angle-change for any ray, regardless of whether the relativistic model in use is special relativity, Newtonian theory, or something else.

---

i … ignoring the downward gravitational deflection. To eliminate this, we can suppress the *z*-axis and make the experiment two-dimensional (*x*-, *y*-axes only), or we can have the room floating in deep space.

4.3. The relativistic ellipse

The condition that the rays must converge on a point at the same moment (for all observers) as an *event* provides us with a **relativistic angle-change diagram**, in which the round-trip light-distance e_1 ⟶ **reflector** ⟶ e_2 is the same in every direction. The shape that collects signals from one focus and reconverges them at another is an elongated ellipse, [i] and the ellipse geometry gives us the necessary relativistic angle-change formula. Notice that we have not distinguished between light and little thrown balls (Nineteenth-Century ballistic emission theory) – our result holds under special relativity (Einstein 1905, §7 [1]), under Newtonian theory, and under any other relativistic system.

Geometry then gives us the "relativistic ellipse" diagrams [25] of Figure 3. [ii]

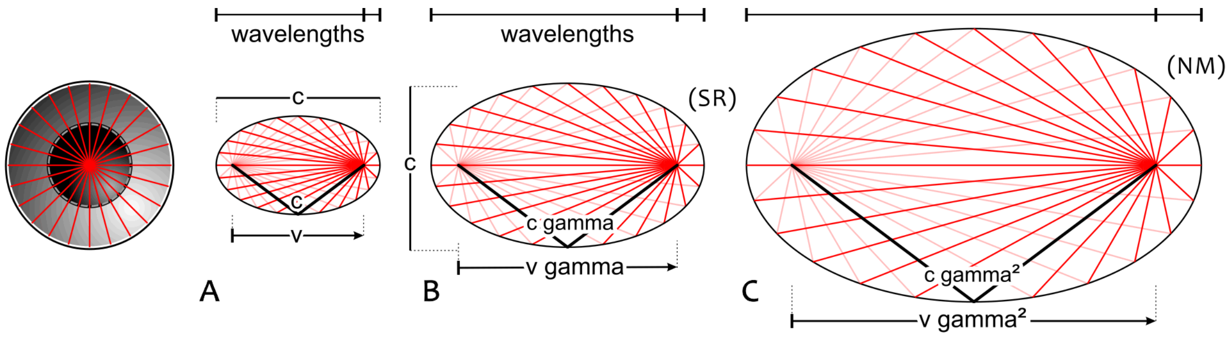

*Figure 3: **Relativistic Ellipses:** "stationary" sphere, and relativistic ellipses for v=0.8c, showing angles and wavelengths for **A**, an unspecified theory, **B**, special relativity, and **C**, Newtonian theory.*

> Relativistic aberration is a general result of the principle of relativity, and is not exclusive to special relativity.

4.4. The relativistic ellipse under special relativity

Special relativity "owns" the simplest implementation of a relativistic ellipse: the elongated shape of **B** has a constant width (constant minor diameter), and can be packed back into the original outline with a simple Lorentz contraction along the *x*-axis (Moreau 1994 [24]). This contraction can be visualised in four dimensions as the result of tilting the ellipse diagram out of the plane of the paper (tilted "plane of simultaneity" [24]), until its "shadow" produces the original circular outline. The ellipse can also be considered as a cross-section (or a shadow of a cross-section) though a **Minkowski light-cone**, created by intersecting the cone with a moving system's (tilted) plane of simultaneity, to obtain a conic section (which is, of course, an elongated ellipse).

The lengths of the individual rays can also be considered as wavelengths, and when we calculate their lengths in ellipse **B**, they do turn out to exactly correspond to the Doppler predictions of special relativity. The **B** ellipse therefore usefully expresses all the critical intersecting relationships of special relativity (for a given velocity) in a single geometrical figure. [24] [iii]

---

i Applying a Lorentz contraction to this elongated shape then turns it back into a sphere.

ii This diagram is not what one might expect from Einstein's 1905 description, which implies that a viewer sees the shape compacting *from* a sphere to a compressed ellipsoid (*see* "Gamow", section 43.3). The "contracted sphere" description is not to be taken literally – it has the wrong wavelengths, it has the wrong angles, and it was dismissed by peer review as a geometrical misunderstanding in the late 1950s and early 1960s (*see* Terrell 1959, [26] 1989, [27] Penrose 1959, [28] Weinstein 1960, [29] and Boas 1960 [30]).

iii Anyone unconvinced by the elongated ellipsoid geometry is invited to try plotting a diagram of the aberrated ray-angles for special relativity for a velocity $v<c$, and then extending and contracting the ray-lengths according to the SR Doppler predictions. They will find themselves recreating the **B** version of the ellipse. NM generates **C**.

4.5. Ellipse scaling for multiple relativistic theories

The ellipse angles and proportions come from the requirement that our rays need to originate at one event and converge at another, relativistic theories can only disagree about the wavelength values. If we know a relativistic theory's forward and rearward wavelengths for a given nominal velocity, the ellipse lets us reconstruct its predictions for all other angles. This gives us a way of calculating the expected Newtonian "transverse" redshift: starting with forward and rearward wavelengths that are redder than their SR counterparts by a Lorentz factor, the entire ellipse magnifies by a Lorentz factor compared to SR, giving a Lorentz-squared transverse effect.

For comparative relativity theory, the ellipse's magnification is the only free parameter:

- If we declare that the round-trip cycle time for a pulse is one second, that the equipment moves at v m/s, and that the distance between two foci is therefore v metres, we would obtain diagram **A**, whose length is a constant but whose volume *shrinks* as speed increases (associating negative curvature with positive kinetic energy, a bad result).
- If we declare that spacetime must be "flat" (no curvature associated with kinetic energy), the ellipse *width* must remain constant regardless of velocity. We magnify all **A** distances by a Lorentz factor, giving diagram **B**, Minkowski spacetime, and special relativity.
  The direct interfocal distance $\boldsymbol{e_1} \longrightarrow \boldsymbol{e_2}$ travelled during one cycle is then not v, but a longer distance, scaled by the Lorentz factor. This is traditionally explained by saying that, if we treat the pulsing apparatus as a "clock", the increased distance that light has to travel when the apparatus is moving causes it to "tick" more slowly (Lorentz time dilation), so that it travels a greater distance between "ticks". While **A** distances are literal ("naive"), **B** introduces the principle that a theory is allowed to redefine distances and times.
- If we were to assume that positive kinetic energy is associated with *positive* curvature, we would require a larger ellipse than SR, to express the excess of distance packed into the region as a function of velocity …
  - … with the maximally-curved solution then being the ellipse that we get by constructing the ellipse around the Newtonian wavelength distances, **C**. In order to return this shape to the original outline, we'd have to extrude the surface out of the plane to from a sort of "tilted gravity-well" shape.

  The different versions of an ellipse all have *identical proportions* for the same nominal velocity – they are always longer than they are wide, by the Lorentz factor.

These three different sets of predictions (and those of an infinite number of intermediate relativistic solutions) are able to associate different physics with the same aberration angles, because the shape of the region containing a moving system changes differently with relative velocity for each solution.

In general, we can see that spacetime is always flat in **B**, in the range to the right of **B** (towards **C**) spacetime becomes more positively curved with increasing kinetic energy, and in the range to the left (towards **A**), spacetime becomes more *negatively* curved as a consequence of kinetic energy. All the wavelengths of the **B** ellipse are increased compared those of **A** by a Lorentz redshift, and those of extremal solution **C** are increased with respect to those of **B**, by a further Lorentz redshift.

At this point we have solved the puzzle of why Nineteenth-Century theorists were unable to produce a wavelike description of Newtonian optics: the Newtonian predictions do not work geometrically in flat spacetime.

4.6. Lorentzlike factors, and the relativistic continuum

Since any relativistic theory has to produce exactly these angles, this exercise shows that we can define a *continuous spectrum* of potential theories of relativity, and uniquely identify any individual solution by a known reference theory and a single number − its relative ellipse magnification factor.

If energy has *no* associated curvature effects, we have **B** and special relativity, if recoverable kinetic energy produces positive curvature we have a magnified ellipse and redder, longer wavelengths in the range **B**-**C** (with **B** excluded). The properties of ellipse geometry allow the relative magnification factor between theories to be expressed as a **Lorentzlike factor,**

$$E'/E = (1 - v^2/c^2)^{exp}$$

, where the value of the undefined exponent, "**exp**", is different for each relativistic solution.

As we move between different relativistic theories' predictions for a given nominal velocity, the entire ellipse diagram magnifies or reduces to match.

4.7. Summary

The aberration/relativistic ellipse exercise is extremely productive, as it not only lets us express all the key relationships of special relativity for a given velocity in a single diagram (Moreau [24]), it also gives us the exact relationships for *any other possible relativistic theory*, which is invaluable information for testing purposes. Once we know the divergence of a different relativistic theory's non-transverse Doppler prediction from SR, we know that theory's predictions for any other angle, providing a basis for checking which hypothetical theory of relativity is the most accurate.

As well as letting us conceptualise the possibility of non-SR relativistic theories, and showing exactly how the different relativistic solutions interrelate, it puts extreme constraints on the relativistic possibilities, and lets us catalogue the spectrum of potential relativistic models and their varying predictions according to a single number. [i]

> The SR aberration formula is not specific to SR. All possible relativistic theories predict the same change in angles for a given nominal velocity (although there may be some theory-dependent disagreement about how we define that velocity value).

> Different relativistic theories diverge, not by having different aberration formulae, but by overlaying the same formula onto backgrounds with different degrees of velocity-dependent curvature.

> Of these, special relativity is *provably* the only possible relativistic solution for flat spacetime. The ellipse exercise tells us that if the principle of relativity is correct, and spacetime is perfectly flat, then special relativity's relationships are geometrically unavoidable. We can rewrite SR's postulates as being "relativity, plus flat spacetime".

> However, it is not unreasonable to suggest that perhaps the energy tied up in the relative motion of masses may have associated curvature, in which case we also need to consider solutions that are redder than SR by an additional Lorentzlike factor, in the range **B** to **C**.

i Other critical interlinked relationships, such as the relationship between the momentum and velocity of a mass, also have to vary across the spectrum of theories in the same way by a corresponding Lorentzlike factor.

5. SR Proof Five: Transverse redshifts

5.1. The "unique" SR transverse redshift

The "transverse redshift" argument for special relativity is simple: it says that the effect is "***purely relativistic**, for there is no transverse Doppler effect in classical physics at all*" (Resnick 1968 [31]). Under SR, if we aim a detector at 90° to a straight set of railway tracks and a train passes by, the train's redshifted signal will make the train seem time-dilated: since time-dilation is considered to be a "new" effect introduced by SR, [i] the existence of the redshift verifies special relativity. Surely, if it wasn't for special relativity, the signal wouldn't be redshifted?

Here are some more typical statements, in date order:

> Ray d'Inverno, **Introducing Einstein's Relativity** (1992) [32] " … *the transverse Doppler shift ... This is a purely relativistic effect due to the time dilation of the moving source.* "

> Richard A. Mould, **Basic Relativity**, (1994) [33] " *If a source is observed from a direction perpendicular to its motion, the resulting change in frequency is called a transverse Doppler effect. This is a relativistic effect, for classically one would not expect a frequency shift from a source that moves by right angles.* "

> Wolfgang Rindler, **Relativity … second edition** (2006) [34] " *Time dilation is the **only** cause of the frequency shift whenever there is no radial motion between source and observer. This is the so-called **transverse Doppler effect**, and has long been considered as a possible basis for time dilation experiments. …* "

The "upshot" of all these statements seems to be pretty clear: transverse redshifts don't appear without special relativity, they can be interpreted as demonstrating the existence of the SR time dilation effect, the effect is unique to SR (and Lorentzian electrodynamics), and since no other theory predicts similar effects, a test that shows that these redshifts are real makes for a solid validation of special relativity. [ii]

This all seems very convincing. Unfortunately, this convenient argument (like our earlier convenient arguments) cannot be reconciled with geometry, math, or logic. If it *was* true, it would also invalidate special relativity (section 5.3).

5.2. Aberration redshifts in C19th Newtonian theory

Suppose that under the (very bad) ballistic emission theory of light commonly used to represent Newtonian optics in the Nineteenth Century, a train moving along a straight track aims a beam of light at (what seems *to it* to be) 90° to the track. To trackside observers, this "thrown" light must be seen to be advancing along the track at the same speed at the train: its path, or "ray" must be seen to be tilted forwards (see section 4.1). If the trackside observer now aims a telescope at exactly 90° to the track (as measured with a trackside set-square), the "transverse" signal that enters the optics to be analysed when the train goes past will not belong to the ray that was *aimed* at 90°: the angular aberration effect will cause the scope to intercept and evaluate a *different* ray, which (according to the emitting body) was originally aimed to point a little more rearward (and has ended up in our "transverse" detector by being deflected a little forward). This ray is therefore contaminated by a small recession redshift component (Lodge 1909, "*a spurious or*

---

i … and, arguably, "proto-SR" theories such as Lorentz aether theory.

ii Since "true transverse" experiments are quite sensitive to small angular errors, we often test SR by looking for the existence of a transverse redshift "time dilation" component in longitudinal Doppler shifts.

apparent Doppler effect due to common aberration." *i* [35]). When the strength of the Newtonian "aberration redshift" component is calculated, [36] it turns out to be a Lorentz-squared redshift, *stronger* than the SR prediction, [37] and in precise agreement with the results of the ellipse geometry shown in section 4.5. The transverse-aimed detector should therefore report a redshift regardless of whether we are doing NM or SR calculations.

> Textbook descriptions of how SR's predictions compare to those of older theories are often somewhat "creative" with facts, and are not to be taken too literally (or too seriously).

> The belief that a predicted redshift in this situation is unique to SR seems to be based more on "guessing", on convenience, and on faith that someone, somewhere will have carried out an actual geometrical analysis of the situation. This seems not to have happened.

5.3. Special relativity requires the non-SR redshift to be real

The usual response from SR experts at this point is to say that the expected broad agreement between the SR and pre-SR effects is pure coincidence, that the two are *obviously* quite different effects with different causes, and that there is no deeper theoretical link between the "old" C19th effect and the "new" SR version.

Unfortunately, this isn't true, either: special relativity allows us to make precisely the same final physical predictions by declaring the speed of light to be fixed *in any legitimate inertial frame*. If we should choose to assume that light propagates preferentially in the *observer's* frame, then we expect no aberration redshift, and a Lorentz time-dilation redshift. But if we choose to assume (as we are entitled to) that light instead propagates preferentially in the *emitter's* frame, any propagation effect must now be multiplied by a Lorentz *blueshift* (since *we* are now said to be moving and *our* clocks are considered time-dilated) ... which means that in order to arrive at the same final physical prediction as before (a single Lorentz redshift), the propagation effect associated with lightspeed propagating at *c* with respect to the emitter ***must* be a Lorentz-squared redshift**, in agreement with the earlier referenced calculations.

If the "moving observer" pre-SR transverse redshift due to propagation effects were *not* exactly a Lorentz-squared redshift, special relativity's logical structure would fail.

5.4. Interchangeability of Newtonian and SR transverse redshifts

Further, since the same recorded *physical* SR transverse redshift can be described under SR as either being due to time dilation of the moving body, or as being due to simple signal-propagation effects (partly cancelled by the time dilation of the observer), we cannot *on principle* tell the two classes of effect apart within SR, except by their magnitude, with the pre-SR effect being necessarily stronger than the SR version.

If there was any way to tell (qualitatively) whether a detected redshift was due to time dilation or due to propagation effects, then we would be able to identify a preferred frame for light propagation, breaking the principle of relativity.

> Within SR, there MUST be a propagation-based redshift at $90°_{\text{OBSERVER}}$ if the speed of light is fixed in a different frame to the observer's, and there is *not allowed* to be any qualitative way to tell the two effects apart.

---

i It is difficult to find any mention of aberration redshifts in the literature between 1909 (Lodge [35]) and the mid-1990s.

5.5. Transverse redshift effects under aether theories

Since an absolute aether stationary with respect to the observer gives no transverse redshift, and an aether moving with a moving source generates a Lorentz-squared aberration redshift (equivalent to the "emission theory" result), the range of "transverse" predictions for aether theories is between "no effect" and a Lorentz-squared redshift. Similar arguments apply to more complex aether models (such as "dragged" or "aerodynamic" aether theories): when the motion of the emitter has *any effect at all on* on light-propagation, even if it's only at close range, we expect to see a corresponding "transverse" redshift.

5.6. Invoking theory-dependent definitions of "transverseness"

In all these calculations we are assuming that the "transverse" signal that is being considered is the signal that is "*received at 90 degrees in the observer's frame, according to the observer.*"

A counter-argument is that the term "transverse redshift" *originates* with special relativity, *belongs* to special relativity (like a trademark), and should not be used to refer to the analogous predicted effects that appear under other theories ("*transverse effects are unique to SR by definition*"). Other theories must then use different language (*e.g.* "aberration redshift"). With this argument, though, there is no longer any physical content to a statement that transverse redshifts are unique to SR, it becomes purely a matter of naming conventions. A naming convention cannot be used as the basis of an experimental testing programme.

Another counter-argument is that while the shift predictions are correct, we must not *refer* to them as transverse, because angles should be defined within the frame in which lightspeed is fixed (making the redshift effect disappear when we select a different ray). But this definition is useless for experimental testing. We need some way to decide where to aim our detector, and then we need different theories to tell us what would be seen in that given situation. We cannot compare SR's predictions against other predictions that need the detector to be pointed somewhere else. We can also object that under SR this choice of propagation frame is arbitrary, and that dragged-light theories don't have a propagation frame.

5.7. Transverse components in non-transverse Doppler shifts

We often test for a transverse redshift *components* in non-transverse signals: we take the received signal, divide out an assumed propagation shift of $E'/E = c/(c+v)$, and if SR is correct, end up with a residual Lorentz redshift that we say can only be explained with SR and time dilation.

But with the Newtonian relationship $(c-v)/c$, this exercise gives $E'/E = (c-v)/c \; / \; c/(c+v) = (c^2+v^2)/c^2 = 1 - v^2/c^2$, so the corresponding Newtonian prediction is again a Lorentz-squared redshift.

5.8. Summary

Special relativity's "transverse" redshift predictions are not qualitatively novel or unique, as commonly presented: they are right in the middle of the range of physical predictions made by Nineteenth-Century theories, with Newtonian theory actually predicting a stronger effect for a given nominal velocity.

In experimental scenarios where SR predicts a transverse redshift, most other theories will also predict a redshift, regardless of what we choose to call it.

Textbook characterisations of other theories' predictions are not to be considered trustworthy.

6. SR Proof Six: particle accelerator time dilation (straight path)

Some physics professionals are fond of saying that we know for a fact that high-speed particle decay times and lengths in accelerator complexes have been experimentally shown to be Lorentz-dilated (proving special relativity), even when the particles are travelling in straight lines at constant speed.

This cannot be correct: part of the point of special relativity's solution is that it allows us to make the same physical predictions in flat spacetime regardless of which frame we use as a reference for global lightspeed constancy: Even assuming that the SR predictions are *numerically* correct, we would not on principle be able to determine experimentally whether this outcome was due to:

(a) the speed of light being fixed in our frame and the particles being time dilated,

(b) the speed of light being fixed in the particle frame, and the *lab* being time dilated with respect to the particles (and also length-contracted), or

(c) the speed of light being fixed in an *intermediate* frame, with the lab and muons ageing at exactly the same rate. [37] , or,

(d) the speed of light being fixed in some other SR-legal frame, with an appropriate application of Lorentz effects

We are not allowed under "core" special relativity [i] to say whether, after assumed propagation effects are taken into account, we are "really" seeing a residual time dilation redshift, a residual blueshift, or no shift at all. [ii]

... A proof would be a disproof

If we *really could demonstrate* experimentally that particles moving through the lab in a straight line at constant speed were *unambiguously* time-dilated, then this would demonstrate that it was the particles that were "really" moving rather than the laboratory – we'd have identified a preferred frame for global lightspeed propagation, disproving the principle of relativity [iii]).

When pressed on this, particle accelerator professional will usually be forced to back down and admit that no such unambiguous experimental result is known to exist, but they will then usually reply that this doesn't matter, because we know that particles *do* unambiguously age more slowly in particle storage rings. We'll now examine this much more complicated case ...

---

i It is common for people using special relativity to say that we know the speed of light is fixed for the observer. This can imply (wrongly) that there might be something special about that frame. SR actually *forbids* us from knowing which frame lightspeed is supposed to be globally fixed in: we are free to assume global *c* in any SR-legal frame, and after we apply Lorentz correction effects, the final physical outcome will always be precisely the same. We can say that it's quite valid to *assume* that the speed of light is fixed in the observer frame when we do our calculations, but it can also be assumed with equal validity to be fixed for any other SR-legal reference system.

ii Referring back to the muon case in section 2.2, special relativity does not let tell whether the muon is "really" time dilated with respect to the Earth, or *vice versa*: if we say (under SR) that the muon is subluminal and that the speed of light is constant for the *Earth*, the "SR" muon's ability to penetrate as far as it does is blamed on time dilation. But from the point of view of the muon, we can argue that the speed of light appears fixed for the muon, that the Earth is moving, and is time-dilated and length-contracted, and that the muon manages to pass through more atmosphere than expected because the depth of the Earth's atmosphere is Lorentz-contracted.
Although these arguments can seem slippery and contrived, this is arguably the whole point of special relativity: it is a system that allows us to declare that the speed of light is globally fixed to any legal frame, and still predict precisely the same final physical outcome.

iii Textbook writers tend not to make this mistake: Taylor and Wheeler (1992) [38]: Box 3.4, "Does a moving clock **really** 'run slow'?", *"Does something about a clock really change when it moves, resulting in the observed change in tick rate? Absolutely not!* "

7. SR Proof Seven: Time dilation in particle storage rings

Confronted with the slightly metaphysical nature of SR's time dilation effects for simple rectilinear motion, the particle accelerator professional will call one's attention to the real, verifiable, and *clearly unambiguous* time dilation effect that is seen in particle storage rings.

These storage rings are essentially circular sections of particle accelerator-like hardware, whose purpose is to keep high-velocity charged particles contained. The *length* of a storage ring in Earth-coordinates is unambiguous (because we can slowly pace around the perimeter with measuring-wheel to find its circumference), the *speed* of the particles in Earth coordinates is *also* unambiguous (because we can inject a pulse of particles, and count how many times this pulse passes our position per second), and the *time dilation* of these circling particles is also a physically "real" effect (because we can send a particle-pulse around the ring and measure how many circuits are achieved before the particle decays).

Putting high-speed particles into a storage ring is like putting fresh food into a refrigerator: it makes them last longer.

7.1. Mach-Einstein principle

Unfortunately, this agreed time dilation of circling particles doesn't amount to a "clean" physical proof of special relativity, because the circling particles are no longer moving inertially in a straight line with constant velocity ("core" assumptions for SR), and once particles are undergoing "gee-forces", other explanations are available:

According to **Ernst Mach** (1838-1916), and Albert Einstein (~1920 [39], 1921 [40], the general principle of relativity says that when a particle circles a rotation axis, the apparent outward gee-forces that it feels are to be regarded as a *real* gravitational field (for the particle).

> Einstein, "*Ideas and Methods...*": *circa* ~1920- [39] " *Next we introduce a second co-ordinate system **K'** that uniformly rotates relative to **K**; we symbolize this systems as a circular disk uniformly rotating relative to **K** (see fig 3) ...*
> *According to the theory of general relativity, we can also view the* [rotating] *system **K'** as 'at rest.' However, then we have to perceive the field of centrifugal forces existing relative to **K'** as a (real) gravitational field that acts upon all masses that are at rest relative to **K'** in proportion to their masses. (For the sake of completeness, we have to add that this gravitational field does not only consist of this (gravitational) centrifugal field, but it also has other components that act upon moving masses.* " [i]

If we stand at the centre and rotate *with* the circling particle, the particle appears to be stationary but surrounded by a hollow shell of environmental matter (stars, galaxies) spinning around the axis – we see this rotating shell to be associated with what appears to be an "unconventional" gravitational field pointing *away* from the rotation axis whose strength increases with distance from the axis. In this description, the particle is artificially suspended in this field by the lab hardware, and is only prevented from flying outward (away from the central rotation axis), by the containment fields generated by the ring's coils.

By treating this apparent gravitational field as real, we can argue that the suspended, accelerated particle is actually undergoing a gravitational time dilation and redshift, an assumption that explains the reduced ageing rate of the circling particle, correct to the available experimental accuracy (Harwell group, 1960 [41], [42]).

---

i The "other components" include a velocity-dependent dragging component – a gravitomagnetic effect that pulls light in the direction of the moving matter, and appears to conflict with special relativity.

7.2. Until ~1960: Dual explanations

At this point, the subject starts to become somewhat controversial.

Until around 1960, Einstein's "gravitational" explanation of physical time dilation for a circling mass was considered legitimate, and in February 1960 a group at the Atomic Energy Research Establishment at Harwell, England published their measurements of centrifuge time dilations, as measurements of "effective" gravitational shifts. [42] Invoking the principle of equivalence let the group use the huge gee-forces produced by a high-speed centrifuge rather than having to rely on puny Earth gravity to produce a redshift [i] (as managed by the Harvard group [43], [44], [45] in the US).

The "SR" and "gravitational" explanations were assumed to be "dual" – if an inertial observer stood at the centre of the centrifuge and watched the perimeter, the redshift would be entirely due to the relative speed (transverse redshift, section 5) and not gravitation. If we rotated *with* the centrifuge, we should see the same redshift, but there would now be no relative motion within the apparatus to blame it on, but a very considerable apparent gravitational field operating across the centrifuge body, which could then be held responsible for the energy-loss in the signals.

The two observers would have two very different explanations for *why* there was a redshift in the apparatus, but their final physical predictions for the outcome should agree.

7.3. Post-1960, Competing explanations, GR rotation paradox

After the publication of the 1960 Harwell paper, the physics community was thrown into a state of crisis when it began to be appreciated that the flat-spacetime prediction (SR) and the curved-spacetime prediction (principle of equivalence) were *not* necessarily dual, and might even be competitors.

If the spacetime curvature that existed across the region in the rotating frame ("gravitational explanation") was *intrinsic* curvature, it would have to exist for *both* the inertial *and* noninertial observers, threatening to make the SR flat-spacetime explanation redundant.

- If the curvature effect existed for both classes of observer, but the SR effect *only* existed for the *non*-rotating observer, then *this* observer would see roughly twice the effect seen by their colleague (the curvature effect, *plus* the SR effect). This result would not be viable. [ii]
- On the other hand, we could try to reintegrate the two descriptions by associating the relative motion of all masses with curvature (using a gravitomagnetic theory). This would allow the central rotating observer to blame the redshift on curvature due to relative *acceleration*, while a co-rotating observer could blame it on gravitomagnetic curvature due relative *velocity*. While the gravitomagnetic approach would solve the rotation paradox, it would also require a redesign of inertial physics to associate the relative motion of masses with curvature – the result would no longer be a flat-spacetime theory, and would therefore no longer agree with special relativity.

The preservation of SR therefore requires a rejection of some basic aspect(s) of general relativity.

---

i Hay, Schiffer, Cranshaw and Egelstaff (1960): [42] " *Einstein's principle of equivalence states that a gravitational field is locally indistinguishable from an accelerated system. … The shift of the gamma-ray energy in the effective gravitational field of a rotating system …* "

ii The co-rotating and non-rotating central observers have to agree as to the amount of redshift seen in light arriving at their location. Since an identical "strong" redshift can be obtained by making the centrifuge rotation arbitrarily fast (and the radius small), or the rotation arbitrarily slow (and the radius large), the apparatus' central physics (effects of relative rotation of the masses of the two types of central observer) doesn't seem to be relevant.

7.4. The 1960 breakdown of Einstein's general theory

If the principle of equivalence and the general principle of relativity, applied to rotation, appeared to invalidate special relativity, [iii] then – since Einstein's 1916 general theory had assumed that all three things were fundamentally correct, this also meant the loss of Einstein's 1916 theory. We could still have *a* general theory, just not Einstein's, or one that reduced to SR. While a gravitomagnetic replacement for Einstein's 1916 theory might be considered desirable in the long term, in the short term we would have simultaneously lost *both* our mainstream theories of relativity, without having an obvious replacement theory waiting to take over.

As recounted by Schild (1960 [46]), the idea that the equivalence principle invalidated SR was unacceptable.

> Schild (1960): [46] " *The group at Harwell has also measured the red-shift produced in rotating disks. The question arises whether, by virtue of the equivalence principle, such effects in accelerated systems are to be regarded as verifications of general relativity. There seems to be some confusion on this point and even some lack of unanimity among theoretical physicists. It is one of the purposes of this note to clear up this question. The confusion is unnecessary, because within the framework of the theory of relativity the answer is simple and definite. It is "no!"* "

Schild's arguments are slightly odd.

Firstly he starts off with an argument that the experimental evidence supporting SR and flat Minkowski spacetime is "*well established*", as opposed to the general theory, "*where the empirical evidence is slight and perhaps inconclusive.*" The issue of whether a particular result should or should not be considered to be a result of a fundamental principle ("*Should result X be considered evidence for principle Y?*"), should really try to be based on pure logic and geometry, rather than "looking over our shoulder" at the physical evidence to try to guess which result we would prefer to be right.

Schild next says that "*The special theory of relativity completely describes and predicts the effects which one would expect to see in an accelerating system*" (so that by proving the principle of equivalence, we were proving SR). Unfortunately, it doesn't, not if our "expectations" are based on the GPoR. According to the principle of equivalence and the general principle of relativity, SR *cannot* completely describe the effects of relatively accelerated masses, due to its failure to take into account the physical spacetime curvature associated with masses that experience gee-forces.

We are *not compelled* to assume the total validity of SR when applying the principle of equivalence, and if the two turn out to be in conflict, we actually *mustn't.* By using an SR-centric approach to gravity rather than a GPoR-based approach, Schild is "assuming SR to prove SR".

Acceleration and the stars

Schild goes on to say that we know that the principle of equivalence is wrong: if someone is in an elevator in deep space, suspended in a real gravitational field, the region's geometry is curved, while if the elevator is being physically accelerated, and modelled by SR, the region *appears* curved in that they see lightbeams appearing to bend, but since all the beams bend by the same amount, the underlying geometry is still deemed to be "flat", with the lightbeams all still seen to be travelling along straight lines, if we select an appropriate observer.

---

iii Schild 1960: " *It is important to formulate this question very carefully because special relativity and the equivalence principle do not form a consistent theoretical system. …* "

Suppose that we start with a large and effectively "flat" region of spacetime dotted with stars and galaxies (an "effectively flat" piece of universe). Now imagine that we are in a spaceship in that region, firing its engines to produce a relative acceleration between us and this environment. Within the spaceship, we feel an apparent gravitational field, pushing us back into our seats, and if we look out of the spaceship window we see further evidence that the field is real, because we see stars and galaxies all apparently free-fall-accelerating in the same direction, due to the same field. The universe appears to be permeated by a uniform gravitational field, and the only reason we are not falling with it is because we are firing our engines in an attempt to remain stationary.

According to the argument invoked by Schild, since everything in the universe is freefalling, we can eliminate the field by a suitable choice of (inertial) observer, and reveal its inherent underlying flatness – since this doesn't work for a "real" gravitational field, the two situations are not equivalent.

Schild's idea that we can distinguish between "real" and "apparent" gravitational fields seems to be reasonably widespread (e.g. Møller 1955):

> Møller (1955), [47] VIII:83, page 221 " *It is thus quite natural to assume that both types of gravitational fields are of the same nature and obey the same fundamental laws. This assumption is often referred to as the **principle of equivalence**. It is true that the gravitational fields due to the distant masses can be made to disappear by a suitable choice of the system of reference, viz. by choosing a system of inertia as system of reference, while the gravitational fields arising from 'close' masses such as that of the earth or the sun cannot be 'transformed away' by a proper choice of the system of reference; the latter fields will therefore be referred to as **permanent gravitational fields**.* " [i]

But **this argument is not correct according to the general principle of relativity**. The region is *not* intrinsically flat, because all bodies are *not* freefalling in the same direction at the same rate. *One* body is having differently – the spaceship occupied by the observer! If the principle of equivalence argues that acceleration effects are purely relative, and that apparent lightbeam curvature results from the relative acceleration of masses, then just as the relative acceleration of the background stars to the spaceship causes the spaceship to feel gravitational side-effects, the principle of mutuality (or Wheeler's "democratic principle" [48]) requires that the relative acceleration of the spaceship's mass to the starfield must similarly cause the background observers to feel gravitational side-effects, too (emanating from the location of the spaceship). The background observers feel the forced acceleration of the spaceship to be causing a physical distortion of the region's spacetime around the accelerating ship's structure (Einstein 1921 [40]), giving a warping of the lightbeam geometry that cannot be removed by a convenient choice of observer.

The distortion (around the ship) due to the relative acceleration of its mass is geometrically real, and can be considered as the difference in the states of motion of the two types of system, expressed as physical geometry, and smoothed across the intervening regions of space as a *geometrical transition*. Although this distortion might only operate significantly-strongly over a region that seems insignificantly small to us in terms of astronomy (encouraging us to classify the larger expanse as "effectively flat"), everything that the spaceship occupants do to interact with their environment (and everything the environment onlookers do to interact with the ship) involves signals passing directly through this distorted region, and being affected by it. The distortion is part of how a moving mass interacts with its neighbours, and its environment.

---

i … in other words, all apparent gravitational fields are equivalent, but some are more equivalent then others. This is not correct use of the general principle of relativity. Where the observer has *mass*, we require full equivalence.

Rotation and the stars

A similar common mistake involves special relativity applied to rotation. In an SR-centric argument we might start by saying that signals passing through a region travel in straight lines according to the background stars. We then place a rotating disc in the region, and within the disc's frame of reference, the same lightbeams appear to be curved. Matter loosely attached to the disc (and attempting to travel in an inertial straight line with respect to the background starfield) finds itself pulled outward from the centre by an apparent gravitational field and also experiences a sideways pull in the direction that the nearest stars appear to be travelling in, giving radial and Coriolis gravitational fields (Einstein, 1921 [40]).

In the SR-centric description we say: these apparent gravitational fields can be completely modelled using special relativity … but they are not *real* gravitational fields because we can map back from the rotating disc's experience to the original reference with respect to the background stars and find that all the lightbeams are actually straight.

General relativity disagrees. The "logical disjunct" here is the unwarranted assumption that the light-beams that we initially defined as being straight *before* we introduced the rotating disc, *remain* straight when we introduce a lump of rotating matter – and according to general relativity this is not what happens.

Topology

A geometrical theory allows us to apply topological transforms without changing the inherent physics, or the outcome of an experiment. If we start with a rotating hollow environmental shell facing inwards at a central star, a topological transform allows us to turn the situation "inside-out", so that the shell becomes a central sphere of mass facing outward, and the star becomes a surrounding surface facing inward. If the rotating external shell generates gravitomagnetic dragging effects on the contained body, a rotating contained body (such as our rotating disc) must also be associated with gravitomagnetic dragging effects on what surrounds it. [i]

Under a general theory of relativity, adding a rotating central mass physically changes the region's lightbeam geometry, causing the region to be warped, even for an inertial onlooker.

No reduction to the "extended SR" description

There is no choice of frame in which this distortion disappears. In the frame of the rotating body, lightbeams are curved due to the rotating shell of background stars, for the background stars, lightbeams are curved in the vicinity of the body by dragging effects, and in intermediate frames, the external shell and the contained body, rotating in different directions, both contribute to the "twisting up" of the intervening region of spacetime. [ii]

> According to the general principle of relativity, the presence of rotating or forcibly accelerated matter experiencing gee-forces physically warps spacetime – a "flat spacetime" SR-based analysis will not give exact correct answers.

i A geometrical theory of relativity and gravity that *didn't* support topological transformations would not be credible.

ii This faulty pro-SR argument might be classified as an example of **process blindness**, where we insist on continuing to use an initial temporary definition even after subsequent additional processes should have modified or invalidated it.
An example might be:
"Teacher*: I have two apples, and then I buy another three apples. How many apples do I have in total?*"
"Problem Student*: You have two apples. You told us this at the start of the problem. The part about you buying another three apples is therefore a lie. Two apples! Two apples!*"

7.5. The failure of GR1916 as a principle-based theory

Before 1960, general relativity was seen as a fully-geometrical "principle" theory [50] with exact and non-negotiable results, whose structure and principles were inviolate, and whose predictions were rock-solid. This encouraged Karl Popper to present Einstein's general theory as a prototypical example of a *scientifically falsifiable* theory, [51] in that nothing in it could be fudged or fiddled, and that if any parts of the structure refused to fit, the whole theory was invalidated (1919 [50]). i

Post-1960, Einstein's general theory could be said to have been *logically* invalidated due to a failure to agree with *itself* (incompatibility of the SR component with the principle of equivalence and the GPoR) and the community's earlier enthusiastic approval of Popper and the principle of scientific falsifiability became more guarded. ii

The 1960 episode appears to mark the loss of Einstein's general theory as a principle-based system, and its replacement with a looser, more *ad-hoc* system of rules and approximations, in which concepts previously regarded as foundational principles were allowed to be overridden in order to protect special relativity.

> The incompatibility between SR and the general principle of relativity (discovered in 1960) makes Einstein's general theory *structurally pathological.*

> A valid general theory of relativity cannot incorporate special relativity.

> In the case of the rotating Earth, the physical existence of these rotational dragging effects were confirmed experimentally by the Gravity Probe B experiment, in 2004-2005. [48], [49]

7.6. Aftermath: The SR clock hypothesis

In order to continue claiming that centrifuge time dilation *was* entirely in accordance with special relativity, the community invoked the **SR clock hypothesis**, an idea that had been used by Laue in around ~1913 [52] in an attempt to allow SR-based analysis of problems involving acceleration.

The essence of the hypothesis is that when we model the physical changes in clockrate of an accelerating body, these changes are purely a function of the sequence of instantaneous velocities that the body experiences, and the results of those velocities according to special relativity – it assumes that there is are no additional effects or distortions due to the acceleration itself.

> Møller (1955) [47] page 49: " *This equation is now assumed to be valid also for an arbitrarily moving clock where u is the momentary velocity of the clock. Hence we assume that the acceleration of the clock relative to an inertial system has no influence on the rate of the clock, and that the increase in the proper time of the clock at any time is the same as that of the standard clocks in the rest system S0), i.e. the system in which the clock is momentarily at rest.* "

This is arguably a reasonable assumption within "extended SR", as special relativity doesn't attempt to apply the principle of relativity to acceleration to find out *whether or not* there are any additional effects that arise due to acceleration – that's the job of general relativity.

---

i Einstein 1919, writing for the London Times: "*The great attraction of the theory is its logical consistency. If any deduction from it should prove untenable, it must be given up. A modification of it seems impossible without destruction of the whole.*"

ii "*The code is more what you call guidelines than actual rules. Welcome aboard the Black Pearl, Miss Turner!*" – Hector Barbossa, **Pirates of the Caribbean: The Curse of the Black Pearl** (2003)

At the time that Laue's book was first published, we didn't yet have a general theory of relativity. The problem with invoking the hypothesis in the 1950s and 1960s was that by this time we *did* have a general theory, and by then, Einstein had already pointed out that as a result of this theory, the "physical" acceleration of a mass should be associated with physical distortions in spacetime that would then also be identifiable for other, non-accelerating observers.

> Einstein (1921): [40] " *A rotating hollow body must generate inside of itself a 'Coriolis field' which deflects moving bodies in the sense of the rotation, and a radial centrifugal field as well. ...*
> *A material particle, moving perpendicularly to the axis of rotation inside a rotating hollow body, is deflected in the sense of the rotation (Coriolis field)*"

, and then we could use the principle of mutuality to argue that if the rotation of a massed shell creates a dragging effects on bodies inside it, then the rotation of bodies inside a non-rotating shell must create dragging effects on the shell. Spin a wheel, and nearby masses should be pulled around with the wheel by a gravitomagnetic effect. Also,

> Einstein (1921): " *A body must experience an accelerating force when neighbouring masses are accelerated, and, in fact, the force must be in the same direction as that acceleration. …*
> *There is an inductive action of accelerated masses, of the same sign, upon the test body.* "

Fire a pellet from a rubber slingshot, and nearby matter should experience a gravitational "tug" in the direction of the pellet's forced acceleration.

If the relative acceleration and rotation of masses *physically deforms spacetime*, producing intrinsic curvature, then the geometry is not the same as that of a a perfect SR test object changing speed against a perfectly flat, undisturbed, "Minkowski" background. If we are lucky, the SR-based description might be good enough for engineering purposes, but it will not be technically exact under a general theory of relativity, and can't be used as foundation theory.

> The SR clock hypothesis – that acceleration and rotation has no effect on lightbeam geometry – is incompatible with the general principle of relativity.
>
> Since the clock hypothesis is necessary to prevent the GPoR disproving SR, special relativity is still geometrically incompatible with the general principle of relativity.

7.7. Inconsistencies

This subject is still controversial: it is possible to find some sources arguing that *of course* it is perfectly legal to apply SR to acceleration, and others arguing that of course it isn't, and that we must use "full GR" for an exact answer. Proponents of both points of view argue that there is no controversy: their own position is provably correct, and "the other team" have misunderstood basic theoretical principles.

These disagreements are the result of the 1916 theory's pathological structure, which allows both groups to *prove* that their preferred interpretation is the correct one, as a pathological system lets us prove the mutually exclusive outcomes "**A**" and "**Not A**", depending on which part of the structure we start from. If we start from SR-centric arguments, we can prove that parts of GR1916 must be *in perfect agreement* with SR, while if we start from the general principle of relativity, we can prove that the same parts must *disagree* with special relativity.

Gravitational textbooks have a habit of fudging the issue: according to MTW, section 6.1, [53]

> Misner, Thorne and Wheeler ("MTW"), **Gravitation** (1973), §6.1 "Accelerated observers can be analyzed using Special Relativity" page 164: " *When spacetime is flat, move however one will, special relativity can handle the job* "

This is a little like someone asking whether it is safe to jump out of an aeroplane at altitude without a parachute and being reassured, "*Oh yes, it's perfectly safe ... as long as there's no danger*". MTW's qualification "*when spacetime is flat*" makes the answer almost meaningless.

Where there is a physically accelerated mass under general relativity, ***the region is never flat***. [i] We may as well say that "*special relativity can handle the job ...*", "*... as long as the 'observer' has zero mass and is purely mathematical*", or, "*... as long as we do not need the answer to be credible, reliable, describing physical reality, or compatible with the general principle of relativity*".

7.8. The clock hypothesis applied to accelerator storage rings

Defending SR by invoking the clock hypothesis for storage rings amounts to a declaration that since, within SR, the measured centrifuge effect *must* be explicable using SR, we *cannot allow* there to be any complicating effects due to acceleration. Since flat-spacetime SR makes such a good match to the data, we "*know for an experimental fact*" that there are no significant additional effects due to spacetime distortions.

Although this sounds convincing, we can also run the argument in reverse. In a universe in which the principle of equivalence and Einstein's general principle of relativity apply, we can explain the centrifuge outcome using purely gravitational principles for both inertial and noninertial observers, and can then say that since *these* calculations make such a great match to the data that we "*know for an experimental fact*" that there are no significant additional detectable effects due to *special relativity.*

> If we start by assuming the validity of SR, the SR clock hypothesis is correct, and SR is vindicated. If we start by assuming the validity of the general principle of relativity, the clock hypothesis is wrong, and so is the special theory.

7.9. Testing the SR clock hypothesis

If the SR clock hypothesis is correct (and the entire time-dilation effect in a muon storage ring *really is* due to relative speed rather than acceleration), then muons with a fixed speed should decay at precisely the same rate regardless of whether their track is straight or curved. This makes the clock hypothesis *in theory* physically testable. [54], [55]

Suppose that we take a circular muon storage ring of perimeter *c* metres (figure 4), and populate it with hypothetical particles with a rest-frame decay time of slightly over one second, which would only normally be expected to make at most a single circuit before decaying. Suppose also that in practice, these particles actually manage to make *ten* circuits before decaying when moving at a sufficiently high speed. Since the speed and the acceleration both scale up by the same amount when we change the ring's radius or the particle speed, [ii] how do we analyse the data to show

---

i In fact, according to the principle of equivalence, the region ceases to be flat as soon as we add a physical mass capable of acting as an observer, and becomes less flat if the observer-mass moves, even without acceleration.

ii If we double the speed of a circling particle, we also double its acceleration: if we double the radius of the ring but keep the number of circuits per second constant, the speed and acceleration both double. This lets us calculate the same time-dilation effect in a circular storage ring by blaming it *either* on particle speed or acceleration.

whether it is "really" the particles' speed or their acceleration that is responsible for the effect?

Separating the effects of speed from acceleration requires us to change the geometry of the particle's path.

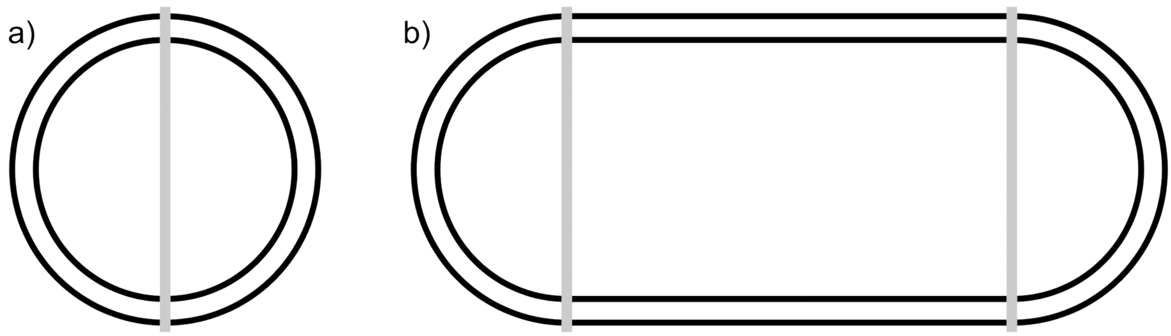

Figure 4: ***Particle accelerator racetrack***

In theory we can establish which explanation is the correct one by cutting the ring in two and inserting two additional straight sections of length "one half *c*", so that the total perimeter distance is now 2*c*.

- If particle decay times and distances are identical in the straight and curved sections, then the muons will complete around five circuits of the doubled-length track before decaying.
- If physical time dilation only happens in the curved sections, then the muons will decay before completing a single circuit. [i]

Enquiries of the particle accelerator community as to whether anyone is considering this experiment have prompted different responses, ranging from "*We don't need to test the SR clock hypothesis because we already know it to be a fact*", to "*We do not need to test the SR clock hypothesis because it is explicitly only SR-specific (hence the name), and a full analysis of the problem requires full GR*". [ii]

7.10. Gravity from velocity

If we attach a line of clocks to a radius of a disc, and spin the disc, the "centrifuge" behaviour should result in the faster-moving clocks at the perimeter physically ageing more slowly.

In the 1905 paper, Einstein used this to make one of the less successful predictions of special relativity:

> Einstein, "… *Electrodynamics* ..." (1905) §4: " ... *Thence we conclude that a balance-clock at the* [Earth's] *equator must go more slowly, by a very small amount,* [iii] *than a precisely similar clock situated at one of the poles under otherwise identical conditions.* "

The modern view (supported by measurements of transported atomic clocks) is that these two clocks will actually "tick" at *an identical* rate. [56], [57] This is not due to anything wrong with Einstein's basic argument, but is because the 1905 description didn't take into account the compensating gravitational consequences of a physical variation in clockrate.

---

i A potential complication here is the question of whether individual charged particles suspended against a gravitational field by an *electric* field can still be said to experience gee-forces.

ii … in other words, we do not need to test the theorem, because we do not particularly expect it to be correct, and consider it to be an "engineering" extrapolation of SR without a deeper significance.

iii According to Alley (1979), ~102 nanoseconds per day.

The equitemporal geoid

According to Einstein's argument, clocks distributed around the Earth's surface ought to tick more slowly towards the equator: this would mean that we would have a *temporal gradient* operating across the Earth's surface, and just as Einstein's 1911 paper [12] said that a variation in clockrate between locations should cause lightbending due to Huygens' principle, deflecting light to the region of slowest lightspeed, a similar effect should deflect light (and matter) towards the equator.

A ball placed on a smooth frictionless track at the pole would then run "downhill" towards the equator, and rocks and dirt would migrate away from the poles until the resulting heap at the equator was piled sufficiently high for the reduced gravitation there (due to the height of the pile) to exactly cancel out the first effect. Ignoring complicating effects such as atmospheric variations and ocean currents being obstructed by continental masses, and tides, sea level (and to a lesser extent, ground level) reaches gravitational equilibrium when all points on the surface have *exactly the same* clockrate. The idealised ocean surface is an **equitemporal geoid**.

- **From the point of view of the background stars**, the Earth has an equatorial bulge because its rotation is throwing mass outward from the centre, and the bulge is caused by the circling material's *inertial* mass, which is trying (unsuccessfully) to continue travelling in a straight line.
- **From the point of view of the Earth**, we can say that the rotating shell of background stars creates an outward-pointing gravitational field that is responsible for lifting the Earth's equatorial crust to form the bulge, which is caused by the material's *gravitational* mass.

Peer-reviewed descriptions tend to say that if we calculate the SR effect and then calculate the GR gravitational effect separately, they cancel. This is unnecessarily complicated – starting from the 1905 prediction, the correct answer ("no effect") – courtesy of Huygens principle and a little thinking – doesn't require any mathematical calculations at all.

7.11. "Extended" special relativity is not a valid subset of general relativity

The idea of a physical distinction between "natural" gravitational fields and the effects of relative rotation or acceleration (Schild/Møller/etc.), is not just contrary to the general principle of relativity, and gravitational theory, it is also now (thanks to Gravity Probe B) contradicted by the available physical evidence.

If we define a region as containing straight lightbeams, insert a rotating mass, map how the straight beams appear in the rotating frame to be curved by an apparent gravitational field ***and then stop***, then we we have the Schild/Møller position that the field is not real, and can be made to disappear by reverting to the initial nonrotating frame of reference. But a theorist trying to derive a general theory of relativity, will then *iterate*. They will take the initial exercise as a provisional demonstration that relative rotation appears to be associated with a field, and they will then apply the general principle of relativity to argue that … since we do not believe in absolute space, and we do not believe that the matter making up the background stars has different properties to matter here on Earth … that the relative rotation of matter must *also* be seen to be associated with a field by an inertial bystander. The lightbeam geometry defined in the first stage can only be regarded as a provisional "first approximation" geometry, giving us a description of the broad phenomenology that the GPoR *must then support for all observers.*

"Extended special relativity" is a faltering step *towards* some of the effects of general relativity, but its geometry is "disposable" – it does not live on within an actual general theory. The distinction between "real" and "apparent" fields exists temporarily in an attempted extension of special relativity, as we head *towards* a general theory, but the distinction is supposed to disappear once a full general theory has been achieved. The necessity of this disappearance gives the "new" physical effects described by Einstein in 1921 [40] (section 7.6). If we persist in using the Extended SR arguments, we are doing a form of "**SR Plus**", not GR.

> If distortions due to rotating masses were "fictitious", and *really could* be made to disappear by referring our description to the frame of the background stars, then there would be no rotational dragging effect under GR. Since the rotational effect is now experimentally confirmed by Gravity Probe B, the Schild argument defending SR and downgrading the GPoR appears to be at odds with the currently available experimental evidence.

7.12. Summary: Storage rings

The problem of how special relativity does or doesn't apply to particle storage rings is quite a complex problem

> Not only does the situation with particle storage-rings *not prove* special relativity, it sets in motion a chain of logic that appears to destroy both of Einstein's theories, special and general.

Analysis of rotating-body problems seems to have revealed (back in 1960 [46]) a previously unnoticed fundamental geometrical incompatibility between special relativity and both the principle of equivalence and the 1916 general principle of relativity, preventing SR from being able to coexist with the PoE and the GPoR as exact solutions within a single logical framework.

Since Einstein's general theory had been clearly defined as supporting both the GPoR/PoE *and* special relativity – a combination that we now appreciate to be geometrically impossible to implement – Einstein's 1916 theory, as originally presented, has to be considered invalidated on the grounds of logical inconsistency.

According to (Schild 1960 [46]), we cannot construct a logical system that incorporates both special relativity and the general principle of relativity as exact solutions, and (since SR cannot be wrong), we therefore need to downgrade the GPoR.

A legitimate alternative conclusion might be that since the GPoR is based on fundamental logic and geometry (and the principle of equivalence of inertial and gravitational mass), while the case for SR is arguably more based on the *convenience* of overlooking curvature/dragging effects in order to assume flat spacetime and simplify the problem of inertial physics, perhaps we should be putting principles ahead of convenience, and exploring the consequences of giving the GPoR priority, and downgrading SR.

The science community is supposed to be open and honest about potential problems in major theories, so it is awkward to think that Einstein's special theory could be refuted by the GPoR and the PoE, and Einstein's general theory could be found to be structurally pathological, without the physics and math communities widely publicising the news, and encouraging debate. Instead, we have one paper stating that a policy-change is being made, and no apparent discussion. The 1960 episode remains obscure, and is apparently not often discussed in polite company. [i]

---

i A community "policy decision" to downgrade the general principle and principle of equivalence would require some form of peer-reviewed document setting out a statement of position, which can then be referred to by

8. SR Proof Eight: The finite speed of light

8.1. History

Some theorists have characterised the main technical difference between Newtonian and Einsteinian physics as being that Newton thought that the speed of light was infinite whereas Einstein in 1905 more correctly treated it as finite. This narrative makes Newtonian mechanics a low-velocity approximation of special relativity, based on an incorrect and naive idealisation by Newton, which SR then corrects. It is then argued that since we know that lightspeed ***is*** finite, this means that we know that the SR description is the correct one [i].

According to this narrative, Newton and his contemporaries did not appreciate that the speed of light was not unlimited, and failed to take "finite-*c*" into account in their models, with special relativity's big advance being that it was more in agreement with how physics really worked in the real world.

Although the argument sounds convincing, it is historically and mathematically wrong, and is contradicted by a basic reading of Newton's *Opticks*. By the time of the book's publication, observations of the eclipses of the moons of Jupiter (Roemer) had already demonstrated that we saw different offsets in the timing of the eclipses that related to how far away Jupiter was when the observations occurred, and Newton not only documented the effect, he quoted a decent resulting value for the speed of light, as a multiple of the speed of sound.

> Newton, *Opticks*, " **Prop. XI.** ***Light is propagated from luminous Bodies in time, and spends about seven or eight Minutes of an Hour in passing from the Sun to the Earth.***
>
> *This was observed first by Roemer, and then by others, by means of the Eclipses of the Satellites of Jupiter. For these Eclipses, when the Earth is between the Sun and Jupiter, happen about seven or eight Minutes sooner than they ought to do by the Tables, and when the Earth is beyond the Sun they happen about seven or eight Minutes later than they ought to do; the reason being, that the Light of the Satellites has farther to go in the latter case than in the former by the Diameter of the Earth's Orbit. "*

Newton's descriptions of light bending under the influence of gravitation, due to *proportional changes* in lightspeed caused by a "gravitational" variation in density of the underlying medium (*Opticks*, query 30 [15]) would also not make sense if he believed that the speed of light was infinite.

What about the speed of gravity?

> *Bertschinger & Taylor (2017):* [58] *" Without quite saying so, Newton assumed that gravitational interaction propagates instantaneously. "*

By "*without quite saying so*", the authors would seem to mean, "*we have been unable to find any quotes that support our assertion*". In Newton's aether model, the deflection of light by gravity was a local effect caused by the local aether density-variation. It did not require any form of "spooky" instantaneous interaction or communication with a distant gravitational source. In this sense, the speed of gravitational interaction was no more "infinite" than it is under general relativity, where

subsequent peer-reviewers. This seems to be the function of the first part of the Schild paper (the second part being more personal theory). The paper's first part is strange in that it seems to represent a discontinuous and emphatic change in how peer-review was to deal with general relativity ("*It used to be believed that X but from now on we must say Y instead*") – a sort of "*palace coup*" – but attracted almost no further peer-reviewed discussion. The community members that had taken part in previous discussion were not identified, and the arguments that led to the conclusion that SR and the GPoR were incompatible were not presented for analysis. This is not usual scientific behaviour.

i The "Einstein good, Newton bad" characterisation also assumes that these are the only two options.

light responds locally to the curvature variations that it encounters along its path. When it comes to the speed at which *gravitational changes* are communicated, this would depend on the rate at which *density changes* are propagated though the aether. Although we can have aether models in which *persistent flows* can be constant, it is difficult to imagine an aether model in which *acceleration-related* density-variations (the equivalent of gravitational waves under general relativity) propagate infinitely quickly.

Newton probably *didn't* attempt to model gravitational waves, for the simple reason that any calculation would have to depend on other parts of the theory that were not yet entirely solid, and that in any case the technology required to measure gravitational waves was not exactly within the reach of contemporary physicists (who had only recently come to terms with the idea that light's colour was associated with wavelength). This absence does not mean that Newton *disbelieved* in gravitational waves, or chose not to model them due to a belief that c_g was infinite.

While these arguments are again founded on bad physics and bad history, it is worth briefly looking at how lightspeed issues *really* affect relativity theory.

8.2. Finite c requires finite c_g

If we chose to believe that the speed of light was finite but **the speed of gravity, c_g, was infinite**, then if multiple gravitational signals were received by one observer at the same moment they would be received by *all* observers at the same moment. We would then have a way of establishing distant simultaneity, bypassing light-signal delays, and would be able to establish a preferred frame, losing the principle of relativity for inertial physics.

Relativity theory (whether NM, GR, or something else) therefore requires a link between the speed of light (c) and the speed of gravitational signals (c_g).

8.3. Finite c_g gives gravitomagnetism

If the speed of gravity is finite, then gravitational changes take time to propagate, as do gravitational measurements of position. [i] Physics then requires an explanation of how gravitational signalling obeys the principle of relativity (the simplest approach being to assume that the speed of gravity equals the speed of light, $c_g=c$). A finite speed of gravitational signals means that the field of a moving gravity-source is distorted, causing it to seem to pull more weakly when it approaches and more strongly when it recedes, the net effect being a dragging effect on nearby light (with the dragging becoming complete at a body's gravitational horizon, if it has one).

We then have a velocity-dependent deviation from flat spacetime for moving gravitational bodies which must obey the principle of relativity, requiring a relativistic theory of light-dragging effects (a gravitomagnetic theory of relativity).

8.4. Gravitomagnetic theory invalidates special relativity

Unfortunately, this required theory of relativistic gravitomagnetism cannot coexist with special relativity.

If we have a strong-gravity body that drags light, exchanging signals with, say, an individual hydrogen atom, then the principle of relativity requires us to be able to predict the same final outcome regardless of which of the two is said to be "moving". If the strong-gravity body's signals are said to deviate from the SR relationships when it moves due to the complicating effects of

i Since gravitational effects in *Opticks* were supposed to be the result of variations in aether-density, this system didn't require any form of instantaneous gravitational action-at-a-distance.

gravitomagnetism (associated with a finite speed of gravity), then we must be able to predict precisely the same deviation by saying that the strong gravity-source is "stationary" and that that it is instead the "moving" hydrogen atom that is responsible for dragging the light (Fizeau effect), and causing the same gravitomagnetic deviation.

Relativistic gravitation requires the dragging effects associated with moving gravitational masses to be universal, and therefore requires all inertial physics to deviate from the flat-spacetime relationships of special relativity, by the same law.

A finite speed of light combined with the PoR does not automatically prove special relativity to be right: in a universe with gravity, it requires special relativity's equations to be wrong.

8.5. Newton and lightbending

A second common mis-statement that used to be made in "Einstein *vs.* Newton" comparisons was that Newton believed that light moved in straight lines, and that we thought that this was true until Einstein discovered that light was bent by gravity. This impression was encouraged by Einstein's style of exposition:

> Einstein (1914): [59] " *… In recent years it has turned out that such an extension of relativity theory can be carried out, and that it leads to a general theory of gravitation which contains the Newtonian theory as a first approximation. According to this theory, lightrays suffer a curvature in a gravitational field; though minute, it is just within the range of astronomical measurement.* "

A casual reader may feel that they are being told that the new theory predicts the bending of light, and that this implies that the old one didn't. This assumption would be wrong – the idea of lightbending is novel *compared to special relativity*, but not to Newtonian theory or other systems.

> Isaac Newton, ***Principia***, [60] Book I: " *… because of the analogy there is between the propagation of the rays of light and the motion of bodies, I thought it not amiss to add the following Propositions for optical uses; not at all considering the nature of the rays of light, or inquiring whether they are bodies or not; but only determining the trajectories of bodies which are extremely like the trajectories of the rays.* "

While the context for the 1919 Eddington result [61] seems to have been understood correctly at the time (the "Deutsche Physik" movement even tried to accuse Einstein of plagiarism based on an earlier Newtonian lightbending calculation by Soldner in 1804 [63]), in later years it was often misrepresented as having proved a class of effect (gravitational lightbending) that supposedly hadn't been predicted before Einstein.

In more recent years, the community has been forced to acknowledge that the Eddington experiment didn't demonstrate a revolutionary new effect, but showed the superiority of Einstein's *revised 1916 prediction*, which was twice as strong as both the "historical" Newtonian prediction and his own earlier 1911 time-dilation-based prediction. [i]

---

i The Newtonian description of gravitational light-bending can be thought of as a consequence of variations in density of an aetheric medium ("*curved space*"), while Einstein's 1911 prediction applies Huygens' principle to variations in lightspeed due to gravitational time dilation ("*curved time*"). When it came to deciding whether these two effects were cumulative or dual, Einstein's 1916 theory made them cumulative, doubling the previous prediction. However, since the 1911 paper shows that we can predict gravitational time dilation from Newtonian theory, we could presumably update the Newtonian predictions to have the same doubling. Newton himself does not seem to have published a mathematical prediction for the strength of the bending of light by gravity, but did

8.6. Newton and absolute time

A third common "ahistorical" statement is that "while Newton said that time was absolute, Einstein recognised that it was relative". Unfortunately for this nice story, Newton did *not* say that time was absolute: Chapter One of Principia [60] carefully defines concepts of both absolute ***and relative*** time, *relative* times being observer-specific, affected by physical factors such as lightspeed delays (hence the variation in apparent eclipse times of the moons of Jupiter).

> Isaac Newton, ***Principia***, " Absolute, true, and mathematical time, of itself, and from its own nature flows equably without regard to anything external, and by another name is called duration: relative, apparent, and common time, is some sensible and external (whether accurate or unequable) measure of duration by the means of motion, which is commonly used instead of true time: such as an hour, a day, a month, a year. …
>
> Absolute time, in astronomy, is distinguished from relative, by the equation or correction of the vulgar time. For the natural days are truly unequal, though they are commonly considered as equal, and used for a measure of time; astronomers correct this inequality for their more accurate deducing of the celestial motions. It may be that there is no such thing as an equable motion, whereby time may be accurately measured.
>
> The duration or perseverance of the existence of things remains the same, whether the motions are swift or slow, or none at all: and therefore it ought to be distinguished from what are only sensible measures thereof; and out of which we collect it, by means of the astronomical equation. The necessity of which equation, for determining the times of a phaenomenon, is evinced as well from the experiments of the pendulum clock, as by eclipses of the satellites of Jupiter.
>
> … the order of the parts of time is immutable … "

Newton defines absolute time as an abstract *adjusted* or *corrected* time, that does not have to refer to the rate at which clocks actually run (or are seen to run), and acknowledges that since no clocks are perfect, there may well be no actual clock in the universe that runs at that idealised rate.

This is really not the same thing as saying that "time is absolute" and that all clocks run at the same rate.

8.7. Scholarship

In the case of Newton and absolute time, it would seem that readers have looked at the first part of the paragraph defining "*absolute, true and mathematical time*", and have stopped and gone away satisfied, without going on to read the rest of the paragraph about "*relative, common and apparent time*". Either that, or that have not read the original text and have relied on selective quotes in secondary sources.

It can also be difficult to decode shifts in in meaning and context across the centuries. Some commentators have described Newton's physics as a "clockwork" universe, or described Newton's "absolute time" as saying that time functioned as a form of universal clock. However, in the Seventeenth and Eighteenth Centuries, while "clockwork" might have been a synonym for "deterministic" (Kepler, 1605) it would not *necessarily* have implied reliable timekeeping. Principia contains a mention of the fact that we had already been able to measure that our most reliable *pendulum* clocks had been measured as running at different rates at different altitudes as a result of the variation in gravitational field, and Newton points out that, due to perturbing influences, we cannot even trust *celestial* clockwork to run at the "right" speed. [i]

---

publish tables of the equivalent effect of lensing by the Earth's variable-density atmosphere.

i This paper is not meant to be a defence of Newton, or of Nineteenth-Century theory: Newton inverted the relationship between energy and light-frequency, got lightspeed variations back-to front, and failed to predict gravitational time dilation. While we can correct and retrofit these effects (and others) to Newtonian theory with the

8.8. The danger of "educational" information

It is common for educators to make history more "convenient" by misattributing faulty beliefs to earlier generations: many of us will have been taught at school that it was believed that the Earth was flat until **Christopher Columbus** proved otherwise (an "educational" narrative rather spoiled by the fact **Eratosthenes of Cyrene** produced a surprisingly accurate estimate of the circumference of the (round!) Earth back in the BC era).

It is supposed to be an attribute of mature professionals that we can distinguish between "purely educational" data (where the "usefulness" of a narrative for helping students learn sometimes seems to be more important than its truthfulness) and "scientific" data (which is hopefully less compromised).

With many subjects, graduates first learn of shortcomings in the "textbook" version of a subject when they leave the university system and move out into industry, or start fieldwork, or start doing serious experiments, or start interacting with engineers. In "pure" theoretical physics, a researcher's career may keep them almost entirely within the university system, and may not find themselves so likely to be confronted by external peer-groups with different cultural values. Without an external community of sceptical fellow-professionals to "keep us on our toes", we may not feel as strong a need to question and analyse our own core beliefs, and may not realise that some pieces of "educational" information that we have received and taken at face value are not actually correct.

8.9. The "tidying" problem

The tendency of well-meaning physicists to try to "tidy things up" means that when a physics theory is believed to be true, history has a tendency to be rewritten and "optimised" to reflect that believed truth. Narratives that support an inevitable progression towards the theory are emphasised, narratives that conflict with that journey are dropped. We repeat the story that Einstein set out to explain the Michelson-Morley result because it corresponds to an existing literary theme that this is how science works, *even though the story itself appears to be untrue.* The false story is *more efficient* at helping us to achieve our immediate goals (giving an example how science progresses in an orderly fashion) than the real facts, as far as we can ascertain them.

While false narratives that misrepresent a *current* theory's predictions tend to be stamped out (eventually), when we misrepresent an older theory that is already believed to be wrong, and that nobody uses, it is more difficult to find anybody to "cry foul", other than historians.

8.10. Summary

Physics histories written by physicists *who were actually there* are often quite excellent. But accounts of much older history are often untrustworthy, and are often composites of "educational" narratives that never seem to have been fact-checked. Most physicists and educators explaining "what Newton thought" appear not to have even read the introductory chapter of *Principia* or the final section of *Opticks* – since "*everybody knows*" what Newton thought, why bother checking?

> Statements supporting Einstein's special or general theories by comparing them to supposedly more primitive beliefs assigned to Newton and other researchers before Einstein have a habit of being untrue, and should not normally be taken at face value.

---

benefit of hindsight, the resulting system will end up diverging significantly from C18th and C19th theory.

9. SR Proof Nine: "SR is unavoidable because it only has two postulates, which are both correct"

9.1. SR's two "official" postulates

Reading Einstein's 1905 paper, [1] we are told that the theory depends only on two postulates: **(1)** the principle of relativity, and **(2)** the constancy of the speed of light.

> **Einstein, 1905:** " *We will raise this conjecture (… 'the Principle of Relativity') to the status of a postulate, and also introduce another postulate, which is only apparently irreconcilable with the former, namely, that the speed of light is always propagated in empty space with a definite velocity **c** which is independent of the state of motion of the emitting body. …* " i

While this may have seemed true to many readers in 1905, general relativity has since expanded our conceptual vocabulary to the extent that we now appreciate that the phrase "the constancy of the speed of light" can mean different things.

9.2. "Global" *vs.* "local" lightspeed constancy

Under a general theory of relativity, lightspeed is (in a sense) *variable*. If we rotate or accelerate, then light that we would have thought had a constant speed will be seen to apparently be travelling along curved paths, and deflecting as a result of (apparent) lightspeed variations between regions. We can then invoke nominal lightspeed variations associated with gravitational fields to predict the gravitational bending due to light. While special relativity is concerned with the idea that lightspeed is constant, much of general relativity is concerned with the idea that it is "effectively" variable, with this variability causing gravitational phenomena, and allowing us to say that, if we use light signals to define geometry, spacetime is then warped, both by conventional gravity and by the relative acceleration and rotation of masses. ii

> **Einstein (1911)** [12] "[on the calculation of gravitational lightbending] *The principle of the constancy of the velocity of light holds good according to this theory in a different form from that which usually underlies the ordinary theory of relativity* "

What we *can* say according to general relativity is that **the speed of light is still everywhere *locally* constant**. That is, if a gravitational field slows light by causing an excess of space in the region, and by causing time in the region to run more slowly – the result of which we can show by measuring the external perimeter dimensions of the region and then firing a light-signal through it – any colleagues that we might have situated *inside* the region will not be able to detect any change. As they enter the "denser" region of space, the dimensions of their experimental equipment will automatically adjust to the new spatial properties, and the time dilation effect that makes the light travel even slower will also slow the reference clocks used by the apparatus. The effects that alter the region's speed of light also alter the properties of any local measuring hardware by precisely the same amount, in such a way as to make it unable to detect the change

i We might question Einstein's use of the phrase "*definite velocity c*", given that he later argues that (under SR) we cannot *on principle* determine the definite one-way velocity of light, only its apparent round-trip *speed* ("Relativity …", [65] section 8). Subsequent presentations of SR tended to talk about the constancy of the *speed* of light.

ii In singly-connected space, a curved-spacetime theory with local lightspeed constancy can be projected onto a flat background and redescribed in terms of gravitational fields that are associated with variations in lightspeed (Thorne 1994, [22], figure 11.1, page 399) – in an equivalent curved-spacetime interpretation, the theory will normalise away these variations in light-velocity and/or lightspeed by warping distances and times to compensate. If there was *no* apparent variation in lightspeeds, this warping would not be necessary.

with local measurements. The speed of light can be characterised as being ***globally variable*** **but *locally constant*.** [i]

> The second postulate, implemented in the special theory as ***global*** *c*-constancy, is not a law of nature – it is routinely violated.
> What *does* appear to be a physical law is the weaker condition of ***local*** lightspeed constancy.

9.3. SR's third, implicit postulate

In order to arrive at special relativity … in order for an experimenter to be justified in extrapolating their own purely local sense of lightspeed over a wider region containing other bodies with different states of motion … in order to justify converting strictly local *c* to wider, global *c* … we require a third postulate, **that spacetime be totally flat**. [ii]

We require,

(3 i) that the *initial geometry* of the region of spacetime is flat – empty space, no gravitational or curvature/density fields (Einstein does specify empty space), and,

(3 ii) that it is unaffected by the presence of matter we use to conduct our experiments, and,

(3 iii) that it is unaffected by any relative motions of that matter.

If the presence of matter alters lightbeam geometry (**3ii**), then our geometrical derivation will not apply when matter is present, and we risk ending up with a theory of inertial physics that only holds in the absence of matter. If the relative motion of matter affects lightbeam geometry (**3iii**), our flat-spacetime geometrical theory will not give correct answers when matter is moving.

The "flatness" condition amounts to saying that an observer's local sense of constant *c* can legitimately be extended out to cover a wider region, that incorporates other bodies with significantly different states of motion ("no difference between local *c* and global *c*").

Another alternative way of writing the missing third postulate might be that, after having established how the principle of relativity would appear to play out in empty space, the resulting relationships are supposed to still hold when we violate the empty space condition in order to insert physical bodies (given that the result is supposed to be "*a … theory of the electrodynamics of moving bodies*"). To obtain special relativity as *physics*, we need the initial assumed flatness implicit in the idea of empty space to still be correct after we have added some fast-moving massed particles in order to do physics.

To arrive at special relativity as a theory of the interactions of *moving bodies*, the assumption of empty space is not enough – we also have to assume that ***the relative motion of matter in a region has no effect on the propagation of light***, and that this absence of complicating effects makes it valid to model inertial physics as a problem in flat spacetime. [iii]

i Under special relativity we say "*You can't travel at the speed of light*". Post-SR we have to be more careful about how we phrase things: you *can* travel faster than *some* people's light (if they are immersed in an intense gravitational field, and their speed of light is slowed to a crawl) … what you *cannot* do is travel faster, locally, than *your own* light. You cannot catch up with or overtake your own lightsignals along the same path.

ii We might try to rescue the two postulates by saying that "*empty space*" means "also empty of gravitational field gradients". But given that any variation in *c* that preserves "local *c*" can be *described as* gravitational, this becomes a slightly circular argument ("*it is a law of nature that the the speed of light is globally constant in situations where it is globally constant*"). It is simpler to add an explicit third postulate requiring flatness.

iii Without the condition of flatness (or global *c*), other theories may be available, for instance, a fully-dragged-aether theory might have locally constant *c* for every massed particle, but only in the immediate vicinity of those particles. The regions *between* the particles would be associated with lightspeed gradients. The system might also conform to

9.4. Is special relativity robust?

The art of modelling is the art of *data reduction* – the identification of the critical features of a system that need to be preserved in a model, and the identification of others that are irrelevant, or insufficiently important to justify the effort of implementing them. In theoretical physics we distil out only the key features of a behaviour that are universal to a problem, and use this minimalist skeletal description to derive universal laws.

It is legitimate to derive theories from idealised (and often unrealistic) assumptions, provided that we can then show that the final form of the theory doesn't seem to change when those conditions are relaxed – that the theory does not have a *critical dependency* on the idealisations used to derive it. To say that a theory is "robust" is to say that when we perturb the initial idealisations, the essential character of the theory and relationships remain.

If we derive the SR relationships and Minkowski geometry for the principle of relativity applied to empty space, do these results carry over to spaces populated by real masses? Or are the calculations *sensitive* to departures from perfect flatness associated with matter?

Unfortunately, there is sensitivity. If we allow particles to have associated curvature, we break condition (**3ii**), and when curvature-sources move we get further complicating gravitomagnetic effects breaking (**3iii**). In a universe that associates gravitational mass with inertial mass (principle of equivalence) and has a finite speed of gravity, there must *always on principle* be a deviation from flat Minkowski spacetime as a function of masses' relative velocities, and since SR and Minkowski geometry are inseparable and interchangeable, (section 10) this translates into a velocity-dependent deviation from the equations of special relativity.

If the principle of relativity still holds, the departure from SR must be *Lorentzlike*, and if we accept that the distortions due to increased energy must be associated with *positive* rather than negative curvature, this Lorentzlike divergence must be to the red. While we may still *hope* that the divergence will be infinitesimally small, hoping for a result is not the same thing as deriving it.

9.5. Summary

While the idea that Einstein's special theory only has two postulates (relativity and constant lightspeed) makes the theory compelling and apparently inescapable, a more careful analysis shows that it also depends critically on *global* lightspeed and on the idea that massed particles have no associated curvature ($c_{\text{LOCAL}} \equiv c_{\text{GLOBAL}}$).

If particles (and bodies) *can* have associated curvature, the rules of the game alter, and special relativity is no longer the only possibility.

If particles (or bodies) *do* have associated curvature, special relativity is ruled out.

---

the principle of relativity and the two 1905 postulates, but would not give us special relativity.

10. SR Proof 10: "Minkowski spacetime is perfect and unavoidable"

10.1. A perfect solution

Minkowski spacetime, presented by Hermann Minkowski (1864-1909) in a lecture in Cologne in 1908 [10] is a four-dimensional geometrical expression of the relationships of special relativity, and of how an observer's perception of the alignment of space and time coordinates is supposed to change when their worldline is aligned differently to background events. Minkowski spacetime is a *mathematical entity* and a *geometry*, and is the *only* geometrical solution that combines the principle of relativity with flat spacetime. If the principle of relativity is correct and spacetime is flat, then the geometry of spacetime *must* conform to Minkowski's description, and special relativity must be fundamentally correct. This is not negotiable.

But this is still not sufficient for Minkowski spacetime to necessarily be *physics.*

10.2. Dependency on initial assumptions

Minkowski spacetime, like special relativity, relies on the implicit assumption that spacetime is flat and/or empty, and that the presence and relative motion of any masses subsequently introduced into a region has zero effect on lightbeam geometry. Experience tells us that this is wrong: place a chunk of glass in the signal beam and the light slows, and the relative difference in speeds can cause parts of the light to deflect to one side or another (refractive index). This is the how optical lenses work: if the presence of matter *didn't* disturb lightspeeds, the human eye would not be able to focus (and you probably wouldn't be able to read this). Similarly, if we cause that chunk of glass to *move*, experience tells us that the moving particulate medium should cause a measurable asymmetry in the one-way velocity of light in the region (Fizeau ~1850 [102], [103], section 19).

These particulate-matter dragging effects have counterparts in gravitational theory: Place a strong-gravity body in the signal path, and the signal slows (Shapiro effect [66], [67]), and velocity gradients cause the deflection of light, giving gravitational lensing. [68] Move the strong-gravity body and momentum exchange causes an asymmetry in light velocities due to dragging effects, analogous to the Fizeau effect. [102], [103] This suggests that the relative motion of particulate matter may be better described by a curvature-based theory than by flat Minkowski spacetime. [i]

10.3. Robustness and extensibility

What makes Minkowski spacetime excellent as a scientific construct is that it is eminently falsifiable – its predictions cannot be altered or fudged, and it is not open to small modifications. It is either ***utterly correct***, or it is wrong, in which case we need to construct an alternative relativistic system based on different geometrical principles. [ii]

When we move on to gravitational theory, we find that Minkowski spacetime cannot be correct for regions that contain bodies with associated gravitational fields … and under a general theory, *all* bodies with mass have associated gravitational fields, even if these fields are normally thought of as weak (*see:* section 37).

---

i While SR and Minkowski spacetime are sometimes described as "general relativity with gravity switched off", "switching off" gravity under a general theory is an illegal operation, as "switching off" gravitational mass means that we also lose *inertial* mass, and cannot have a theory of inertial physics.

ii Any departure from Minkowski spacetime presumably also needs to be relativistic, and if we are going to construct a new relativistic curved-spacetime theory to deal with just the departures from Minkowski spacetime, with new principles and new methods, we may as well start over and create a whole new theory from scratch.

10.4. Curvature plus relative motion gives gravitomagnetism

If we believe that strong localised massenergy distributions (representing massed particles) are associated with curvature effects, then Minkowski spacetime is falsified. If two bodies have curvatures, then when they move relative to one another the motion of those distortions in spacetime, combined with a finite speed of gravity, produces additional velocity-dependent distortion effects (sections 8.2, 8.3). These effects cannot be retrofitted to the Minkowski system. The sheer crystalline *perfection* of Minkowski spacetime's match to flat spacetime means that the system is *not extensible* to cope with moving bodies with associated curvature, as required by general relativity and particle physics. It is a perfect, final answer to a question of how physics behaves in an idealised universe different to ours, but its rules do not transfer across. [i]

Minkowski spacetime is like the joke about the scientist who produces a "theory of chickens" that "only works for spherical chickens in a vacuum" (section 11.7) ... the difference being that in the case of Minkowski spacetime, the solution *still* doesn't work until we also get rid of the chickens.

10.5. Over-simplicity of Minkowski spacetime

While working towards a general theory, Einstein described distinction between inertial and noninertial physics as "*an inherent epistemological defect.*" (Einstein 1916, "§2. *The need for an extension of the postulate of relativity*" [2]).

Philosophically speaking, Minkowski spacetime represented a **prior cause** – an entity that imposed behaviour without being in any way affected by the results. We could say, "*Space tells matter how to move, and matter does whatever it's bloody well told.*" It was action without reaction (or back-reaction), a strictly one-way causal relationship of the sort that Einstein (after having had more time to think about) decided that he deeply disproved of.

> Einstein (1921): [40] " *… from the standpoint of the special theory of relativity, we must say, continuum spatii et temporis est absolutum … absolutum means not only 'physically real' but also 'independent in its physical properties, having a physical effect, but not itself influenced by physical conditions' ...*
> *It is contrary to the mode of thinking in science to conceive of a thing (the space-time continuum) which acts itself, but cannot be acted upon.* "

The new, reinvented spacetime of general relativity was more "democratic" (Wheeler: "*Space tells matter how to move, matter tells space how to curve.*" MTW [53] page 5), with matter and space having a more delicate equilibrium: Matter could be regarded as a sort of condensation of space, space could be regarded as an extension of matter-fields, and spacetime curvature was something in between.

> Albert Einstein, "**Relativity** ...", Notes to the Fifteenth Edition, "*Physical objects are not* ***in space****, but these objects are* ***spatially extended****. In this way the concept "empty space" loses its meaning.* "

However, since the 1916 general theory incorporated special relativity, its spacetime had to incorporate Minkowski spacetime as a flat limit, meaning that the "defect" was still present.

---

i A quote sometimes attributed to Einstein is that "*Everything should be made as simple as possible, but not simpler*". In the case of Minkowski spacetime and matter, we have a system based on the idea of empty spacetime that is beautifully minimalistic, but just too idealised to generate the more messy laws of physics that must operate when real particulate matter is involved, where inertial mass is associated with gravitational mass: in such a universe, Minkowski spacetime relies on "abstracting away" a critical property of matter: it is "simpler than possible".

10.6. Atoms as curvature

Clifford had already suggested that perhaps some problems in physics might be explicable as results of small-scale curvature – just because space seemed flattish at *our* scales, it didn't mean that there weren't strong curvature effects operating at the scale of the very small – this gave us a potential opportunity to try to model particle physics using curvature principles:

> W.K. Clifford (1870): [69] " *In particular, the axioms of plane geometry are true within the limits of experiment on the surface of a sheet of paper, and yet we know that the sheet is really covered with a number of small ridges and furrows, upon which (the total curvature not being zero) these axioms are not true. Similarly, [Riemann] says, although the axioms of solid geometry are true within the limits of experiment for finite portions of our space, yet we have no reason to conclude that they are true for very small portions; and if any help can be got thereby for the explanation of physical phenomena, we may have reason to conclude that they are not true for very small portions of space.* "

Einstein also "had a shot" at attempting to assign gravitational fields to atoms (1919 [71]), and treating fundamental particles as curvature singularities. [72], [73]

10.7. Replacing Minkowski spacetime

In a more realistic physics, velocity-dependent distortions cause a region's spacetime geometry to distort, to reflect the relative motion of particles within it. The geometry is fully **dynamic** and the physics of the region – which particles it contains, where they are, and how they are moving – is described by these distortions. The distortions are the region's data-storage, and also its computing power – an extrapolation of how the distortions interact dynamically over time tells us how the particles will move, and their past and future positions.

> W.K. Clifford, "On the Space-Theory of Matter" (1870) [69]
>
> " *I hold in fact*
>
> *(1) That small portions of space are in fact of a nature analogous to little hills on a surface which is on the average flat; namely, that the ordinary laws of geometry are not valid in them.*
>
> *(2) That this property of being curved or distorted is continually being passed on from one portion of space to another after the manner of a wave.*
>
> *(3) That this variation of the curvature of space is what really happens in that phenomenon which we call the* ***motion of matter****, whether ponderable or etherial.*
>
> *(4) That in the physical world nothing else takes place but this variation, subject (possibly) to the law of continuity. …* "

None of this is possible if we presuppose that the critical geometry describing inertial physics is flat.

This next-generation system of physics cannot be obtained using Minkowski spacetime, because the presence of velocity-dependent distortions means that we have velocity-dependent *deviations* from the Minkowski geometry, and (since Minkowski geometry and SR are mutually defining) therefore also velocity-dependent deviations from special relativity.

> Minkowski geometry only applies if the concentration of massenergy associated with massed particles (which is quite large, thanks to $E=mc^2$) has *no associated curvature.*
>
> If a massed particle has *any gravitational field whatsoever*, then its physics (including its Doppler relationships) cannot be correctly described geometrically using Minkowski spacetime or special relativity.

10.8. Learning to let go of Minkowski spacetime

It is tempting to say that because Minkowski spacetime is *so* nice, that we want to keep it – perhaps we could retain it and just make some tiny, *tiny* minimal correction to it? This doesn't work. Minkowski spacetime is already a completely finished solution, and cannot be developed any further. In theoretical terms, it is a dead end.

> Richard Feynman, "*Seeking New Laws*" (1964) [74]
>
> *"Newton's ideas … agreed with experiment very well but in order to get the correct motion of the orbit of Mercury, which was a **tiny**, tiny difference, the difference in the character of the theory with which you started was enormous. The reason is these are so simple and so perfect they're produced definite results.*
>
> *In order to get something that produced a little different result, it has to be completely different.*
>
> ***You can't make imperfections on a perfect thing. You have to have another perfect thing.***"
> (emphasis added)

Suppose that, as geometers, we had recently discovered the square, and loved the geometry. We loved the way it broke space into perpendicular x and y coordinates, and when it came time to explore the properties of a circle, we said, "*We know that squares are fundamental geometry, let's build on what we know and define a circle based on the properties of tiled squares!*" Since "pi" is irrational, this would be a fundamentally bad idea. To go from the description of a square to a circle, we need to introduce new concepts (rotation), and discard others (identifiable corners). Some of what the square has taught us about *geometry* will carry over, but *the solution itself* will not.

10.9. Summary

Minkowski spacetime is a perfect description (the *only possible* description) of how classical relativistic inertial physics must operate in a universe that does not obey the general principle of relativity or the principle of equivalence, in which matter does not affect the propagation of light, (no refractive index), and in which gravitational effects do not exist. If we inhabit such a universe, we can justifiably mock the logic of anyone who suggests that special relativity is wrong.

However, if we live in a universe in which *any* of these conditions are broken, the geometrical proofs go into reverse: the *very perfection* of the Minkowskian fit to perfectly flat spacetime then means that if moving bodies have any velocity-dependent departure at all from flatness, the resulting geometry *absolutely cannot* be Minkowski's. Special relativity then cannot be the correct physical description (in our universe) of the relativistic interaction of inertial masses, and some other system must apply.

Since the principle of equivalence says that *every* massed particle has a gravitational field, Minkowski spacetime is a perfect hypothetical answer to an inappropriate question. Every inertial physics problem including real matter *must* include a curvature deviation from special relativity, and the means by which those deviations obey the principle of relativity then requires a new relativistic system – not SR (because SR assumes flatness), and also not GR1916 (because GR1916 assumes a perfect reduction to the SR relationships), but Something Else.

> If we require an exact solution from a theory that incorporates the principle of equivalence of inertia and gravity, the perfection of Minkowski geometry's fit to flatness does not make special relativity *unavoidable*, it makes it *unachievable.*
>
> This is not a matter of personal aesthetics, or convictions or beliefs about how physics "ought" to work: it is a matter of geometry.

11. SR Proof 11: "Curved-spacetime theories must reduce geometrically to SR"

11.1. The reduction argument

A key theoretical argument for the correctness of special relativity is geometrical reduction.

We can say:

> "*Just as curved classical geometry reduces, if we zoom in on a line segment sufficiently far, to an arbitrarily-close agreement with a straight line, so must curved-spacetime physics, if we look at a small enough region, reduce to flat-spacetime physics. Since the only possible relativistic theory of inertial physics in flat spacetime is special relativity, all gravitational theories that support the principle of relativity for inertial motion* ***must*** *reduce to SR physics*".

> Einstein, **Relativity**, chapter 22: [65] " *… it has often been contended by opponents of the theory of relativity that [SR] is overthrown by [GR] … No fairer destiny could be allotted to any physical theory than that it should of itself point out the way to the introduction of a more comprehensive theory, in which it lives on as a limiting case.* "
>
> " *… the general theory of relativity enables us to derive theoretically the influence of a gravitational field on the course of natural processes, the laws of which are already known when a gravitational field is absent.* "

While this is certainly an *argument* for the validity of special relativity, it is not a *proof*, because there is an obvious case in which it fails: it does not work if we assign curvature to particles.

The view of Nineteenth Century geometer **William Kingdon Clifford** can be summarised as "*all physics is curvature*". [69] In a "Cliffordian" universe, all massed particles have associated curvature, and the physics of how the particles interact is described by the interactions of these curvatures – the curvature *is* the physics.

We can compare this to Eddington's similar description of particles under a general theory:

> **Eddington 1920,** [61] page xi : " *Matter does not cause the curvature of spacetime (G): it* ***is*** *the curvature* "
> page 46 (Einstein's Law of Gravitation): " *It will be seen that the measured space around a particle is not Euclidean.* "

In a "Cliffordian" universe, the reduction to flat spacetime does not yield *flat-spacetime physics*, as in this type of universe there is *no such thing* as flat spacetime physics: it instead yields the limit (the "zero particles" solution) at which meaningful physics can be said to have already disappeared. [70] It is the limit at which relativistic observerspace physics has already vanished, because there is *by geometrical definition*, nothing to observe, and also no possible physical observer.

11.2. How can a geometrical proof fail?

The key to understanding the failure of the "reduction" argument is to recognise that even if "*all physics is geometry*", not all geometry is physics. In a Cliffordian universe, where all massed particles have associated distortions, we can only obtain effective flatness by removing all matter some distance away from from the region under study. The equations that we derive for flat spacetime will be *critically dependent* on no matter being present (or nearby), and any equations that we derive for curvature-sources with relative motion (the physics of interactions between *actual* matter) will then necessarily be different.

We can still derive special relativity as mathematics, but only as an unphysical "null" solution.

- **If particles have *zero* distortion**, then we do *not* live in a Cliffordian universe, the reduction argument holds, and the PoR applied to inertial physics gives us special relativity.
- **If particles have *positive* distortion**, then we *do* live in a Cliffordian universe, the relative motion of particles must be associated with velocity dependent curvature, particle-particle interactions cannot conform to flat Minkowski spacetime, and the PoR ***cannot*** give us special relativity.

If particles have associated curvature, then as we examine the interaction of these particles in a smaller and smaller arena, the *background* field gradients will effectively disappear, leaving us with inertial physics played out against a flat backdrop ... but that physics will still not be "flat spacetime" physics, because the process of "zooming in" does not eliminate the distortion-fields that belong to the particles that perform the physics.

11.3. The free-fall argument

> Einstein "Ideas and Methods" *circa* ~1920+ [39] " *I got the happiest thought of my life in the following form: In an example worth considering, the gravitational field has a relative existence only in a manner similar to the electric field generated by magneto-electric induction.* ***Because for an observer in free-fall from the roof of a house there is during the fall*** *– at least in his immediate vicinity –* ***no gravitational field.*** "
>
> " *… the entire conceptual system of the theory of special relativity can claim rigorous validity only for those space-time domains where gravitational fields (under appropriately chosen coordinate systems) are absent. The theory of special relativity, therefore, applies only to a limiting case that is nowhere precisely realized in the real world. Nevertheless, this limiting case (also) is of fundamental significance for the theory of general relativity; because the fact from which we started out, namely that no gravitational field exists in the immediate vicinity of a free-falling observer, this very fact shows that in the vicinity of every worldpoint the results of the theory of special relativity are valid (in the infinitesimal) for a suitably chosen local coordinate system.* "

The shortcomings of Einstein's argument are that,

(a) ... while the freefall argument says that explicitly "gravitational" physics should reduce to "inertial" physics over smallish regions, this is not enough to decide *which* inertial physics is to be reduced to.

(b) ... while "zooming in" makes the background field gradients disappear, it does not make any field gradients disappear that are associated *with the particles themselves*. Zooming in on a patch of spacetime that contains a small experiment allows us to say that the experiment is being performed against an effectively "flat" *background*, but does not let us say that the internal physics *of the experiment itself* is therefore flat. We can only know that no gravitational field exists in the immediate vicinity of a free-falling observer if we define *the observer themself* as not having a gravitational field, which violates the PoE.

(c) Invoking infinitesimals is perhaps slightly "naughty". The distinguishing feature of special relativity is that it assumes global *c*-constancy, and if *c*-constancy is merely *local*, we can have a different type of relativistic model in which local lightspeed constancy is regulated by curvature. Having realised that *c*-constancy is only local, we might reasonably ask "*How does this change the predictions?*" Instead, Einstein argues that a global lightspeed constancy-based theory is still correct, but that it only operates over vanishingly small pointlike regions. At this point we are entitled to raise a quizzical eyebrow.

Another definitional "logical disjunct" worth mentioning is the idea that special relativity derives the laws of physics for moving particles exchanging signals, in a vacuum. If we have a vacuum, then by definition we have no particles, and if we have particles, then by definition, the region is not a vacuum.

11.4. Theory and logic

If Einstein's argument for the absence of curvature isn't valid, can we use other theoretical arguments to prove that particles have zero associated distortion? Under a general theory, no ... because the principle of equivalence of inertial and gravitational mass *requires* particles with inertial mass to also have gravitational mass (and therefore curvature). Thanks to to $E=mc^2$, a massed particle represents a significant variation in the region's energy-density, and needs to be associated with a corresponding deviation from flatness. While we would seem to be able to obtain SR's relationships for the physical behaviour of matter by specifying the *absence* of matter (empty space), there is no guarantee that the region remains flat when it is populated with real particles with significant relative velocities. When curvature-sources have relative velocities, we get additional velocity-dependent curvature effects, and the resulting "dynamic" geometry no longer conforms to the rules of flat, "fixed" Minkowski spacetime.

11.5. Gravitational theories that conform to the PoE cannot on principle reduce exactly to the equations of SR physics.

Not only is the reduction argument not general, in the context of the GPoR it is not correct. Since special relativity is so emphatically and unambiguously "the theory of relativity in flat spacetime", the existence of any velocity-dependent deviations from flat spacetime tells us immediately that ... whatever the correct equations of motion may then turn out to be ... the one thing that we know *immediately*, without performing a single calculation, is that they *cannot* be those of flat-spacetime special relativity. In a geometrical theory, in which physical relationships are derived from geometrical relationships, we cannot change the geometry without also changing the physics.

Physical law is then required geometrically to be Something Else. A correct analysis of the situation does not enforce SR: it excludes it.

11.6. Experimental testing of particle curvature

How might we carry out experimental tests to see if this supposed particle-curvature exists, and whether the effect is strong enough to be detectable?

1. If massed particles all have associated gravity-wells, then light will take longer to cross a region populated by such particles (Shapiro effect [66]), and even longer if the number and density of the particles is increased, and if the particles are substituted for others with greater mass.

2. If a body of massed particles *moves*, then the expected gravitomagnetic dragging effect should cause a measurable offset in the speed of light in the region (again, stronger if the particles are more numerous and more densely packed).

3. If we use wavelengths of light comparable to the scale of the curvature, then using light to measure the distance across a region will report a greater distance if the wavelength of the light is smaller (and can penetrate further into the region's smaller geometrical features). The previous effects **(1)** and **(2)** will also be stronger for shorter wavelengths.

In reality we find that,

1. The speed of light *is* indeed slower in particulate media than in vacuum, and is slower in glass than air, and slower in lead glass than in normal glass (refractive index).
2. The speed of light *is* dragged when the particulate medium moves (Fizeau effect). [102], [103]
3. The speed-drop and dragging effect *is* more pronounced for bluer light than red (diffraction).

The basic phenomenology of C19th and C20th optics is in accordance with the idea that particles do seem to show significant associated distortions and measurable light-dragging effects, in agreement with the PoE, but in disagreement with special relativity's requirement that interactions between particles and light behave as if moving masses have zero effect on lightbeam geometry.

If we compare known effects against Clifford's idea and Einstein's 1905 concept, Clifford's seems to make a better match to reality – the phenomenology seems to suggest that our universe is Cliffordian.

11.7. Mathematically provable results are not always valid physics

A key question that practitioners of mathematical physics should ask themselves is: "*Do I want to to be able to **prove** my results, or would I prefer to get the correct answer?*"

Sometimes, what is correct cannot be derived and proved using traditional induction, and sometimes what can be derived and proved is not correct. This idea, expressed by both Einstein (1921: [75] "*... as far as the propositions of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality. ...* ") and Hawking ("*I'd rather be right than rigorous*") is not quite as paradoxical as it seems:

- **In order to obtain a *guaranteed, incremental* proof**, we often need to impose harsh initial idealisations and assumptions onto the problem to make the process of obtaining a strict proof easier ... and if a proof is especially challenging, the required idealisations can be so severe as to change the form of the final answer. Assuming flat empty spacetime makes it easier to derive and prove a set of relativistic equations ... but the assumption generates a *different* set of equations to those that we would have arrived at by saying that "real" physics requires physical masses to be associated with curvature.
  The "mathematically provable" result won't necessarily be the correct set of equations. [i]
- **If our universe is self-contained and self-sufficient**, then we might expect its fundamental laws to be self-contained and self-sufficient, too. We may be able to arrive at these final laws through a process of trained intuition, inspired deduction, and artistic sensibility ("*what would be the most efficient laws with which to build a universe?*") but find that the resulting convergent system, which simultaneously modifies multiple disciplines in order to get them to fit together, might not allow an incremental "bottom-up" construction – it might be "top-down". The rule-set that runs our universe might be an emergent property of theory-space, with properties based less on hierarchy and more on duality, parallelism, holographism and circular logic. If we are especially unlucky, they might turn out not to be derivable by traditional incremental logic, *on principle.* [ii]

---

i This is again the "spherical chicken" problem (Stellman, 1973 [76]). The "spherical chicken" idealisation may be useful if we want to comply with animal welfare laws by guaranteeing a certain minimum distance between chickens, but should not persuade us that we can efficiently stack chickens in a self-supporting three-dimensional array.

ii Nature does not produce an adult horse by constructing the torso, head and limbs separately, and then fitting them all together. Similarly, our universe's laws might be an "organic" solution that defies an incremental piecemeal approach.

11.8. Previous examples of failed proofs

It is not difficult to find previous examples of "unbreakable" proofs that turned out to be bad physics, [i] and which were responsible for holding back progress in their respective fields of research:

- **The impossibility of people-carrying heavier-than-air aircraft** – famous polymath **Simon Newcomb** (1805-1909) argued that even if one could build a small flying scale model aeroplane, the design would never "scale up", and could never be built with then-current materials. Each doubling of size meant 2×2 times the lift area, but 2×2×2 times the volume and weight, halving the lift-to-weight ratio and efficiency, which was why there were no very large flying birds. Partly as a result of Newcomb's geometrical disproof, early aeronautics research had to be almost entirely carried out by self-funded amateurs.
  In the 1890s, Alexander Graham Bell publicly disproved Newcomb's argument by inventing the field of fractal engineering, flying a person-carrying kite based on a fractal tetrahedral box-kite, whose weight scaled linearly with lift area.
  The math community retaliated by studiously refusing to recognise or cite Bell's work, and later naming his fractal tetrahedral pyramid shape after Wacław Sierpiński. [ii]

- **The impracticality of liquid-fuelled rockets** – Due to the "dead weight" of fuel tanks and pipes, it was said that one could prove the mathematical impossibility of using liquid-fuelled rockets to raise a payload into orbit, making the subject seem like a dead end, so that further early research again had to be carried out by persistent self-funders. The US and Soviet space programs sidestepped the proof by simply discarding the unnecessary "spent" infrastructure as a craft climbed.

- **William Thomson's disproofs of Darwinian evolution** – For the known complexity of life on Earth to evolve naturally would take billions of years. But the Sun's chemical fuel (said Thompson, *a.k.a.* **Lord Kelvin**) would then have been exhausted long ago, and temperature readings in boreholes showed that the Earth's interior was still so hot that (extrapolating backwards to take into account cooling effects) the surface would have been molten mere hundreds of millions of years ago. Thompson argued that physics was real science, geology and evolutionary biology weren't, that members of those disciplines who disagreed weren't real scientists, and that *thermodynamics itself* proved that the Earth was young, and that Humankind had been created by a benign and benevolent God.
  Thompson's public campaigning wrecked careers and held back geology and evolutionary theory – his fame and position as President of the Royal Society meant that to argue with Thompson was career suicide. We now know that stars are powered by nuclear fusion, and the Earth's core is heated by nuclear fission, from heavy radioactive elements that sink and concentrate at the centre. A more open-minded physicist could have noted that geology required the Earth to be old and that thermodynamics gave a different answer, and used the discrepancy to postulate the existence of some new, previously-undiscovered power source – if Thomson had instead embraced geological findings and Darwinian theory, he could have used geology and biology to predict the existence of nuclear energy.

This is not to say that mathematical proofs are completely useless in physics – invalidating them can be an enjoyable pastime and has educational value.

i Bertrand Russell, 1917: [77] "*Mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true.*"

ii The math community may also have had a dark suspicion from Bell's cheerfully irreverent tone that he may have been implicitly suggesting that Newcomb "go fly a kite", a rather rude suggestion.

11.9. "Small" deviations from Minkowski spacetime

Having established the principle that special relativity *must* be at least *nominally* wrong, the next obvious question is, "by how much?" [iii] According to the standard view, even if general relativity cannot after all reduce *exactly* to the flat spacetime physics of SR, the divergence must surely be so vanishingly small that we can safely ignore it.

But the curious scientist will not be satisfied by this. A deviation from SR due to particle curvature alters the Doppler equations and equations of motion, and the principle of relativity requires that any such deviation must apply identically to all matter in our universe. Additionally, we know that in order to be "relativistic", the deviation must be "Lorentzlike", of the form $(1\text{-}v^2/c^2)^{exp}$, and must be a deviation to the red rather to the blue. The deviation can therefore be expressed as a single number ("exp"), with a value greater than zero, but no more than one half.

The inquisitive theorist will not be satisfied with being fobbed off by being told that this number is "something really very small" ... they will demand to know its actual value (at the very least to an order of magnitude), and will not be able to rest until they know something more about this number.

- The number might be some simple geometrical ratio, such as "a half", or "one-over-pi", the justification for which might then be found later from fundamental principles. But since most basic ratios are reasonably sizeable (compared to the range 0-0.5), this would suggest that the deviation was not small, meaning that it would count as a more major change to existing theory.
- The number might be tiny, but related so some fundamental physics ratio (as the gravitational constant *might* be, according to the Large Numbers Hypothesis).
- The number might be tiny, with no obvious justification. In this case, it would count as one of the universe's fundamental constants – numbers that together "seed" the physics of our universe, in much the same way that a pair of numbers used to generate a Julia set define which version of the set we end up creating. As with the Julia "seed values", our universe might be one of a continuum of potential universes defined by coordinates that locate where our reality exists on a larger map or superverse of potential realities with different laws of physics.

Since the principle of universality requires this "special" number to apply to all matter in the universe, and its value must be fixed, independently of the nature of the matter, it would seem to have to be one of the fundamental constants of nature, and we will be quite anxious to find out its actual value.

Even if we believe the value to be tiny, this does not make it unimportant. The gravitational constant is also tiny, but is desperately important to gravitational theory. Similarly, Einstein's relationship between mass and energy, expressed as $m=E/c^2$ gives values for the change in mass of a clockwork motor, electrical battery or chemical reaction that are far too small to ever be measured ... but setting the "insignificant" change in mass to zero would be a terrible mistake. Understanding $E=mc^2$ helped usher in a new science of nuclear energy and helped us to understand how stars "burned".

What new discoveries might be born from embracing and exploring the implications of a new piece of fundamental physics?

---

iii An answer to this question is given in other parts of this paper.

12. SR Argument 12: "How could a general theory NOT reduce to a special theory?"

12.1. Minimal structures

We may feel that the idea that a general theory *not* reducing to a "special-case" theory as a limit is perverse, as *surely* a general case reduces to a special case *by definition*?

This expectation may be true for geometry, but does not always work for physics. An efficient *theory* tends to be an assemblage of laws and principles, cross-connected in such a way as to produce a multiply-redundant **minimal structure**, the defining property of a minimal structure being that if any one element is removed, the structure fails.

A good theory will then often *not* be a perfect subset or superset of another good theory that tries to do a similar job: each will tend to represent the most efficient possible use of the number of available degrees of freedom.

12.2. Components *vs.* structures

To illustrate this principle, we can imagine trying to build a two-lane road bridge to a particular minimal design, from a kit of parts consisting of pre-made metal girders and other components of various exact standardised sizes. If we remove all instances of one of the more common parts (such as a critical support beam), then we can no longer build our bridge.

We may be able to come up with a *different* design that creatively puts the remaining parts together in different ways to make a *different* bridge with reduced specifications (perhaps single-lane rather than two-lane), but our smaller bridge spanning the same chasm will then need to have a distinctly different configuration to to the larger version, and will need to connect the remaining parts together in different ways, and perhaps use some of them for different purposes. Although there may well be thematic similarities between the two bridges, the construction of the smaller bridge will *not* exist within the larger arrangement.

We then have a distinction between the two "*kits of parts*" (lists of features used in a theory), where one *is* a subset of the other, and the resulting designs (theories), where the architecture is uniquely tailored to use the parts available in each case, and one design is *not* a simple superset of the other.

12.3. Non-reducing theories

If we want more literal examples of cases where geometrical reduction doesn't work, we can consider theories of physics that have different numbers of equivalent spatial dimensions:

A three-dimensional universe will obviously yield two-dimensional and one-dimensional descriptions as "slices" through the main model, and if we draw a square or a circle in two dimensions, those same shapes can exist in three dimensions as cross-sections.

But consider what happens when we add signals or fields: if we have one dimension of space and a pointlike emitter, its signal or field will have nowhere to spread out into, and will have the same amplitude or intensity at any distance. If we have *two* dimensions of space, the signal or field will spread out into the *plane*, and weaken as a function of distance, $1/r$. If we have *three* dimensions, the signal or field will spread out into a *volume*, and will weaken as a function of distance-*squared*, $1/r^2$.

Theories constructed for one, two and three spatial dimensions have different fundamental relationships, and the physics of a two-dimensional universe (giving $1/r$) is not the same as the internal physics of a *two-dimensional slice* through a *three* dimensional universe (which should show $1/r^2$). 3-geometry is a superset of 2-geometry, but a valid 3-theory does not reduce to a valid 2-theory. A "larger" theory is not always a superset of a "smaller" theory.

Quantum from classical

If an object moves in 3-space and passes through a 2D plane, it seems (within the plane) to appear discontinuously from nowhere, to grow and evolve, to shrivel away, and then to disappear back to nowhere again. An attempt to describe the fluctuating behaviour of a slice through a classical 3-theory can appear as a partly *non-classical* 2-theory, in which some behaviour has to be described by purely statistical laws. Similarly, a GR1916-style projection of a (classical) acoustic metric's physics generates apparently non-classical QM-style effects as an artefact of the projection.

12.4. "Generality of application" *vs.* "inclusivity"

A further cause of confusion is the word "general", which in mathematics tends to mean that a general law includes all possible sub-laws.

If general relativity used the word "general" in this sense, it might be expected to be a more abstract set of general mathematical laws that could be used to generate all other possible theories of relativity, including Newtonian mechanics and special relativity, and presumably also other theories tailored for other numbers of dimensions. It would be a "meta-theory" rather than a physical theory itself (it might, for instance, have a parameter that lets us select whether we are using, say, a Minkowski spacetime or an acoustic metric, another parameter for the number of dimensions, and so on).

But "general relativity" is not "general" in this sense of inclusivity: its name refers to its (attempted) support of the general principle of relativity, which says that, in Nature, *the relativity principle applies generally.* [i]

12.5. The "Newtonian approximation"

An objection at this point is that surely Newtonian mechanics breaks this pattern by somehow managing to be both a subset of SR *and* GR1916?

This, again, is not quite true. Special relativity is a subset of GR1916 *by definition*, in that all of its laws are *defined* as carrying over into the larger theory. This leads to illegal results, but … it is how GR1916 was constructed, and it lets us argue that SR is technically a fully-contained subset.

By contrast, the laws of Newtonian theory do *not* carry over unchanged into SR and GR1916. The Newtonian Doppler relationships are not a subset of the SR Doppler relationships (or an SR-centric GR): they are a *different solution*, and normally only coincide *exactly* with the SR predictions when v=0.

The key NM relationships can be more accurately considered a *low-velocity approximation* of SR, just as the SR set can be considered a *low-velocity approximation* of NM – but a mutual convergence as v tends to zero does not does not generate a hierarchy. [ii]

---

i We *could* construct a "general theory of relativity" (referring to relativity-as-a-subject") that is "general" in terms of inclusivity, and which includes and parameterises all possible theories of relativity, across a range of potential mutually-incompatible universes. This would not be a single physical theory, but would be a "theory of theories".

12.6. SR reconsidered as a subset of more advanced Newtonian theory

If we really *insist* on creating some sort of logical hierarchy where one of these two theories generates the other as a limiting case, we could argue instead that *SR* is technically a limiting case of *NM*:

Since the NM equation-set is time-asymmetrical, requires curvature to be implemented as a more complex "acoustic metric" theory, and departs from classical energy conservation, we could argue that if we start with an advanced NM-based curved-spacetime model, we can "erase" its curvature effects and time asymmetry (and simplify it) by multiplying its predictions in forward time against its predictions in reversed time, and square rooting – an operation that then results in the flat and time-symmetrical predictions of special relativity. [8]

This process produces special relativity as a simplified but unphysical subset of NM … we can then say that SR represents the more advanced Newtonian theory with curvature "switched off", or with time-asymmetries cancelled, or with energy-asymmetries erased. This makes SR a reasonable idealised flat first approximation of an "advanced" Newtonian theory, but its laws are still not a perfect subset, because the loss of critical physical behaviours present in the more complex theory forces the simpler theory to achieve completeness in different ways, with different equations and principles. In the larger theory, *c*-constancy is achieved locally with curvature, in the smaller theory (without the luxury of curvature), *c*-constancy has to appear differently, expressed as global *c*-constancy and absolute (Minkowski) geometry.

12.7. Summary

> If we were building a general theory of relativity from scratch, around the principle of equivalence and the general principle of relativity, we would have no real reason to expect our curved-spacetime theory to reduce to the physics of Einstein's "flat" special theory.

Physical theories are to some extent tailored to the geometrical options available in a given universe (especially if they are highly-efficient **minimal structures**, in which removing any component destroys the theory). While lower-dimensional geometry is a subset of higher-dimensional geometry, *a lower-dimensional physics is not necessarily a subset of a higher-dimensional physics.* The different set of connections within the more complex system does not necessarily "contain" all the same connections and interrelationships of the simpler system.

Although Einstein's special and general theories are both four-dimensional, the existence of curvature effects in the general theory can be considered as an additional degree of freedom, that might be considered to play the part of a form of extra dimension, and general relativity does arguably require an extra dimension for its *embedding-space.*

> Physical theories that encompass different ranges of effects or that have different degrees of freedom, do not necessarily "nest" like Russian dolls. If we want to design a general theory of relativity from scratch, as a minimal structure, we will want to use gravitomagnetism to physically regulate lightspeeds. If we want to design a special theory assuming the absence of curvature, we need Minkowski spacetime. These two approaches to obtaining local c-constancy are physically distinct, and incompatible.

---

ii Geometrical physics is an exact science. The SR equation-set is a gravity-free solution that (in a gravitational universe) is at best a first approximation for weak fields. By contrast, the NM Doppler equation-set set is the *exact* solution for *maximally strong* fields. Exact solutions *are* possible – every time a geometrical theory avoids a breakdown by invoking yet another first approximation, such as "small" velocities or "weak "fields, a kitten dies.

13. SR Argument 13: "Aesthetics: How can anything as elegant as SR/Minkowski spacetime NOT be correct?"

13.1. Beauty as truth

The idea that the natural sciences can seize on existing bodies of mathematics in the hope that physical law gives meaning to mathematical beauty was described by Eugene Wigner as "The empirical law of epistemology": [140]

> Wigner: "*The Unreasonable Effectiveness of Mathematics in the Natural Sciences*" (1960)
> "*… if the empirical law of epistemology were not correct, we would lack the encouragement and reassurance which are emotional necessities, without which the "laws of nature" could not have been successfully explored. Dr. R. G. Sachs, with whom I discussed the empirical law of epistemology, called it an article of faith of the theoretical physicist, and it is surely that.* "

This correspondence encourages the mathematical physicist to believe that they are "doing science right".

But mathematics is *not* unreasonably effective: as a symbolic representation of classes of logic, it is quite reasonable that … if our universe is logical … its patterns and themes will be found *somewhere* in the Great Book of All Mathematics. To associate successes with some sort of mysticism is both unscientific and counterproductive, and if we want to consider whether mathematics is *unreasonably*, *reasonably*, or ***merely reasonably*** effective we have to look for cases where the results of the mathematical / theological approach has been really rather bad.

Aside from the cases given in section 43, we can see with hindsight that surprisingly many major breakthroughs appeared shockingly late in the historical record.

Gravitational time dilation is a *childishly* simple idea, and yet Riemann failed to see it because it involved querying an initial mathematical assumption, that timeflow was *obviously* universal. $E=mc^2$ should have been derived in around the 1860s, but wasn't … because Newton's idea of interconvertibility was no longer in circulation, and the idea violated conservation laws. Fractals arguably didn't become "respectable" until the Twentieth Century, and something as trivial, screamingly obvious and desperately important as the method of incorporating infinities and zeroes into normal calculations using a hierarchy of parallel number scales, didn't seem to appear until the 1960s (Robinson 1961 [142], [143]). Some of our greatest achievements, blindingly obvious with hindsight, were not made in a timely manner, not because of any technical shortcomings, but because anyone working on those problems considered the correct logical conclusion to be wrong, as it was not in accordance with contemporary aesthetics or received wisdom.

We currently seem to be on the threshold of a range of scientific breakthroughs, that will collapse many principles into few, but are unable to take the necessary step because our existing sense of the aesthetics of natural law has been trained on special relativity and Minkowski spacetime.

13.2. Ideas and concepts

The development of science and mathematics is often portrayed as a series of advances that are only made possible by building on a foundation of highly-technical previous work. An alternative interpretation is that the presumed correctness of each generation of theory may be rather less important to the development of science as the introduction and establishing of its *associated ideas*. If the mathematics is correct, then this obviously adds further to an idea's credibility and the rate at which it is adopted, but correctness is sometimes of secondary importance to how a piece of research extends our **conceptual vocabulary**.

Einstein's general theory is generally considered to have been one of the most successful, influential and celebrated theories of the Twentieth Century, and yet … its internal architecture, geometrically speaking, is junkyard scrap. The value of GR1916 is as a *proof of concept* or *archetype*, a symbol or icon embodying a series of ideas that extends our ability to think, and provides a new series of conceptual platforms on which we can build further ideas. The fact that the internal geometrical "machinery" of GR1916 doesn't actually fit together properly does not seem to keep the physics community awake at night, and also doesn't seem to bother the math community terribly much. Apart from the 1960 Schild paper, it seems to have gone almost without comment.

While it is comparatively easy to do an analysis of the dependencies within GR1916, and identify where the theory goes wrong and why, over half a century after the Schild paper, nobody in the math or physics communities appears to have done this work. This is not due to any inherent technical difficulty, but because **(a)** people are reluctant to consider the idea that a "celebrity" theory *can* be wrong, and **(b)** because the inevitable solution to GR1916's internal inconsistencies means "taking down" special relativity, a theory which we all "know" *cannot* be wrong.

The solution to general relativity's problems cannot be considered without our being prepared to *lose an idea* – that curved-spacetime physics is built on flat-spacetime physics – which we find useful and comforting. It is our unwillingness to give up on an idea that explains why we have not made any worthwhile progress in general relativity for half a century.

13.3. Einstein's general theory as an "unscientific" system

Given the amount of geometry and mathematical notation involved in textbook explanations of general relativity, it may seem perverse to suggest that the 1916 theory is not a truly mathematical/geometrical theory.

However, we can made a fairly convincing case that the theory's architectural specification violates basic mathematical/geometrical laws, and takes the form that it does partly because of emotional/historical/pragmatic reasons.

In terms of set theory, we can write:

SR is a member of the set of theories in which particle curvature is forbidden.

GR is a member of the set of theories in which particle curvature is compulsory.

Can SR be a valid subset of GR?

(spoiler: the answer is "no")

The 1916 theory is not just *not mathematically rigorous*, it is ***not technically valid***.

Despite the fact that mathematics, geometry and logic all tell that a general theory ***absolutely cannot*** reduce legally to special relativity, this seems to be of secondary importance to us to preserving our traditional cultural values. Rigour (in this situation) is discarded when it gives results that we do not want to hear, and we are prepared to set aside our entire structure of logic when logic tells us something inconvenient that disagrees with our deeply-held beliefs.

We do not appear to be especially interested in solving this problem. Our actions (or lack thereof) suggest that we do not seem to be any more enthusiastic about the idea of disruptive change or scientific revolution in relativity theory than the owners of horse-drawn cabs in the early Twentieth Century were enthusiastic about the introduction of the motor car.

13.4. The primacy of the Idea

What appears to be the thing that allows us to evolve mentally, and transcend our petty parochial received wisdoms is not incremental math or logic, but the Idea (ideally, the "Big Idea").

The Idea provides a sense of *potential* legitimacy and persistent identity to a set of thought processes that may eventually turn into "proper mathematics". Einstein's most productive period was in his earlier career where he based his work on The Idea. The idea that only round-trip lightspeeds mattered, that moving bodies contracted, that coordinate systems allowed one to extend local concepts globally, the idea that inertial and gravitational descriptions were interchangeable, the idea that a falling person feels no gravity, and that rate of timeflow might be a variable based on regional gravitational field density ... these led to breakthrough physics. Einstein's later career, which often seemed to involve shuffling mathematics around in the hope of spotting interesting relationships, like a child poking a rockpool with a stick, in the *absence* of new ideas, was rather less fruitful.

Research can be important if it is "interesting" even it it is not correct: similarly a theory's importance may be judged not by the correctness of its mathematics, but by its influence in introducing new ideas. Once we have ideas, the mathematics often follows – the 1916 theory's general concept of curved spacetime is more important than the theory's (flawed) attempted implementation – GR1916 has been more important as a transmission vector for concepts than as a set of mathematics or geometry.

> Einstein's general theory was not important or unimportant because of the quality of its geometry or the consistency of its logic – these, in some parts, were terrible – it was important as a *proof of concept.* [i] Once the concept is defined, any number of mathematicians can implement the details.

13.5. The aesthetic principle

Just because something is geometrically compelling to humans does not mean that the thing arises in Nature (until humans, as agencies of Nature, create it) – the shape of the Mandelbrot Set is a profound and magnificent piece of geometry, but we are not aware of Nature making use of it anywhere.

Humans are occupied (almost to the point of obsession) with squares and rectangles and right-angles. If we look at the buildings and artefacts produced by *Homo sapiens*, or the street grids of modern cities, one could forgive an alien anthropologist for believing that our species is so obsessed with rectangles that we must be rectangular ourselves.

When we create maps, we create rectangular coordinate systems, even to the point of specifying positions on the surface of the round Earth by trying to divide the surface into rectangles. When we create a system of physics in order to try to make sense of our surroundings, we (usually) break it down into x, y and z coordinates. [ii]

i Einstein and Infeld (1938) [{141}] page 95: " *The formulation of a problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science.* "

ii Other coordinate systems are available (*e.g.* polar coordinates).

This is partly because the approach allows incremental thought and rigorous-yet-simple geometrical logic – we love the idea that we can create a unit line, extend it into a unit square and extend that into a unit cube. Rectangles are excellent and often optimally-efficient when it comes to *building and assembling* things, from components (which humans do a lot of).

However rectangle-based shapes are not as fundamental when it comes to *growing* things, and when we look around at Nature, with a few exceptions (such as some crystals and perhaps some radiolaria), there is a notable and profound absence of squares and rectangles. Nature stubbornly refuses to build cube-shaped stars or square galaxies.

We hope that our internal sense of aesthetics may be able to let us decode the rules that govern the universe, partly because we (and our brains, and our minds) are products of that universe. We are *natives*, and this is where we live. But since we design systems to be assembled rather than grown, the geometrical elements that we try to break reality down into to please a sense of aesthetics influenced by engineering and construction do not necessarily correspond to what is actually in front of us. The cleanest rules for creating a universe are different depending on whether we want to construct a universe out of smaller self-contained parts, or whether we wish to grow one, organically.

Being surrounded by Nature does not always seem to give us great intuitive insights into how the mathematical laws of Nature ought to behave. We did not embrace fractal geometry until the Twentieth Century despite having been surrounded by trees for longer then we have been humans, and despite the fact that some of our predecessor species would have actually lived in the things. Instead, we initially denounced fractals as monstrous, and still regard turbulence with distaste. Fractals were offensive, not because they did not correspond to Nature, but because they *did.* [i]

> Traditional mathematical aesthetics can be counterproductive to attempts to derive the correct mathematical laws of Nature.

> Mathematics is not the language of Nature, it is the language of mathematicians.
> Nature does not require a language.

13.6. The profound relationship between mathematics and bad theory

Deep mathematical results often stem from wrong answers. If we try to express the area of a circle as an integer ratio of its radius squared, we fail. It's a fundamentally misguided idea. But the attempt … a cascading infinite series of error-corrections corresponding to recursive tilings that attempt to express an irrational number with an unending series of digits … is interesting. It gets us (for instance) into the theory of fractals, whereas the correct answer, writing simply "2 Pi *r*" is a bit dull.

The mathematics of error, and failure, and bad starting assumptions, is a deep, *deep* subject that cuts across multiple disciplines and has things to say about the nature of reality (or what things look like when they're *not* reality), whereas getting the answer exactly right first time can be a lot less interesting.

---

i Attributed to Heisenberg: "*When I meet God, I'm going to ask him two questions: why relativity? And why turbulence? I really believe he'll have an answer for the first.*" The answer to turbulence and chaotic behaviour in general is that it efficiently generates massive complexity from simple inputs.
The "unsatisfying universe" conjecture: A universe simple enough to satisfy mathematicians would be too simple to allow mathematicians.

Wigner (1960) [140]: “ *Let us consider a few examples of "false" theories which give, in view of their falseness, alarmingly accurate descriptions of groups of phenomena. … the so-called free-electron theory, which gives a marvelously accurate picture of many, if not most, properties of metals, semiconductors, and insulators. In particular, it explains the fact, never properly understood on the basis of the "real theory," that insulators show a specific resistance to electricity which may be 10^{26} times greater than that of metals. In fact, there is no experimental evidence to show that the resistance is not infinite under the conditions under which the free-electron theory would lead us to expect an infinite resistance. Nevertheless, we are convinced that the free-electron theory is a crude approximation which should be replaced, in the description of all phenomena concerning solids, by a more accurate picture.* ”

“ *… The free-electron theory raises doubts as to how much we should trust numerical agreement between theory and experiment as evidence for the correctness of the theory. We are used to such doubts.* ”

13.7. The parable of the flat Moon

Consider the Moon. If we feed a library of astronomical images into a computer-based artificial intelligence (“AI”) system and ask it to make inferences, then, if the first body it considers is the Moon, the system might choke on the idea that regardless of where we take an image of the Moon from (on Earth), it always looks (pretty much) the same. The system, not yet knowing about tidal locking, might decide that it was desperately improbable for the Moon to always show the same face to the Earth, and for its revolution rate to equal its orbital period, correct to an absurd number of decimal places. Is this some cosmological conspiracy? Perhaps the Moon has no other side? Perhaps it’s a flat disc stuck to the interior of a hollow sphere?

In an attempt to show our hypothetical AI system that it is in error, we bring it to the AI’s attention that the the Moon is covered with a series of round(ish) craters, and that the shape of these craters varies, showing that the surface is curved. The AI system analyses these craters and derives an underlying law: if the Moon’s radius is **R**, and the distance of the centre of a crater from the centre of the Moon-disc is ***r***, every crater appears as an ellipse whose minor radius is aligned with a disc radius, with the ratio of minor to major radius being 1: $\sqrt{1 - r^2/R^2}$

In other words, if we assume that the Moon is a flat disc, then details on its surface are radially contracted by the mathematical equivalent of the Lorentz factor. The “flat disc” interpretation generates the Lorentz relationships and a fascinating and compelling body of mathematics.

This mathematics is then so engrossing that … how could the Moon possibly NOT be flat? If assuming flatness generates so much great math, which agree so well with the experimental data, and flatness is obviously simpler than curvature, then wouldn’t it be perverse to go against Occam’s Razor and argue based on some misguided sense of aesthetics or correct physical behaviour that the Moon is *not* flat?

This is similar to the situation we have with special relativity. In the “Moon” example, we can consider lunar craters to be Lorentz-contracted because our fundamental geometrical assumptions (flatness) are wrong. Until the 1960s, we had no images proving that the Moon was ~spherical and that the far side of the Moon actually existed. We *assumed* that the unseen half of the Moon was there because of wider patterns of behaviour: we noted that (thanks to rotation) every other known circular-looking astronomical body in the universe where we had sufficient data to check was a spheroid, after which the simplest interpretation of the Moon (under Occam’s Razor) was that *it* was a spheroid too.

Similarly, the idea that the Earth was the fixed centre of the universe had no direct counter-evidence and seemed to be the simplest and most useful interpretation of the data … until we looked at external systems and found that other planets in turn had their own moons circling them. Bearing in mind this wider pattern of behaviours, it became simpler to assume that the Earth was just another planet.

13.8. Summary

> Occam's Razor can "cut" in different directions depending on the range of phenomena that we wish to be able to explain.

With inertial physics studied in isolation, it may seem simpler to assume that inertial behaviour is a "flat" phenomenon, because "flat" is simpler than "curved", and why would anyone want a curvature-based model when the flat system (with Lorentz corrections) fits the data so well? It's only when we consider the wider spread of physical behaviours, the behaviour of particulate media, cosmology, gravitational theory, horizon behaviour, and quantum mechanics – that a set of unifying themes and principles emerge, in which massed particles pretty much *need* to have rest curvature (and gravitomagnetic curvature when they move), and with which special relativity is incompatible.

The advantage in efficiency of a geometrical explanation that does not require spacetime curvature is somewhat lost when we move to a larger theory that has to include curvature anyway in order to model gravity. Once we have been forced to grudgingly accept the idea of curvature (or its field equivalent), it becomes more efficient to try to use the new idea to explain as much as possible, rather than to try to support two different parallel systems.

In that wider context, it is simpler to embrace unification and ditch the idea of a separate "special" flat physics with its own "special" relativistic laws, just as we discarded the idea of a "special" flat Moon, or a special Earth.

14. SR Argument 14: "The deSitter and Brecher results disprove Newtonian theory, and what's left is SR"

14.1. Historical ballistic emission theory

In the simplest emission theory of light ("Ballistic Emission Theory", "BET"), light-corpuscles are is "thrown" from a moving body at a definite speed of c_{BODY} , and then continue indefinitely at that same speed until they are absorbed by a receiver.

The features that gave BET credibility as the default application of Newtonian theory to light were that **(a)** it seemed to agree with Newton's "corpuscular" description of light, **(b)** it gave the appropriate relativistic aberration formula (section 4), and **(c)** it generated the right Doppler relationships for NM (with only one emitting body, the BET calculations are basically the same as those for an aether calculation, in which the speed of light is fixed with respect to the emitter).

Where BET was terrible was that it offered no possibility of reconciling Newtonian theory with a wave-compatible description. Wave theory (before we take into account nonlinear effects where the signal energy distorts the metric) requires a lightsignal to move at a speed entirely dictated by the local properties of the region it moves through. Even if we take into account nonlinear effects, a signal moves through a region speed dictated by the local region's properties, modified by the signal's properties. Two identical signals that arrive in a region will still propagate identically, regardless of any different physical characteristics (motion, gravity, etc.) of the distant sources that originally spawned them.

Under BET (however), the signal from a distant approaching star will move towards us faster than one from an identical distant receding star, allowing identical signals to travel at *different* speeds and overtake each other along the same path, depending on the properties of distant past events.

14.2. The de Sitter experiment, 1913

The eventual experimental disproof of BET by **Willem de Sitter** (1872-1934) in 1913 [144], [145] (replicated by Kenneth Brecher in 1977 [146]) involved observations of double-star systems. If we were in the rotation plane of a binary star system, and the two stars were identical, there would be a moment where the stars were the same distance from us, with one star (**S₁**) approaching at v m/s, and the other (**S₂**) receding at v m/s. In a pure BET model, the signal from the approaching $\mathbf{S_1}$ would travel towards us at the faster speed ($c+v$), while the signal from the receding $\mathbf{S_2}$ would travel at the slower speed ($c-v$), with the final discrepancy for the two signals being proportional to the distance of the stars from the observer. Half a cycle later, the stars would have exchanged positions, with $\mathbf{S_2}$ approaching and $\mathbf{S_1}$ receding. For signals emitted in the new alignment, $\mathbf{S_2}$'s signals would now be faster and $\mathbf{S_1}$'s slower. The faster, later $\mathbf{S_2}$ signal would eventually catch up with and overtake the earlier, slower $\mathbf{S_2}$ signal, and if an observer was at the exact overtaking point, they would see the star to be in both positions at once. For greater distances, we would see an increasingly scrambled mess of different signals originating from different cycles (Einstein: "*the emission theory would lead to phase relations such that the propagated light would be all badly 'mixed up'* " [147]).

DeSitter's first paper pointed out that when we observed double-stars, their images always seemed to appear politely well-behaved, regardless of distance. If the speed of signals c' was constant throughout their journey, then this fixed speed couldn't be $c'=c+v$. A second (perhaps slightly grumpy) follow-up paper addressed an objection raised that perhaps the fixed speed of the light along its entire journey was not *totally* dictated by the speed of the source, but fractionally affected. De Sitter responded that if we wrote a fractional dependency as $c'=c+kv$, then to explain the well-behaved observed appearance of double-stars at such very great distances, the constant of proportionality k would have to be so absurdly small that it wasn't really credible to set its value to anything but zero.

We therefore knew, *experimentally*, that "the" speed of light (assuming that there *was* a single speed of light for a signal crossing astronomical distances) did not appear to show any detectable global, uniform dependency on the speed of the original source. [i] [ii]

14.3. "Independent of" vs. "not completely dictated by"

The language used in the deSitter/Brecher arguments was a little misleading. If a paper asks "*Is the Speed of Light Independent of the Velocity of the Source?*", and concludes "*no*", then we might be entitled to think that an experiment has shown that there is *no dependency whatever* between the speed of the signal's propagation and the speed of the source, even at shortish ranges.

This misleading word in the de Sitter and Brecher papers was the word "*the*" in "***the** speed of light*", as both authors presumed that there was *only one* speed of light that applied over the total signal path. In general language, we might say,

> "*No, that's wrong … there **is** a dependency of the speed of light on the velocity of the source, but it's only significant in regions where the gravity of the source dominates. And it's not dependent on the fact that the source **is** the source (apart from the proximity issue), because similar dependencies hold for all other masses that the signal interacts with along its journey, including the mass of the planet that the receiving telescope is resting on, the Earth.*"

If someone wanted to take issue with that explanation by objecting that the de Sitter/Brecher experiments had demonstrated the absence of *any* dependency, then they would probably not have understood the nature of the experiments.

> The de Sitter/Brecher experiments showed that if a lightsignal was assumed to propagate **at a single uniform globally-defined speed** along the entire length of its journey, that this hypothetical single fixed speed could not be visibly dependent on the speed of the source.
>
> The experiments did *not* suggest that there was no short-range dependency of the speed of light on the velocity of the source (although the papers' titles and abstracts made it seem otherwise).

14.4. Short-range signal-speed dependencies are compulsory

It is now an important part of mainstream theory that there should be a short-range signal flight-time effect for binary stars. When we watch the final death-throes of binary systems, as they inspiral and merge, we see the stars' mutual rotation rate speed up and accelerate towards the end, before their merger (with the upward sweep in frequency known as a "tweet"). In order to move from an equilibrium orbital state to a state of mutual collapse, the stars must somehow be shedding large amounts of rotational energy, and the current explanation of this is that as the stars' oscillations become more extreme, they radiate progressively more energetic gravitational waves, allowing the stars to move down into tighter and tighter mutual orbits until they start to merge (Thorne 1994, chapter 10 [22], Hulse and Taylor [132].

i … meaning that if the signal has a *single*, fixed speed along it's entire path, this is not dictated by the speed of the source. Which is not quite the same thing as saying that there's not a short-range dependency if signal-speed varies.

ii In a gravitomagnetic theory, the speed of light is influenced by the speed of the emitter as a proximity effect, but no more so than the speed of any other masses that it encounters along its journey.

In this scenario, the rotation plane around a double-star system contains an outward-spreading double-spiral gravitational wave, with the leading and trailing parts of each pulse of the wave having an accelerative, and then a decelerative effect on lightspeeds. When each star approaches us, its bow-wave has an accelerative effect on nearby light, pushing it towards us, and as it recedes again its associated gravitational signal drags light away from us. As the light travels further away from the double-star system, these fields become weaker, the angular separation between the two stars reduces, and the distant accelerative and decelerative field-effects progressively merge together, until the system effectively appears as a single gravitational point-source, whose merged, averaged gravitomagnetic effects are purely decelerative, appearing as the increased gravitational pull of the system due to its kinetic energy.

> The (varying) one-way speed of light *does* depend on the speed of the source … along with the speed of the receiver, and the speeds of any other bodies that it might encounter along its journey, with the strengths and spatial ranges of these effects being functions of the bodies' various gravitational field strengths.
> However, the effect on total signal flight-times will not scale with distance. If the source is a high-mass star, its individual significant region of influence will be greater than if the source is a small moon, or a particle of interstellar gas.

14.5. Making Newtonian theory wave-compliant

Suppose that we ***forced*** Newtonian optics to be wave-theory compliant? To assign only one speed of light (for a specified direction) to every point in space, we'd need light sent between two bodies to be *emitted* at c_{EMITTER}, but to be *received* at c_{RECEIVER}, and to somehow transition smoothly between these two speeds along the lightpath. Within a classical field theory, the light would start out moving at c_{EMITTER}, transition to some sort of averaged environmental-field-related speed along its journey, and then change speed again to c_{OBSERVER} at the end of its path. We would require the speed changes to be describable as a field, in which a receding body pulls light more strongly, and an approaching body pulls light more weakly. In other words, the field gradients will be *gravitomagnetic*, giving us a gravitomagnetic theory of light.

In other words, in order to "fix" Newtonian optics with respect to signal flight-times and wave theory compatibility, we need to turn it into a curved-space or curved-spacetime model, in which the motion of masses has gravitomagnetic side-effects. This is in broad agreement with the result from section 4.5 that the Doppler equations for Newtonian theory do not fit (or generate) flat spacetime.

This updated wave-compatible iteration of Newtonian theory would bypass the deSitter and Brecher objections … but would no longer be emission theory.

14.6. Summary

> The de Sitter-type experiments disprove the flight-time predictions of simple emission theory superimposed on a flat background: they do not disprove the Newtonian *relationships for energy*, implemented on a *curved-spacetime* background.

> They do not rule out proximity-dependent dragging effects, and in fact proximity-based dependencies between the speed of light and the speed of the emitter (and/or any other nearby bodies) must exist for double-stars to be capable of radiating gravitational waves.

15. SR Argument 15: Aether theories are wrong, therefore special relativity is right

A common "educational" narrative is that Nineteenth-Century physicists believed that the Earth's rotation and orbit around the Sun should cause it to move with respect to the aether, causing daily and yearly variations in the "aether wind", that the 1887 Michelson-Morley experiment [78] showed that no such effect existed, and that this threw physics into a state of disarray until Einstein came along and created special relativity to explain the result. [i]

15.1. Aether/either

While the modern educational position on aether theories has been described as defining "*the aether*" as being "*the thing that Michelson and Morley disproved*", it is better to consider "aether theory" as a collective noun. The late Nineteenth Century saw an embarrassment of different flavours of aether theory in circulation, promoted by their respective creators. Did particles push the aether out of the way, or did the aether permeate particles? Was it denser inside particles than outside? Did the density variation extend outside particles? Was it dragged along? Or was it undragged, and the density variation and flow merely *mimicked* dragging? Were there "aerodynamic" aether effects? Was there a relative aether-density gradient *across* or *around* a moving particle? Different theorists pitted their intuitions against each other, making different declarations as to how they felt an aether *ought* to interact with moving matter.

By the early Twentieth Century, even some of the biggest proponents of aether theory appeared to be getting exasperated with the situation. Aether theory (as a general subject) appeared no longer to be capable of making falsifiable predictions ... or rather (as with the case of string theory in the early Twenty-First Century), it was capable of making *too many* predictions. It could be used as a general modelling approach with multiple variables that could be fitted to almost any physics one could think of with a suitable choice of coefficients. And if there was a hypothetical behaviour that *didn't* already have a matching aether-theory variation, one could create one.

What one needed was an auxiliary principle or set of principles that – from basic logic – could eliminate the clutter and tell us *which* aether theory was required. And once one *had* those basic guiding principles, it was arguably the principles that were defining the physics, rather than the idea that "there is an aether".

George Francis FitzGerald (1851-1901) suggested that if matter moving with respect to an aether underwent a physical change of shape, then this could confound any attempts to measure the aether wind. [ii] Lorentz later independently came up with the same concept, and argued that it atoms were held together by electromagnetic forces in equilibrium, perhaps it was reasonable that lightspeed asymmetries might cause a shortening of $\sqrt{1 - v^2/c^2}$ (which the atoms themselves wouldn't notice). Lorentz promoted and wrote papers developing the idea (culminating in a major paper in 1904), [9] and with multiple theorists onboard and actively developing the idea (including **Henri Poincaré**), [80] Lorentz aether theory ("LET") and the Fitzgerald-Lorentz contraction were considered to be in serious danger of being right.

---

i The 1887 experiment *does* appear to have influenced Lorentz and Poincaré, but on the occasions that Einstein was asked about the influence of the M&M experiment, his response was that it hadn't been an influence. §9 of the 1905 Einstein paper does mention "*Lorentz's theory of the electrodynamics of moving bodies*", so we can suggest an *indirect* influence.

ii Oliver Lodge (1893), [79] page 749: "*… ingeniously suggested by Professor Fitzgerald, viz., that the cohesion force between molecules, and, therefore, the size of bodies, may be a function of their direction of motion through the ether …* "

15.2. Losing the aether: 1905

Einstein's 1905 electrodynamics paper theory took the essence of Lorentz' 1895/1904 aether model (the Lorentz factor) [i] rederived it by combining global *c* with relativity, and ignored the (now unnecessary) aether aspect, pointing out that one could construct what appeared to be a full theory of inertial physics that corresponded broadly to Lorentz theory around very basic considerations, without having to hypothesise the existence of any physical medium for light.

The space of the special theory (as made explicit by Minkowski spacetime in 1909) did not have the usual properties assigned to "aether" rules. It was not interactive, one could not establish the existence of flows or currents, it had no pressure or density variations, and it didn't divert or change properties when matter passed through it, or when it passed through matter.

15.3. General relativity considered as an aether theory

Some physics people weaned on the usual educational narrative that "relativity disproved the aether" often seem to be surprised to be told that Einstein gave a lecture in 1920 in Leyden on general relativity *considered as* an aether theory ("*Aether and the theory of relativity*". [81] Einstein's point was that the spacetime of general relativity now had most of the properties that one would associate with an aetheric medium – it was a variable-density physical entity that could carry signals and information by physically distorting, if cg was finite it acted as a carrier for distortional waves (gravity-waves), and if one removed the medium connecting two regions, nothing could get between those regions (the result of removing the medium not being "empty space" but "no space at all"). [ii]

> Einstein, "*Ether and the Theory of Relativity*" (1920): [81] "*… the special theory of relativity does not compel us to deny ether.* "
>
> "*The special theory of relativity forbids us to assume the ether to consist of particles observable through time, but the hypothesis of ether in itself is not in conflict with the special theory of relativity. Only we must be on our guard against ascribing a state of motion to the ether.* "
>
> "*… according to the general theory of relativity space is endowed with physical qualities; in this sense, therefore, there exists an ether. According to the general theory of relativity space without ether is unthinkable; for in such space there not only would be no propagation of light, but also no possibility of existence for standards of space and time (measuring-rods and clocks), nor therefore any space-time intervals in the physical sense. But this ether may not be thought of as endowed with the quality characteristic of ponderable media, as consisting of parts which may be tracked through time. The idea of motion may not be applied to it.* "

With general relativity applied to larger-scale problems involving galaxies and galaxy-clusters, mathematics developed for "particulate medium" problems involving, say, fluid flows, or pressure-densities, started to become relevant again. The one point that distinguished the GR aether from most traditional aether theories was that it was *not particulate*. Its classical structure was totally smooth at small scales, and one could not identify a distinguishable point in the aether and watch how it moved (any more than one can identify a distinguishable point on the perimeter of a circle and watch it over time to see if the circle is stationary or rotating).

---

i Einstein later said that he had not been aware of Lorentz' 1904 Annalen der Physik paper when writing his 1905 Annalen "Electrodynamics" paper. Although this *seems* slightly unlikely, Einstein did happily admit to having read Lorentz' earlier work, and gave credit to Lorentz' system twice in the paper, so … perhaps he wouldn't have anything obvious to gain by claiming not to have read the 1904 work, if it wasn't true.

ii Einstein's comment "*Newton might no less well have called his absolute space 'Ether'* " and his ascribing aetheric interpretations to Newton's *followers* suggests that Einstein might have been unaware of the contents of "*Opticks*".

15.4. Non-particulate aethers

A common response to an explanation of Einstein's position on "GR aether" is that if an aether is non-particulate then it is not an aether, as historically, aether models have always been based on the idea that the medium is made up of little particles.

This is also not *quite* true. In Newton's *Opticks*, we find:

> Newton, ***Opticks***, Qu.21: "*… And so if any one should suppose that Æther (like our Air) may contain Particles which endeavour to recede from one another (for I do not know what this Æther is) and that its Particles are exceedingly smaller than those of Air, or even than those of Light: …*"

Although Newton *did* tend to treat his aether as being particulate (and explored the consequences), by adding an extra layer of distancing between himself and the idea that the aether *really was* particulate ("*if any one should suppose … for I do not know what this Æther is*"), Newton implicitly also opened the door to the concept of a potentially *non*-particulate aether … but, lacking adequate tools to investigate what this might mean, the idea was not explored further.

15.5. Quantum mechanics and cosmology

Quantum mechanics arguably turns empty space into a form of particulate medium.

> Taylor and Wheeler, **Spacetime Physics: 2nd edition** (1992) [38] " *But is space really empty? "Definitely not!" says modern quantum physics. "Space is a boiling cauldron of virtual particles. To observe this cauldron, sample regions of space much smaller than the proton. Carry out this sampling during times much shorter than the time it takes light to cross the diameter of the proton.* "

Taylor and Wheeler also argue that in cosmology, space is not empty. At large scales we can consider stars as "particles", possessing significant gravitational fields and associated curvatures.

15.6. Summary

The special theory showed that, since it was possible to reproduce the essential desirable characteristics of a Lorentzian aether from just the initial *design conditions* that had led to the idea of a Lorentzian aether, *those design conditions on their own* were sufficient to define the behaviour of physics, directly, cutting out the intermediate stage of requiring a physical medium to express them. The existence of a Lorentz aether was not *disproved* … but it was "*superfluous*" ("*an empty hypothesis*") and supposing that it *did* exist didn't obviously add anything of value to the model's existing physical predictions. [i]

However, once we went beyond the case of simple motion in straight lines, the idea of a "physical", interactive medium returned.

> "We know that there is no aether, therefore SR is correct" is not a sensible argument.
>
> Although special relativity arguably makes a simple *Lorentzian* aether redundant, general relativity reassigns spacetime some "aether-like" properties.

---

i Although special relativity was widely considered to be a more modern replacement for LET, a few physicists stayed with the idea that special relativity was *explained* by a Lorentzian aether, and justified keeping LET as a separate subject in case the two systems turned out to diverge when we tried to extend physics past the condition of simple constant-velocity rectilinear motion. Ives and Stilwell, who produced a notable test in favour of special relativity's relationships in 1938, [82] were specific about presenting their paper as confirming the predictions of LET rather than SR.

16.SR Argument 16: General relativistic effects

16.1. Historical overview

Much of what gets presented as the validated behaviour of Einstein's general theory does not rely on GR1916's specifics, but is an extrapolation of the observation, tested experimentally by **Galileo** (1564-1642), that (after we take into account complicating effects like wind resistance and buoyancy effects) different bodies seem to fall at the same speed in a gravitational field, regardless of what they are made of. Gravity appears to pull everything equally. [i]

This was a critically important assumption, if only for pragmatic reasons – if Nature was more complicated and pulled on different materials differently, we would not be able to derive a simple set of laws for gravitation. We would not be able to safely calculate planetary orbits without knowing what the different planets were made of.

Once we have this idea, almost everything else falls into place:

(a) If all bodies fall at the same rate, then light must also fall at that same exact rate. If light was affected *any less* by gravity (or not at all), then a box containing light would always be trying to fall more quickly than its contents, and would be partly buoyed up by the additional light pressure inside its upper surface. The box would fall more slowly. Similarly, if light fell *faster* than matter, the light would exert a stronger radiation-pressure on the box floor, pushing the box downwards and causing it to fall faster. If a box falls at the same rate regardless of its contents, then light, and electric and magnetic fields, and every other possible component of the box and its components must be affected by gravity in the same way. Gravity must be universal.

(b) As well as allowing simple equations for gravity, "universality" allows us to use classical theory to express gravity either as a simple field effect, or as a geometrical effect. If the field deflects *absolutely everything* by the same amount, then as far as a region's inhabitants are concerned, the deflection can be considered a property of the region's *space*, and can be modelled as an apparent spatial distortion. We can then describe gravity using **metric theories**, which model the properties of a region's effective geometry according to the behaviour of light-beams.

(c) The deflection of light produces gravitational lightbending, gravitational lensing and gravitational shifts in the energy of light. If light "falls" in a gravitational field, it must undergo a deflection, [12] so gravity bends light-beams. (Newton, section 8.5) If an initially-horizontal beam is deflected downwards, it applies a downward force, and has downwards momentum, and to someone below it, has an increase in energy. For a falling box to increase in energy by the same proportion when it falls regardless of its contents, the energy-change in light must match the Doppler shift relationship. If we drop a flashlight onto a detector, then if the flashlight is switched on at the last moment, its light-energy will be changed by a Doppler shift due to the flashlight's motion. If the flashlight is switched on at the *start* of its fall, then if the energy taken from the battery undergoes exactly the same change (regardless of how it crosses the region) the change must be the same as before. We can then calculate a *gravitational shift* by calculating the velocity-

i There is a *possible* exception for bodies with more complex interacting gravitational fields: gravitomagnetic theory suggests that if a spinning disc is dropped horizontally over one pole of a spinning dense star, and both have the same rotation rate, that the disc *might* fall differently depending on whether it co-rotates with the star or counter-rotates (*if* there's no compensating inertial effect). We can avoid the subject of gravitomagnetic complications by saying that all ***simple*** bodies are expected to fall identically regardless of their composition in a ***simple*** gravitational field.

change associated with a gravitational differential, and applying whichever Doppler relationship applies to motion shifts.

(d) If light in a box is deflected downwards by gravity, then if we *prevent* the box from falling (by placing it on a set of scales), the light inside the box will continue to be deflected downward, pushing against the box floor, and increasing the apparent weight of the box as reported by the scales. [i] The energy of the trapped light therefore contributes "weight", to the box, and is associated with gravitational mass.

(e) For all bodies to fall at the same rate, the force causing the object to accelerate (its gravitational mass) and the resistance to acceleration (its inertial mass) must have precisely the same ratio for all objects. This gives us the principle of equivalence of inertial and gravitational mass – and applying Occam's razor and simplifying further, instead of saying that matter has two properties – inertial and gravitational mass – we can then say that there is only a single property, "mass", that has inertial and gravitational *aspects*.

(f) Since trapped light has *gravitational* mass (paragraph (d), above), it must then also have *inertial* mass. Calculating this inertial mass using either the Newtonian Doppler equations or some other relativistic equation-set (like SR's) gives us the relationship $E = mc^2$. Since light can be converted into other forms of energy, if we require gravitational bodies such as stars not to change their gravity when they convert energy between forms, we require *energy itself* to have gravitational mass and inertial mass when confined. If we wind up a clockwork motor and place it on a set of scales, it should weight (imperceptibly!) more than if it was unwound.

(g) The gravitational shifting of light means that if we have a satellite in deep space emitting a steady signal, which we receive in on Earth with a gravitational blueshift due to the increase in energy, the signal will be received with a higher frequency than it was sent with. The only way we can see the signal's peaks to be arriving at a faster rate than the satellite generates them … and for this situation to be able to continue *indefinitely* … is if our reference-clocks are genuinely running slower than the satellite's, giving us gravitational time dilation.

In other words, Galileo's idea makes gravitational shifts unavoidable, and gravitational shifts make gravitational time dilation unavoidable. Local timeflow must run at a rate that depends on local gravitational field density. A gravitational field must not just warp spatial coordinates, it must also warp time coordinates, and any metric theory of gravity must be based on the idea of *spacetime* curvature. If gravity distorts everything by the same amount then if our laboratory is free-falling in a uniform gravitational field, none of our instrumentation will be able to detect the existence of the background field gradient. On the other hand, if we are free-falling and decide to force our laboratory to move differently, by applying force, then since we will see objects in our lab experiencing identical gee-forces regardless of their composition, there will be no way to distinguish between our "artificial" gravity and "real" gravity.

Given that pretty much all of this chain of logic is unavoidable, perhaps the surprising thing is that progress in gravitational physics was so wretchedly slow over the last three hundred years. Why did we have to wait for Einstein to come along before we got a proper massenergy calculation? Why did nobody before Einstein seem to point out the gravitational time dilation

---

i The idea of radiation pressure was known in Newton's time: Newton, *Principia,* Book III (page 491), " *Kepler ascribes the ascent of the tails of the comets to the atmospheres of their heads and their direction towards the parts opposite to the sun to the action of the rays of light carrying along with them the matter of the comets tails …* "

effect? Why was most of this work not already done by the mid-Nineteenth century? Was nobody other than Einstein paying attention? [i]

And we can go further. If we continue with this chain of logic, the result that all inertial masses must have associated gravitational mass, and therefore curvature, means that we have to replace the old Minkowski metric with a more advanced relativistic acoustic metric. While it is easy with hindsight to be incredulous that Nineteenth-Century physicists failed to complete the chain of logic that that takes us to GR-level effects, we have to remember that even our own generation of theorists still suffered from a failure of conviction that prevented them from continuing the chain beyond "SR-based" GR and on to the next chapter of the story.

16.2. The "four tests" of Einstein's general theory

The three tests of general relativity suggested by Einstein ("Relativity", [65] Appendix 3) were gravitational shifts, gravitational light-bending and the rotation of the alignment of Mercury's perihelion. The Shapiro time delay is generally considered the fourth test.

The gravitational redshift effect

The gravitational shift on light is a very general effect that appears in pretty much any theory of gravity once we assign light wavelike properties. Einstein's famous 1911 paper on gravitational shifts [147] predates GR1916, and (to keep things simple) derives a "Newtonian" version of the effect rather than using special relativity. [ii] Given that the argument for gravitational shifts was previously laid out by **John Michell** (1724-1793) in a letter published in the Journal of the Royal Society in 1784, [11] it would be not just mathematically but historically perverse to suggest that gravity-shifts wouldn't have been expected or predicted without SR/GR (Michell's study of gravitational shifts even suggested using a prism as a spectrometer to identify gravitational shifts in starlight). [iii]

The main difference between the gravitational shifts predicted by Newtonian theory and GR1916 is that the Newtonian redshift is redder than the SR version by a further Lorentz factor (using $E'/E=(c-v)/c$ rather than special relativity's $E'/E=\sqrt{(c-v)/(c+v)}$). Since the gravitational differential for most bodies is a tiny fraction of the speed of light, and one normally hopes v to be at least a few percent of the speed of light to be able to safely tell the two predictions apart, it's not obvious that there's yet any primary data that can convincingly distinguish between the gravitational shifts of GR1916 and Nineteenth-Century Newtonian gravity.

---

i We have just reproduced most of the phenomenology of general relativity (other than the inclusion of SR), without introducing the concept of relativity. The results of a similar chain of reasoning based on the unrestricted, "general" principle of relativity converge. There are also other parallel logic-chains, such as the observerspace principle that physics should be seen to be consistent, and, more recently, the holographic principle, information theory (and statistical theory, used in QM), all of which seem to be trying to converge on a single massively-dual system of physics. The resulting common system of physics converged on would not include special relativity.

ii Einstein, 1911: "*To avoid unneccessary complications, let us for the moment disregard the theory of relativity, and regard both systems from the point of view of kinematics, and the movements occurring in them from that of ordinary mechanics. …*". Einstein goes on to derive the existence of gravitational shifts, gravitational time dilation and then (via Huygens' principle) gravitational light-bending, *without* introducing the updated Doppler relationships of special relativity.
This is useful, as it shows that not only was it *theoretically possible* to derive these basic effects before general relativity (1916) and without using SR, but that Einstein actually did it.

iii Michell's piece also derives the gravitational horizon radius, and suggests a survey to work out statistically the proportion of matter in the universe that is "dark", by counting the number of known double-stars where only one circling partner is visible. This idea is usually considered not to have been thought of until the Twentieth Century.

What about secondary effects?

- **Horizon behaviour:** Both equations generate a horizon at exactly the same radius (r=2GM/c^2), with GR1916 the horizon is *absolute*, with a perfectly zero surface temperature, with NM the horizon is an *effective*, observer-dependent *relative. acoustic* horizon, able to leak and radiate indirectly, with a statistical temperature above zero. The SR-based version is incompatible with quantum mechanics (which requires horizons to be leaky), while NM version seems to be at least in general agreement with QM.
- **Energy-loss:** With the SR equations, a signal passing onto and out of a gravity-well, and undergoing a successive blueshift and redshift for the same velocity differential, will emerge with precisely the same energy it started out with. With the NM version, the blueshift times the redshift gives $E'/E=(c - (-v))/c \times (c - v)/c = 1 - v^2/c^2$, a Lorentz-squared redshift over the round trip. Over cosmological scales, a signal riding a switchback series of gravitational hills and troughs will be expected to reach us with a cumulative redshift as a function of distance. In a fully gravitomagnetic theory, this redshift looks just like (and can be treated as) a recession redshift, so while the SR Doppler relationship is a better fit to a static fixed universe (which is how Einstein originally saw GR-based cosmology [90]) the non-SR version is a better fit to an expanding-universe cosmology (section 44).

Deflection of light by gravity

This is also an incredibly general effect, and arguably the default behaviour for light and gravity (it appears in Newton's *Opticks*).

> Newton, ***Opticks***,: " *Query 1. Do not Bodies act upon Light at a distance, and by their action bend its Rays; and is not this action (**cæteris paribus**) strongest at the least distance?* "

We can calculate the gravitational deflection of light using Huygens principle, either by arguing that the speed of light is slower near a gravity-source due to increased spacial distance per unit volume (curvature of space), or calculating the effect due to gravitational time dilation (curvature of time, Einstein 1911 [12]), Both calculations, curved-space and curved-time, give essentially the same result. At this point, we have what appears to be a "decision-fork" in how a theory is to be designed: do we say that the two effects are dual, and are the same effect, described in two different domains? Or do we say that the spatial- and temporal-curvature effects are separate and cumulative? Einstein argued (in the 1916 theory) that they were cumulative, meaning that the 1916 theory predicted twice the angular deflection of starlight by the Sun, a result that was confirmed by Eddington's team in 1919. [61]

This was a genuine prediction made in advance of the data being known, and deserves some credit. However, we have to also understand that, after Einstein had presented the idea of gravitational time dilation in 1911, which he'd pointed out was also a result of Newtonian calculations, Newtonian gravity, *updated with the new additional 1911 idea of gravitational time dilation* could claim the same result.

The perihelion effect on Mercury

The planet Mercury has quite an elliptical orbit, and is the closest planet to the Sun, meaning that it experiences a larger proportional variation in gravitational field intensity as it orbits the Sun than the other planets. As the planet swings around the Sun, the alignment of its ellipse changes, like the movement of a pen in a "Spirograph" set (perihelion precession).

The *impression* given by Einstein (in "*The experimental confirmation of the general theory of relativity*"), is that Newtonian mechanics predicts no effect.

> Einstein, "Relativity …", [65] Appendix 3 " *According to Newtonian mechanics and Newton's law of gravitation, a planet which is revolving around the sun would describe an ellipse … If instead of Newton's law we insert a somewhat different law of attraction, we find that … the angle … from perihelion … to perihelion … would differ from 360 degrees* "

Einstein gives a value of 43 seconds of arc per century for the GR perihelion shift.

In fact the perturbing influences of the other planets (and other effects), calculated under Newtonian theory, predict a perihelion shift that is rather larger than the GR effect, just not quite enough to agree with the historical astronomical records. The observed shift is actually supposed to be around ~574 arcseconds per century, of which around ~532 are already accounted for without involving general relativity. This leaves a shortfall of ~43 arcseconds per century, which GR then explains as being due to curvature effects.

The comparison is therefore not (as Einstein's description implies) the difference between no perihelion shift at all and the validated GR shift. The standard Newtonian prediction manages to account for ~92.5% of the observed perihelion rotation, and the GR effect accounts for the missing ~7.5% . Significant, but not quite as compelling as saying that Newtonian theory predicts no effect at all.

We also have to bear in mind that the traditional Newtonian calculation does not seem to take into account that fact that, even under Newton's system, the lightbeam geometry of the region is not "flat", so even *without* any planetary perturbing influences, we would expect the perihelion to advance in a way that is not obviously taken into account in the normal Keplerian "ellipse on a flat background" calculations, shrinking the 7.5% shortfall. If we add the Newtonian curved-space correction, and then an additional correction for Einstein's 1911 Newtonian gravitational time dilation effect, it's not obvious that the result of NM plus curved spacetime would be appreciably different (or any different!) from the GR prediction.

While Einstein's prediction was obviously a scientific advance, the main difference between the results of GR1916 and "updated NM" would seem to be in the Doppler equations used for motion shifts and gravitational shifts, which in GR1916's case are the later SR set. It's not obvious what part (if any) this difference makes to the Mercury calculations.

Shapiro effect

The Shapiro effect is the portion of the time-delay of a light-signal passing through a gravitational field that is not due to light-bending. [66] If two identical stars in a binary system orbit a common centre, and we aim a beam of light along their central orbital axis, the beam will not be deflected to any side (a spatial straight line), but will still take longer to pass through the region.

Shapiro was inspired by the idea of treating gravitational fields as a variations in refractive index, an approach that Einstein had used in 1911 to calculate lightbending, [12] and which Newton had also used, albeit with an unfortunate error.

Newton explicitly describes this approach in Opticks:

> Isaac Newton, ***Opticks*** (words in square brackets have been inverted to match our current understanding of physics):
>
> *Qu. 19. Doth not the Refraction of Light proceed from the different density of this Æthereal Medium in different places, the Light receding always from the ~~denser~~ [rarer] parts of the Medium? ...*
>
> *Qu. 20. Doth not this Æthereal Medium in passing out of Water, Glass, Crystal, and other compact and dense Bodies into empty Spaces, grow ~~denser~~ [rarer] and ~~denser~~ [rarer] by degrees, and by that means refract the Rays of Light not in a point, but by bending them gradually in curve Lines? And doth not the gradual ~~condensation~~ [rarefaction] of this Medium extend to some distance from the Bodies, and thereby cause the Inflexions of the Rays of Light, which pass by the edges of dense Bodies, at some distance from the Bodies? ...*
>
> *Qu. 21. Is not this Medium much ~~rarer~~ [denser] within the dense Bodies of the Sun, Stars, Planets and Comets, than in the empty celestial Spaces between them? And in passing from them to great distances, doth it not grow ~~denser~~ [rarer] and ~~denser~~ [rarer] perpetually, and thereby cause the gravity of those great Bodies towards one another, and of their parts towards the Bodies; every Body endeavouring to go from the ~~denser~~ [rarer] parts of the Medium towards the ~~rarer~~ [denser]? … I see no reason why the ~~Increase~~ [Decrease] of density should stop any where, and not rather be continued through all distances from the Sun to Saturn, and beyond. And though this ~~Increase~~ [Decrease] of density may at great distances be exceeding slow, yet if the elastick force of this Medium be exceeding great, it may suffice to impel Bodies from the ~~denser~~ [rarer] parts of the Medium towards the ~~rarer~~ [denser], with all that power which we call Gravity. "*

With Newton's original system, light accelerated as it entered a gravitational region and slowed again as it exited. By wrongly associating the energy-gain with a *red*shift due to the light being stretched, we could use the wavelengths of light to map space, and conclude that since fewer wavelengths fitted into a more strongly "gravitational" region, the region was spatially less dense, and the aetheric medium therefore had to be described as being *displaced* by matter (*see:* Newton's description, above, before and after correction).

If we substitute the correct *proportional* relationship between energy and frequency, Newton's description inverts. In the corrected description, light entering a gravitationally denser region now blueshifts as it increases in energy,i the wavelengths shorten, and the region as mapped by light now contains an excess of space compared to the background. The aetheric medium is now not *displaced* by a gravitational field, *it is* the gravitational field (section 15.3).

Newton's "refractive index gravity" model, if updated with the correct frequency~energy relationship, immediately gives the Shapiro effect, which wasn't predicted in the context of GR until the 1960s. [66] If Nineteenth-Century theorists had corrected Newton's model to fix the original inverted lightspeed and energy relationships, and bring it into line with Huygens' principle (rather than discarding this section of the theory), the Shapiro effect could (and should) have been predicted in the Nineteenth Century. [i]

---

i … this is assuming that anyone would have cared. There seems to have been a tendency in the C18th and C19th to be dismissive of physical predictions whose verification was impractical with available technologies, so it's possible that some theorists might have been aware of the idea, but considered it too inconsequential to be worth documenting.

16.3. Gravitational lensing

Once we have gravitational lightbending, gravitational lensing is an obvious consequence.

Gravitational lensing (Chwolson, 1924 [68]) was originally considered a trivial and theoretically unimportant adjunct to the physical prediction of gravitational lightbending (when Einstein was cajoled into submitting a paper on the subject, he included a note to the editor apologising for submitting anything so petty). Although Einstein's paper said there was no serious hope of testing the effect, deep-field astronomy has now given us views of so many galaxies that we have found examples where pairs of galaxies line up, and the more distant galaxy appears to us as a series of distorted images surrounding the nearer galaxy, with shared spectral characteristics telling us that they are different views of the same distant object (*e.g.* the "Einstein cross" [83]). Gravitational lensing is now important to humans because it lets us look at a photograph and see lightbending effects in action, in an easily-understandable form, and the ability to see further into deep space using lensing is likely to tell us more about the distant early universe. However, in terms of purely theoretical significance, it doesn't obviously add anything much to the subject of lightbending.

16.4. Summary

We cannot interpret these successes of general relativity to mean that special relativity is right.

- Firstly, it is not obvious that special relativity *ought to exist* as part of general relativity. The GPoR and a range of gravitational arguments (c_g=c, sections 8.2, 8.3) require gravitomagnetism to be an intrinsic part of relativistic physics, while support for special relativity requires that complicating gravitomagnetic and light-dragging effects (which violate SR's flat spacetime assumption) not exist. SR requires particles to have zero associated curvature: the principle of equivalence requires them to have non-zero curvature. SR physics is not a legal logical subset of gravitational physics: Einstein's inclusion of SR physics as a limiting case of GR1916 is geometrically invalid.

- Secondly, the "good bits" of general relativity appear to be founded on the principle of equivalence, the Mach-Einstein general principle of relativity, and on other basic laws and principles that do not use or rely on special relativity.

- Thirdly, the more "problematic" aspects of general relativity – the incompatibility with full gravitomagnetic theory, the incompatibility with quantum mechanics, the incompatibility with modern cosmology and the principle of equivalence, and the appearance of singularities … appear to be side-effects of GR1916's attempt to accommodate SR.

While there *are* physical differences between an NM-based and an SR-based implementation of the principle of equivalence (primarily a different gravitational shift characteristic), these do not yet seem to have been demonstrated in the four tests.

> The results of the standard "four tests" of general relativity appear to be results of the principle of equivalence, and do not obviously depend on Einstein's 1916 implementation of a general theory, or the fact that GR1916 is designed to incorporate special relativity.

> Including special relativity makes GR1916 logically inconsistent, and experiment cannot be used to prove an inconsistent theory correct. Rather than general relativity being a success *because* of SR, it might be that it's managed to be relatively successful *despite* it.

17.SR Argument 17: "We wouldn't have GR without SR"

17.1. Nineteenth Century

Newton's idea of gravitational effects as the consequence of a variable-density gravitational aether can be expressed in modern terms as a "curved space" theory of gravity, and if we then recognise that gravitational shifts require gravitational time dilation, we get a *curved spacetime* theory. Turning this heuristic description into pure geometry requires a mathematical framework.

Geometrical rules and notations for describing curved space started to be explored in the Nineteenth Century by **Karl Friedrich Gauss** (1777-1855) who published a seminal paper in 1827 [84]. Gauss encouraged **Bernhard Riemann** (1826-1866), to work on the problem, and Riemann gave a critical address on curved spaces in multiple dimensions in 1854. Riemann's presentation was published in 1868, [85] and was then translated by **William Kingdon Clifford** (1845-1879) and published in Nature for a wider English-language audience in 1873. [86]

17.2. William Kingdon Clifford

Clifford argued that although physical space appeared to us to obey the traditional "undistorted" rules of Euclidean geometry, and Occam's Razor suggested an initial assumption that that space was everywhere the same (unless we had reason to believe otherwise), this didn't mean that the "uniform" default geometry had to be flat. If the universe had constant curvature across space, and/or the curvature was varying over time, slowly, we wouldn't necessarily have noticed. [i]

> W.K. Clifford, "On the Bending of Space" [87] " *When we assert that our space is everywhere same, we suppose it of constant curvature (like the circle as one- and the sphere as two-dimensioned space); when we suppose it homoloidal we assume that this curvature is zero (like the line as one- and the plane as two-dimensioned space). … This real existence, of which it is clearly impossible for us to be cognizant, we postulate as a result built upon our experience of what happens in a limited portion of space. We may postulate that the portion of space of which we are cognizant is practically homoloidal, but we have clearly no right to dogmatically extend this postulate to all space. A constant curvature, imperceptible for that portion of space upon which we can experiment, or even a curvature which may vary in an almost imperceptible manner with the time, would seem to satisfy all that experience has taught us to be true of the space in which we dwell.* "

Clifford then argued that local variations in curvature might already be playing an overlooked part in current physics, without our recognising the possible geometrical explanation:

> *W.K Clifford, "On the Bending of Space"* [87] " *We may... be treating merely as physical variations effects which are really due to changes in the curvature of our space; ...*
> *… We may conceive our space to have everywhere a nearly uniform curvature, but that slight variations of the curvature may occur from point to point, and themselves vary with the time. These variations of the curvature with the time may produce effects which we not unnaturally attribute to physical causes independent of the geometry of our space. We might even go so far as to assign to this variation of the curvature of space 'what really happens in that phenomenon which we term the motion of matter.'* "

Although Clifford didn't seem to explicitly mention gravitation, he did write that *uniform* curvature was a condition of being able to freely translate geometrical structures without side-effects. If we were to suggest *non*-constant curvature and ask, "*Are there any circumstances in which bodies in otherwise empty space are known to **not** move freely in straight lines with constant orientation?*", the answer would be, "*Yes, in the presence of a gravitational field*".

---

i Lobachevsky discussed a hyperbolic universe in 1829, [88] Zollner a spherical universe in 1872, [89] Einstein 1917. [90], [91] The reason why Clifford's work isn't better known (and why he didn't get the chance to develop his ideas further, despite appearing to have the appropriate mathematical skills and being a world authority on curved space) is that he passed away at the age of 33, just a few months before the publication of his first and only finished book. Clifford died on 3rd March 1879 – Einstein was born just eleven days later, on the 14th March 1879.

17.3. Popularisation of spacetime concepts

The job of judging which ideas were in circulation in the late Nineteenth Century is made more difficult by the fact that some experts had a habit of disseminating their work to a wider audience by giving public lectures, the contents of which were not always published in print form. George Fitzgerald's notion of length contraction, which he gave in lectures, appears to have almost no supporting documentation, and the only record of Clifford's lecture "*On the Space-Theory of Matter*" is a brief published abstract. [69]

The idea of *multiply-connected* topological space is explained in H.G. Wells' 1895 short story "The Remarkable Case of Davidson's Eyes", which explains the principle of a "wormhole". Wells also popularised the idea of time as a fourth dimension via his stories *The Chronic Argonauts* (1888) "*Has it never glimmered upon your consciousness that nothing stood between men and a geometry of four dimensions—length, breadth, thickness, and duration—but the inertia of opinion …*", and *The Time Machine* (1895) "… *space, as our mathematicians have it, is spoken of as having three dimensions, … But some philosophical people have been asking why three dimensions particularly—why not another direction at right angles to the other three?—and have even tried to construct a Four-Dimensional geometry. Professor Simon Newcomb was expounding this to the New York Mathematical Society only a month or so ago*". [i] Wells' story "The New Accelerator" (1901) uses a mysterious potion as a plot device to explore what we might see if timeflow was variable. [ii]

By around 1900, the psychological vocabulary and component concepts necessary for a curved spacetime theory were not just current and known to mathematicians, they were already part of popular fiction. What had apparently been lost (to the science and math communities) was Michell's argument for gravitational shifts, [11] which should have led to the idea of curved time coordinates in the early-to-mid Nineteenth Century.

When Einstein published the general (and simple) argument for gravitational shifts in 1911, presenting gravitational time dilation as an unavoidable consequence, [12] it would have been like lighting a fuse. Mathematicians across Europe whose technical skills outclassed Einstein's might have seen the discovery as a challenge to see who could be first to add curved time (warped by gravity) to Riemann's curved space.

Einstein continued his attempts to develop a general theory (with initial technical help from his good friend Marcel Grossman), [95] and with feedback from various interested mathematicians and figures in theoretical physics, some of whom (such as Max Abraham), had their own competing theories.

17.4. The Great War (1914-1918)

The story of the general theory's development seems to have taken some odd twists and turns with the outbreak of World War One:

i Simon Newcomb's two lectures on curved and higher-dimensional spaces [92], [93] didn't actually associate a fourth dimension with *time* – Wells' understanding here seemed to be more advanced than Newcomb's.

ii Lewis Carroll's "**Sylvie and Bruno**" (1890) [94] also featured "An Outlandish Watch" that allowed one to experience earlier times, or to see time running backwards.

Erwin Finlay-Freundlich (1885-1964) had travelled to the Crimea to measure the bending of light during an eclipse expected on 21st August 1914, but ended up being interned during a period of heightening political tension, with the assassination of Archduke Ferdinand on 28th June 1914, and a state of war officially declared by Germany on 1st August. If this test had gone ahead, and revealed the "doubled" lightbending result – *before* Einstein had been able to predict it – then it would have proved both Newton *and* Einstein wrong. Doubled lightbending would have suggested that curved-space and curved-time lightbending effects were cumulative, requiring curved spacetime, and in turn suggesting a geometrical explanation for the Mercury perihelion anomaly before Einstein had finished constructing his general theory.

David Hilbert (1862-1943) corresponded with Einstein before the the field equations were finalised, and submitted his own version in 1915, triggering debates about priority (although it may be that Hilbert had added the equations to his paper while it was "in proof", to show a link between what he was doing and what Einstein was doing). [96]

Karl Schwarzschild (1873-1916), had been publicly pondering curved-space issues since 1900. With the outbreak of WW1, the patriotic Schwarzschild volunteered to join the army and was sent to the Russian front, but still managed to produce an exact solution to Einstein's tentative field equations in 1915, [97] before Einstein (before dying in 1916).

Over in England, **Arthur Stanley Eddington** (1882-1944) a Quaker, was preparing to declare himself a conscientious objector and spend the war "peeling potatoes" in a camp with his fellow Quakers. His boss Frank Dyson, wanted to avoid the associated scandal, and was keen to devise a special project for Eddington that would excuse him from war work. Eddington had become the *de facto* world expert in Einstein's general theory outside of Germany … partly because the effect of the wartime censor was to stop anyone outside mainland Europe from reading Einstein's wartime work, and Eddington seemed to possess the only known "outside" copy of Einstein's paper, posted to him from "neutral" Holland by his friend Willem de Sitter. [98]

Eddington seemed fairly indifferent to the idea of testing Einstein's theory, taking the position that the theory was simply right, and said at one point that if it had been up to him, he probably wouldn't have bothered … however, the plan to send Eddington off to measure Einstein's revised lightbending prediction during the forthcoming 1919 eclipse solved Dyson's immediate political problem.

Eddington returned to a world profoundly depressed by years of brutal warfare and further mass deaths caused by the subsequent 1918-1919 "Spanish Flu" pandemic. His result, with an *English* pacifist scientist confirming the revolutionary "new physics" of a *German* pacifist scientist, made for an uplifting internationalist "good news" story that went straight to the front page of The Times (in London) "***REVOLUTION IN SCIENCE / NEW THEORY OF THE UNIVERSE / NEWTONIAN IDEAS OVERTHROWN***" (7th November 1919), and earned an "***EINSTEIN THEORY TRIUMPHS***" heading on page 17 of the New York Times (10th November 1919).

If the War hadn't happened, and Freundlich's test had gone ahead, we might have seen Schwarzschild (or Eddington, or someone else) encouraged to produce a general theory that combined Einstein's general 1911 "gravitational time dilation" logic with the knowledge that spacetime curvature seemed to be real, before Einstein. A general theory designed by someone *other* than Einstein might not have been quite so likely to try to incorporate special relativity … a different "architect" might have been satisfied with just a reduction to Newtonian theory.

17.5. Role of special relativity

Lorentzian electrodynamics played a role in encouraging a number of mathematicians to start taking physics seriously as a potentially interesting sideline (by 1905, Minkowski had apparently been pondering the subject for a while and was quite taken aback to see Einstein’s paper).

Special relativity (and extended special relativity) also gave Einstein a platform from which to start investigating the phenomenology that a general theory would need to show. The association between an apparent gravitational field for a centrifuged clock and its SR time dilation would have emboldened Einstein to think that the “gravitational time dilation” idea – a pretty radical concept – was perhaps not such a stupid idea (although Einstein also considered the effect to be a result of Mach’s principle [i]).

17.6. Alternative timelines

The theoretical importance of special relativity in all of this is not obvious. Embedding SR directly into the specifications for his general theory certainly allowed Einstein to complete the theory without having to rederive SR-style inertial physics all over again, from scratch … but if he’d never produced special relativity, wouldn’t his argument that gravitational physics needed to reduce to inertial physics over small regions simply have used Newtonian mechanics for its default inertial physics, instead?

The appearance of the time dilation effect under SR may have encouraged Einstein to think of time as variable, with “extended SR” encouraging him to think of gravitational shifts as a correct result … but if Lorentz’ work hadn’t captured Einstein’s attention and prompted the 1905 electrodynamics paper, then, casting around for different projects to work on, it’s conceivable that Einstein might have tried gravitational theory instead, and found the 1911 gravity-shift arguments sooner.

Even if not having special relativity might have slowed Einstein’s work on GR, this doesn’t mean that we wouldn’t have eventually had a very similar theory: it just would have appeared a few years later under someone else’s name, without the Eddington publicity.

What if Einstein had fallen under a carriage in 1904? Schwarzschild was already interested in curved space, we would most likely have still arrived at $E=mc^2$ not too many years later (Fadner 1988 [99]), and given that **Johann Georg von Soldner** (1776-1833) had published the Newtonian lightbending prediction in 1804, [63], [64] one would hope that someone, somewhere would eventually have thought to test it. If they’d then found that the actual lightbending was double Newton’s prediction, we’d have known that physics needed changing, making a connection with the Mercury anomaly would tell us that gravitational curvature needed to be stronger than that of just curved space, and curved spacetime would have given us gravitational time dilation from a rather more circuitous route.

Producing a general theory by working backwards from lightbending and warped-spacetime precession, it’s not obvious that we’d have felt the need to create an additional underlying flat-spacetime theory: after all, we’d have already have transcended the concepts of globally constant lightspeed and light always travelling in straight lines that generate the 1905 theory. Within curved spacetime and gravitational theory, the easiest way to implement local *c*-constancy is with gravitomagnetism, after which special relativity becomes superfluous.

i If Mach’s Principle was supposed to treat inertia as being the result of an interaction between a mass and background environmental matter, and was to be implemented as a field effect, then by “piling up” matter in a given location, we should be able to increase the inertia (and reduce the ageing-rate) of a given test-mass. [72]

Rindler: GR before SR

Wolfgang Rindler has produced a fictional “alternative history” scenario in which Bernhard Riemann realises that curvature should also affect time coordinates, and goes on to produce a workable general theory of relativity not long after 1854. [100]

In Rindler’s narrative, this general theory then spawns special relativity as a flat-spacetime limit.

However, we can ask whether it might not have been more natural to have this (fictional) Nineteenth-Century version of general relativity reduce by default to Newtonian equations rather than special relativity (which hadn’t yet been devised).

To make the longer NM wavelengths compatible with a metric and wave theory, we would have needed to assign velocity-dependent curvature to the relative motion of physical masses. This means that in some ways, a fictitious Newton-Rindler-Riemann theory might have ended up more advanced than our current general theory: it would need to reduce to an acoustic metric rather than Minkowski spacetime, and would then by default support the classical counterpart of Hawking radiation.

When quantum mechanics then came along, some of the weirder observer-dependent properties of quantum theory might then have been seen as as the natural consequences of applying quantisation to a noisy acoustic metric, and projecting the results onto a flat plane. If a general theory *had* arrived in the mid-C19th, then perhaps Einstein’s career would have instead focused on developing quantum gravity.

17.7. Summary

Physics history is full of tantalising “might-have-beens” and near misses. What else might Clifford have achieved if he hadn’t died young? A footnote to Clifford's further posthumous book (1886 edition [87]) asks whether curved space might not be a better hypothesis than aether. “*It is a question whether physicists might find it simpler to assume that space is capable of a varying curvature, and of a resistance to that variation, than to suppose the existence of a subtle medium pervading an invariable homoloidal space.*”

Without Einstein, Lorentzian electrodynamics would have continued being developed (albeit a little more slowly), and Minkowski would probably still have published a geometrical version. Post-1911, various geometrical theories of gravity would have been developed, although probably mostly by the mathematical community, rather more slowly, without Einstein’s sense of urgency, driving physical insights, and adoption of Mach’s principle. Various researchers were also already groping their way towards $E=mc^2$ by 1905. [99]

Einstein’s biggest single unique contribution to relativity theory seems to have been not SR, but his 1911 argument for gravitational time dilation. Although the argument was childishly simple, and required nothing much more than understanding Doppler effects and Newtonian gravity, the physics and math communities had already failed to notice it for at least half a century. Without someone like Einstein, it might well have stayed unnoticed for another half-century.

> We can imagine a general theory being developed without special relativity.

> A “GR without SR” would have been based on some different principles, would have some different behaviours, and would have required more advanced geometry. It might have taken longer to develop. But it might have been at least as good as what we currently have.

18. SR Argument 18: "The correctness of SR is shown by the velocity-addition formula"

18.1. Argument for SR

We are told that experimental evidence shows us that in real life, velocities don't "add" in the way that they do under classical mechanics: if we have two co-linear same-sign velocities v_1 and v_2, the result of combining them is not the result that we'd expect from the numerical value $v_3 = (v_1+v_2)$, but is a bit smaller. If a signal is normally known to move through a piece of material at velocity v, and we see that material to be moving at V, then the apparent rate of the signal won't be $V+v$, but a smaller rate.

18.2. SR example

Suppose that a signal passes between three massed bodies $\mathbf{M_1}$, $\mathbf{M_2}$, $\mathbf{M_3}$ with relative motion, and all three masses and their motion-vectors lie on a straight line. The signal moving from $\mathbf{M_1}$ to $\mathbf{M_3}$ will undergo two Doppler shifts, one as it moves from $\mathbf{M_1}$ to $\mathbf{M_2}$ (which have a mutual recession velocity v_1), and a second as it moves from $\mathbf{M_2}$ to $\mathbf{M_3}$ (which have mutual recession velocity v_2). The total shift is the product of the two smaller successive shifts. However, we'd sometimes like to calculate the total effect in a single step, from just the relative velocity of $\mathbf{M_1}$ and $\mathbf{M_3}$.

Naively, we'd be tempted to write,

$$\text{Shift}(v_1) \times \text{Shift}(v_2) = \text{Shift}(v_3), \quad _{where} \; v_3 = (v_1 + v_2)$$

Let's try this with the SR shift relationship and set both velocities v_1 and v_2 to half lightspeed:

$$\sqrt{\frac{c-v_1}{c+v_1}} \times \sqrt{\frac{c-v_2}{c+v_2}}, \text{ and } \sqrt{\frac{c-(v_1+v_2)}{c-(v_1+v_2)}}$$

$$\sqrt{\frac{1-0.5}{1+0.5}} \times \sqrt{\frac{1-0.5}{1+0.5}}, \text{ and } \sqrt{\frac{1-1}{1+1}}$$

giving

0.333' , and **0** .

These two sets of calculations clearly don't agree. Since the left-hand operations are non-negotiable, and so is the SR Doppler formula, the only thing that we can change is the assumption that $v_3 = (v_1+v_2)$. We need some more exotic relationship to replace (v_1+v_2) , and this is where the **SR velocity addition formula** comes in.

According to the formula given in §5 of Einstein's "electrodynamics" paper, [1] "0.5c plus 0.5c" should give a total of v_3 = 0.8c, and $\sqrt{(1-0.8)/(1+0.8)}$ then gives us the result we want, 0.333'. Thanks to the formula, we can calculate the SR shifts individually, or as a composite, and get the same answer either way. Under special relativity, v_3 is not just an "effective" velocity … as we are saying that the motion of the bodies has no effect on lightbeam geometry, the SR velocity addition formula has to be regarded as *a structural property of spacetime itself*. Within Minkowski spacetime, 0.5*c* plus 0.5*c* *really does* equal 0.8c.

Under SR, we can then say, **a simple composition of any two velocities less than lightspeed gives a result that is also less than lightspeed, and *c* plus any sub-lightspeed velocity is still *c*.**

18.3. "Newtonian" velocity-addition

Now let's try the equivalent exercise under Newtonian theory, where we're normally told that $v_3 = (v_1+v_2)$ *by definition*. With the same two initial recession velocities of v=0.5c, we would have predicted recession Doppler redshifts of,

$$\frac{c-v_1}{c} \times \frac{c-v_2}{c} \text{ , and } \frac{c-(v_1+v_2)}{c}$$

giving,

$$\mathbf{0.5 \times 0.5\ (= 0.25)}, \text{ and } \mathbf{0}\ .$$

The classical velocity relationship $v_3=v_1+v_2$ doesn't hold under Newtonian theory either!

The "effective" relative velocity of $\mathbf{M_1}$ and $\mathbf{M_3}$, according to a signal sent via $\mathbf{M_2}$, is less than the value we get by simply summing the velocities together – in order to calculate the composite shift in a single stage, and still arrive at a final value of 0.25, Newtonian theory requires "0.5c+0.5c" to equal 0.75c

> Although the NM version of the velocity addition formula is not the same as special relativity's, it has the same critical characteristic, that **a simple composition of any two velocities less than lightspeed gives a result that is also less than lightspeed, and *c* plus any sub-lightspeed velocity is still *c*.**

18.4. Comparisons

Section 6 of Einstein's "Relativity" book says that under classical mechanics, if someone walks along a train at v m/s, and the train is then said to be moving at w m/s, we expect the default total velocity W to be $W=v+w$. Later, in section 13, a more intense SR-based analysis shows that this doesn't work. What Einstein doesn't mention is that the old formula doesn't work even in Nineteenth-Century Newtonian mechanics. [i] A similarly intense and exacting analysis using the Newtonian equations would have come to a similar conclusion to the one using SR.

18.5. Summary

We cannot take the existence of "non-standard" velocity-addition as proof of special relativity, because non-standard velocity-addition is a feature of a range of theories, including Newtonian physics (although this might not have been widely understood or documented at the time). "Simple" velocity addition works for low velocities, and also for a fixed absolute aether (where we apply different Doppler laws to different stages depending on how each stage moves with respect to the absolute reference). "Simple" velocity addition does *not* tend to work when the Doppler equations for relative velocity are identical for all observers, and for each stage of a compound shift (as required by relativity theory).

> A definite departure from $v_3=v_1+v_2$ may be evidence for the principle of relativity, but is not in itself evidence that special relativity is the the correct *implementation* of relativity theory.

i Mathematicians sometimes impose mathematical rigour onto "loose" scenarios where the rules have not yet been fully derived. The textbook Newtonian relationship "$v_3=v_1+v_2$" is a default assumption that doesn't hold up to analysis when velocities approach that of light. The mathematical physicist may respond to this by extrapolating and saying that the original (bad) assumption means that Newton believed that lightspeed was infinite, which is wrong (see section 8). It might be more accurate to say that Newtonian physics never had a fully satisfactory description of the behaviour of light, and that without this, we could not properly review, revise, and "fine-tune" the theory.
We can add the velocity-composition law to our list of problems and issues with Nineteenth-Century Newtonian theory that were never adequately resolved.

19. SR Argument 19: "Special relativity is validated by the Fizeau effect"

19.1. Fresnel and Fizeau

Back in the early 1800s, **Augustin-Jean Fresnel** (1788-1827) decided to "relativise" optics so that Snell's law of refractive index would give the same local results regardless of how a system moved with respect to an aether. [101] Fresnel's conclusion was that if we wanted these relationships preserved, a moving block of glass had to be seen to drag light, with the amount of dragging depending on the block's refractive index *n* (where the speed of light in the block is *c*/*n*).

In around 1850, **Armand Hippolyte Louis Fizeau** (1819-1896) then carried out a painstaking measurement of the difference in one-way velocities of light passing upstream and downstream through tubes of moving water, [102], [103] and concluded that the dragging effect predicted by Fresnel was real, and agreed with Fresnel's formula to the available experimental accuracy.

The dragging effect of matter on light has been repeatedly confirmed since, and an experiment by **R. V. Jones** (1972 [104]) also provided evidence of a transverse deflection of light sent through a spinning transparent disc, aimed parallel to the rotation axis.

19.2. Fizeau *vs.* special relativity

At first sight, the Fizeau result seems more like experimental evidence *against* special relativity. If moving matter drags light, and drags it completely, and if the same effect happens with *all* matter, we immediately have a relativistic light-dragging model in which *c* is locally constant in the proximity of any matter able to function as an observer, and we've already reconciled local *c*-constancy with the principle of relativity, without having to invent special relativity.

Special relativity explains how, if moving matter ***does not*** affect lightbeam geometry, all observers (and potential observers) in a region can agree as to a single underlying shape of spacetime, in which their different experiences relate to different projections of that single agreed underlying geometry (which is Minkowski spacetime).

But if moving matter *does* drag light (as a function of direction, speed and proximity), there will be no single agreed geometry to derive. The light-geometry of a region *diverges* from flat spacetime as a function of the relative velocities of the masses involved, spacetime is dynamic, the shape of spacetime physically changes when we add more moving bodies or change their states of motion, and there is no single underlying shape that applies to all possible situations. Further, since the geometry must diverge further from Minkowski spacetime as relative velocities are increased, and the distances of Minkowski spacetime correspond to the predictions of special relativity, the result of velocity-warped geometry cannot be the SR equations.

> In geometrical physics, changing the lightbeam geometry as a function of relative velocity changes the velocity-dependent physical relationships. The existence of light-dragging effects must be associated with a deviation from the relationships of SR.

19.3. The Einstein view

Einstein dealt with all of this proactively, by declaring that the Fizeau result actually confirmed special relativity. His "take" on the Fizeau experiment ("Relativity" [65] chapter 13) was to point out that SR's special rules for global *c*-constancy only claimed validity for light in a vacuum, not for light travelling between the atoms of a gas, or a transparent liquid or solid.

What we were *supposed* to concentrate on in the Fizeau experiment (said Einstein) was not the fact that the dragging effect *existed* (which seemed to contradict SR) but the fact that the dragging effect of moving matter on light was *less than we might expect* (validating the use of an SR-style velocity addition formula (section 18).

If we treated the light-signal moving through a water-filled glass tube as if it was some other more conventional moving thing (like a bug crawling along the outside of the tube), then the speed that we saw the bug moving at was less than the speed of the bug added to the speed of the thing it was crawling along. We could then invoke the SR velocity addition law, say that this wouldn't happen unless SR was correct, and claim the effect as proving special relativity.

19.4. What do we mean by partial dragging?

Confusingly, the word "partial dragging" has different overlapping understood meanings. The extinction theorem suggests that when light moves from air to a moving glass block, it takes on a new *absolute* velocity of c/n with reference to the block's state of motion. This is in line with the observation that while we *can* measure one-way velocity differentials across a *moving* block, we do not see any such asymmetry (caused by the Earth's motion) in a *stationary* block. For someone *moving with the block*, the velocity of the light entering the block from some otherly-moving medium, quickly takes on the reference frame of the glass, with the relative motion of background material no longer having any detectable influence. We might feel entitled to call this an **absolute** dragging effect, and under the extinction theorem, the new wavefront *really does* move at c/n with respect to the transponder-atoms, as viewed in their own frame.

However, when the velocity of the light in the glass is measured by someone in the *lab* frame, the fact that the system of glass-and-light is *moving* means that the system's distances and times (and therefore also velocities) get redefined. The dragged light appears to show a different speed to that we might expect, because *everything* in the moving system shows different rates to what we might expect. With respect to the system's own *internal system of references*, the light can be considered to be fully dragged, but for someone for whom the glass is moving, the effect on the light seems to be weaker than expected.

19.5. Newtonian visual relativity?

Returning to the exercise in section 18.3 involving three mutually-receding masses, we can treat masses $\mathbf{M_1}$ and $\mathbf{M_2}$ as a single system receding at half lightspeed from $\mathbf{M_3}$, *within* which $\mathbf{M_1}$ recedes at a further half lightspeed. But the receding system is seen to age more slowly by a ratio of $(c\text{-}v)/c = 0.5$ due to its recession, which means that the rate at which $\mathbf{M_1}$ and $\mathbf{M_2}$ *seem* to be separating is now $0.5c{\times}0.5$. So we can obtain the same final result *either* by calculating the two shifts separately, and then multiplying them together ($0.5{\times}0.5 = 0.25$), *or* by adding the local velocity to the apparent velocity, and saying that $v_3 = 0.25c{+}0.5c = 0.75c$, making the total shift once again $E'/E = 1\text{-}0.75 = 0.25$.

Let's try another (more random) example to check that this result isn't a lucky coincidence:

Let's suppose that $v_1{=}0.7c$ and $v_2{=}0.8c$. The two individual Doppler shifts, calculated using $(c\text{-}v)/c$ come out as 0.3 and 0.2, multiplying together to give a total combined redshift of $E'/E = 0.06$.

Masses $\mathbf{M_1}$ and $\mathbf{M_2}$ can be thought of as comprising a system receding from $\mathbf{M_3}$ at $0.8c$, and is seen to age at a rate of 0.2 × normal. The apparent, visual velocity of v_1 (nominally $0.7c$) within the receding system is now $0.7c{\times}0.2 = 0.14c$, the composite velocity is $0.14c{+}0.8c = 0.94c$, and the total shift using $(c\text{-}v)/c$ is then, once again, $E'/E = 0.06$.

19.6. Velocity addition formulae under different theories

The exercise above generates a velocity-addition formula for the Newtonian relationships (for the simple case of objects mutually receding along a straight line, where c=1, and all velocities are quoted as fractions of c), using the "sum-minus-product" rule,

$$v_3 = v_1 + v_2 - v_1 v_2$$

If both velocities are less than c, the result is always less than c, and if one velocity equals c, the result is still c. Special relativity is not the only theory that can play games with velocity-addition!

One caveat that we do have to apply when comparing an SR "composition of velocities" formula to those of other systems is that the associated behaviours will be different under different theories.

- Under special relativity the formula is a geometrical aspect of Minkowski spacetime, and it doesn't matter if intermediate objects referred to by the formula are real or fictitious.
- In an NM-based system, the formula documents how the energetics of a signal passed between two bodies *physically changes* if there is intermediate moving matter.

In Minkowski spacetime, we are free to add and subtract velocities to our heart's content without affecting the underlying geometry. In an NM-based relativistic acoustic metric, we are only allowed to invoke the formula if there *really is* a physical intermediate mass in the signal path. This is essential for modelling the behaviour of acoustic horizons, where classical Hawking radiation (which allows signals to migrate though a horizon) is associated with geometry-changes due to intermediate matter.

19.7. Summary

Einstein's adoption of the Fizeau effect as a confirmation of SR focuses on the effect's associated velocity-addition behaviour, without mentioning that velocities do not add "simply" even with Newtonian theory. Special relativity does not attempt to explain why and how *the Fizeau effect itself* is able to exist – why a group of SR *observers* in motion, exchanging signals, is supposed to be associated with no detectable change in one-way lightspeed, but a group of *atoms* in motion, exchanging signals, is associated with measurable lightspeed offsets.

Do different laws of physics apply to SR observers and to atoms?

The Fizeau effect is not predicted by special relativity, and has to be added "by hand". If SR's description of light behaviour, derived for empty space, does not still apply in the presence of matter, then SR is not a physical theory. If it continues to apply in the presence of matter, then by default, moving matter should not drag light, and there should be no Fizeau effect.

> The existence of the Fizeau effect appears less like an experimental validation of SR and more like an experimental disproof. It corresponds to the class of effect that we would expect to see if special relativity was wrong.

We can try to preserve special relativity by rejecting the idea that moving masses have any *field-like* influence on the behaviour of nearby light, and one way to do this is by invoking the **extinction theorem** …

20. SR Argument 20: "SR is correct because the extinction theorem shows that light-dragging effects don't exist"

20.1. The extinction theorem

The extinction theorem of **Paul Peter Ewald** (1912 [123]) and **Carl Wilhelm Oseen** (1915 [124], [125] has been widely used to model the changes in lightspeed that happen when light enters a stationary or moving particulate medium, while avoiding any smooth changes in velocity (which might suggest a classical mass-field effect and introduce curvature into the description). [i]

In the extinction theorem's description, atoms are treated as signal transponders, or as oscillators that are triggered by an incoming wavefront, absorbing the incoming signal and replacing it with a replacement signal of the same frequency. The incoming signal is said to be absorbed and extinguished over a distance referred to as the **extinction distance**, or **extinction length**.

20.2. Motivation

According to the theorem, when a signal enters a glass block it doesn't *really* slow down, instead the incident wavefront is progressively extinguished and replaced with a *new* wavefront travelling at whatever the speed of light is in the block, c/n (where n is the block's refractive index). This attempts to explain known behaviour without saying that the speed of a light-signal ever *really* changes (which would imply variable-c curvature gradients and a departure from flat spacetime) – instead, we have a mixture of *different* signals, moving at different speeds.

20.3. Problems

The extinction theorem does have some conceptual problems:

(a) It doesn't really explain *why* the speed of the re-emitted light should be different to that of the incident light, or why, if SR says that the speed of light is independent of the speed of the source, the new signal chooses to move at c/n referenced (preferentially) to the speed of the atoms in the block. This is getting perilously close to bad old ballistic emission theory.

(b) Special relativity is supposed to allow us to model the behaviour of arrays of observers, who may choose to measure their distances by exchanging signals and using Einstein's clock-synchronisation method to arrive at definitions of distances and times. These arrays of observers need to assume that the speed of light moving between them is c_{VACUUM} .
However, when we have a moving array of *atoms* with signals passing through the region, experience and the extinction theorem tell us that the light behaves differently to the SR description. Is an atom too complex an object to be able to act as an SR observer? If the rules of special relativity regarding moving "observers" do not apply *even to moving arrays of atoms*, then it is difficult to say that SR is in agreement with all known experimental data.

(c) There is also the problem is what happens when the new wavefront passes the last layer of atoms in the block and re-enters vacuum. When the signal leaves the last row of atoms,

---

i Isaac Newton's description favoured the change in lightspeed when light encountered a particulate medium being a smoother effect, more conducive to a classical field interpretation. The influence of particles extended out some distance, and light started to respond to the presence of the particles before it reached them.
Opticks, "***Qu. 4.*** *Do not the Rays of Light which fall upon Bodies, and are reflected or refracted, begin to bend before they arrive at the Bodies …* "
in Newton's scheme, the smooth variation in lightspeed over small regions as we approached a particulate mass (giving refraction) was the same effect that caused a smooth variation in lightspeeds over larger distances as we approached a gravitational body. This variation in lightspeeds associated with a variation in aether density, was the source of gravitational effects.
"***Qu.21.*** *… it may suffice to impel Bodies ... with all that power which we call Gravity.* "

travelling at c/n, then how does the light *know* that it is then supposed to speed up again? The "extinction theorem" description includes a mechanism of sorts (or at least an association between circumstance and behaviour) for replacing a fast signal with a slow one for an incoming wavefront, but lacks a matching mechanism to explain the subsequent change when the signal leaves the block and re-enters pure vacuum (Keshwari, 2003 [126]).

If the light *is* somehow able to sense that it is leaving the influence of the block – if it has some ability to sense *proximity* – then we are back to a classical field description, and since classical field responsible for altering lightspeeds is the *gravitational* field, we again have curved spacetime, which is what we were trying to avoid.

20.4. The extinction theorem *vs.* Michelson-Morley

The extinction theorem is useful for modelling particulate effects on light, especially those that can complicate astronomical optical calculations, where we can have exceptionally rarefied clouds of particles, spread over vast distances. However, it is not sufficiently complete or internally consistent to count as a full-blown theory, and tends to break SR proofs and derivations. [i]

Since the extinction distance for optical frequencies in air is reckoned to be maybe around one or two millimetres (depending on who calculates it, and how), the Michelson-Morley "aether wind" detection experiment would seem to have been doomed to give a "null" result due to the dragging effect of the Earth's atmosphere, regardless of what the speed of light might or might not have been in the surrounding region of outer space, and regardless of whether or not the Earth's moving gravitational field or the proximal hardware also had any additional influence.

This puts us in an awkward situation:

- A fairly obvious suggestion for why Michelson and Morley might have obtained a null result for aether wind was that perhaps that moving Earth dragged light along with it ("dragged" or "entrained" aether). If so, we wouldn't need Lorentzian arguments or special relativity to explain the Michelson-Morley experiment. We could say, instead, that lightspeed is locally constant for every mass, and that lightspeed varies between relatively-moving masses as some function of proximity. This (conceptually, at least) would be quite a simple model.
- However, the available aether-centric models of light-dragging at the time seemed to be overcomplicated, overspecific, and inconsistent, and measurements of stellar aberration due to the Earth's motion around the Sun were being presented as proof that the Earth did *not* drag light along with it (Airey 1872 [128]). [ii] With relativistic light-dragging models apparently ruled out, this left LET/SR as our main remaining option.

> If aberration observations *really* prove that light is not being carried along with the Earth, and the extinction theorem insists that light at around sea level *must* be carried along with the Earth, because the Earth's surrounding envelope of atmosphere is sufficient to carry light along completely (as seen by Earth observers), then something, somewhere, does not add up. It does not seem possible for both arguments to be right.

---

i Interstellar gases, combined with the extinction theorem, upset de Sitter's original 1913 optical-wavelength disproof of emission theory, by replacing the original signal speeds with a fixed speed with respect to the interstellar medium (Fox, 1962 [127]).

ii Dragging is a proximity effect. Even with no deflection of starlight happening *within* the dragged region, the star will be *outside* the dragged region, and there must still be a deflection *at the boundary*, or *transitional zone*. This *transitional* deflection happens some distance above us, and might have been overlooked in Airey's arguments.

20.5. Is a single atom a particulate medium?

Rarefied media *vs.* vacuum

If we are to consider special relativity as a valid physical theory, we need light emitted by a completely isolated lone atom (in a vacuum), to be describable as travelling at c_{VACUUM}, for any observer (if one atom is too much for the theory to deal with, then it is not a physical theory).

But according to the extinction theorem, if the atom is surrounded by neighbours, the emitted light it emits travels instead at a definite velocity (in the atom's frame) of c/n, where n is the refractive index, This presents us with something of a problem: When is an atom considered to be in empty space, and when is it considered to be part of a particulate medium? And how does the atom *know*?

For interstellar space, the density of matter may be less than one atom per cubic centimetre (a degree of rarefaction that we don't yet attempt with current vacuum pump technology) … and yet, as far as the extinction theorem is concerned, this still counts as a particulate medium (with an extinction distance of, perhaps, a few lightyears, depending on how it is calculated). Does a solitary atom surrounded by a cubic centimetre of utter void count as a legitimate "observer in a vacuum", or does the presence *of the atom itself* invalidate the idea that the region is empty?

Is there *any* situation in which special relativity's behaviour of light "in a vacuum" is considered to actually apply, or is this a mathematical ideal that only holds in the total absence of matter?

A field interpretation for c/n

The extinction theorem (and experience) says that the velocity of light is different when the atom is surrounded by other atoms. The existence of surrounding atoms *could* change the speed of light if the light could sense their presence as a proximity effect … but in that case, the atoms would have *fields* that altered lightspeed (*i.e.* gravitational fields), and special relativity would be invalidated.

"Boxing the atom" – non-field interpretations and scaling effects

What if we were to suppose that neighbouring atoms' fields ***do not*** play a part in influencing lightspeed around our selected atom? We can draw a unphysical "box" or boundary around the atom to exclude its neighbours, and only study the physics inside the box, where the effects of any neighbours (supposedly) do not intrude. An individual atom, emitting light, is then required to behave in exactly the same way regardless of whether it is isolated, or surrounded by other atoms.

Let's initially take a 1 cm^3 volume of interstellar space containing a single atom, corresponding to a volume of interstellar space. We can tile space with identical boxes each containing an atom, to get a particulate medium of density "one atom per cubic centimetre". The speed of light emitted by the atom will then need to be a little lower than c_{VACUUM}, but not by much.

Now let's zoom in on the centre of the box, and look only at the central cubic *nanometre* (10^{-9} of a metre), surrounding the single atom. If the internal physics of the tiny box (10^7 times smaller on each side than the previous box) is unaffected by what's outside (no-fields assumption), then the region will not know whether it is surrounded by total vacuum, or by an array of similar boxes with 1 nm spacing. For light leaving the atom within this box to behave the same way regardless of whether or not the atom has neighbours, it needs to behave as if the box is part of a particulate medium with a density of one atom *per cubic nanometre*, with 10^{21} times the density of the larger one-centimetre cube.

We then need the *effective* value of the smaller region's refractive index to be much larger, and the speed of light leaving the atom in the smaller box must be significantly slower than the overall speed for light leaving the same atom in the larger box.

Reappearance of the field

We can now draw a nested series of hypothetical concentric boxes around the atom, assign a different required effective refractive index to each box, and notice that for an *arbitrarily large* otherwise-empty box containing one atom, the region's speed of light c/n for outgoing light, does indeed tend towards c_{VACUUM} .

The problem we now have is that in order for the same atom to be able to reproduce different refractive index effects over a range of scales, without knowing what's outside a given box, our atom must produce a range of lightspeeds that vary as a function of distance from the atom, with the speed of light being slower closer to the atom (smaller bounding box, denser effective medium), and increasing towards c_{VACUUM} as we get further away (larger bounding box, more rarefied effective medium).

Self-invalidation

Since a smooth variation in lightspeeds with location in otherwise empty space is equivalent to the existence of a gravitational field, what we have inadvertently done by trying to model refractive index *without* fields using the extinction theorem, for a single atom, is to derive the *existence* of a gravitational field associated with the atom, whose presence then explains refractive index. [i]

Once we have the result that our lone "test atom" has a field, the rest of our initial working for section 20.5 is invalidated, since we require all matter to obey the same laws and are then obliged to assign similar fields to all other atoms, contrary to our initial assumption. In a subsequent iteration of the argument, light then *is* able to sense the existence of neighbouring atoms, by the intrusion of those atoms' fields into the region, and refractive index is the result of the combined effects of small-scale gravitational fields.

> If we try to use the extinction theorem to explain refractive index *without* particles having gravitational fields (in order to protect special relativity), we end up deriving the *existence* of gravitational fields, a gravitational mechanism for refractive index, and non-SR physics. The extinction theorem does not protect special relativity: it destroys it.

20.6. Summary

The extinction theorem has to be considered a quick-and-dirty "engineering" approach rather than a considered, fully thought-out theoretical model. While the theorem tries to explain refractive index without involving fields or smooth variations in lightspeed, that would suggest curvature, it is more of a modelling tool than a theory, and doesn't really count as a workable defence of special relativity's concept of global lightspeed constancy and flat spacetime. It also leads to arguments that end up suggesting spacetime curvature and non-SR physics.

An improved version of the argument (a sort of "extinction theorem on steroids") is outlined in section 22, and has a much greater degree of logical consistency.

Before this, we'll have a quick look at how physics diverges from special relativity in a relativistic universe that supports the Fizeau effect.

---

i For wavelength-dependency issues, see section 22.2.

21.SR Argument 21: "Einstein's trains" thought-experiments

21.1. The three-train experiment

In a minor adaptation of Einstein's two-train thought-experiment, we have three parallel sets of railways tracks, with one train on each set. Each train is equipped with an onboard laboratory.

Lightning strikes a low footbridge **B1** straddling the track, and the flash runs parallel to the rails, passing through the bodies of the three trains, and is then seen by an onlooker standing on a second footbridge **B2**.

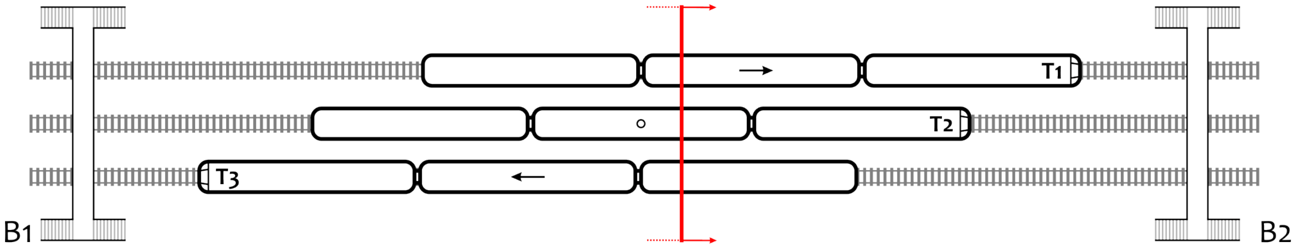

*Figure 5: **Three Trains (SR)**. According to special relativity, the wavefront of a flash of light (red line) should advance through all three trains at exactly the same rate. Observers are entitled to assume that lightspeed is globally fixed throughout the region, with respect to any train.*

Of our three trains, **T1** is moving towards the second bridge at fixed speed, **T3** is moving away from the second bridge at the same fixed speed, and the intermediate train **T2**, on the central set of tracks, is parked, stationary.

21.2. The Einstein description

If we believe that the speed of light is globally constant, and not affected by moving bodies, then the wavefronts of the signals passing through the three trains and through the spaces between them, will advance at the same rate, as a single wavefront. A viewer on the second bridge, watching the four signals will see them arrive at **B2** together. The viewer at **B2** will see a single flash.

We then have a problem: The flash takes a longer time to travel along the length of the train **T1** (which is moving in the same direction as the flash), and takes a shorter time to travel the full length of **T3** (which is moving against the light) How can we reconcile these different train-length transit times with the idea that the velocity of the light is seen to be the same by observers in all three trains?

Einstein's approach tackles this in two stages:

First part

> Einstein points out that experimenters onboard the trains are not physically able to measure the one-way speed of a light-flash, as they cannot be in two places at once. Although they cannot measure the light's one-way velocity, they can measure *the round-trip speed*. Experimenters can place themselves at the end of the train nearest to **B1**, record the time that the light first enters the train and passes their location, then record the later time at which they see the flash reflected off the interior of the train's front end, and divide the time by two to calculate the *averaged speed* of the light, for both directions. [7] For train **T1**, the outward light-trip takes longer, and the return trip is shorter, while for train **T3**, the outward trip is shorter and the return trip takes longer. The physical round-trip times (and therefore the measured round-trip speed of light on these two trains), is identical on trains **T1** and **T3**.

Einstein says that since globally constant *c* makes it impossible to measure the one-way velocity of light, that *this is not to be considered a physical property*. If a thing cannot be measured *on principle*, it is not real. [i] What *is* real is the round-trip average speed., and from this, observers are entitled to *presume*, or *choose to believe* that the underlying one-way velocities are the same in both directions ... as there is nothing that can demonstrate to them that their belief is wrong. Taken in isolation, the interior physics of the two trains $\mathbf{T_1}$ and $\mathbf{T_3}$ is identical, even though one is travelling towards the flash and the other is travelling away from it.

Second part

The next step is to notice that while the measured round-trip time is the same for $\mathbf{T_1}$ and $\mathbf{T_3}$, both times are are slower than the round-trip time in the stationary train $\mathbf{T_2}$. With the SR equations, as the speed of a train approaches lightspeed, the "fast" signal transit time goes to zero, and the "slow" signal time goes to infinity. The sum of the two times (divided by two), increases with velocity, so the faster the train moves, the longer it takes for light to do a round trip.

This is then dealt with by saying that under SR, it *also* takes longer for clocks onboard $\mathbf{T_1}$ and $\mathbf{T_3}$ to record the passing of an agreed amount of $\mathbf{T_2}$-time: if they carry a pocket light-clock in which light bounces between two mirrors, this will "tick" more slowly by the same ratio as the larger bouncing lightbeam, and if *all* forms of clock are affected by the lightspeed variations in the same way, [ii] [iii] the trains $\mathbf{T_1}$ and $\mathbf{T_3}$ will be unable to measure the slowing of their round-trip light, because their local clocks will tick more slowly by the same ratio that affects the round-trip light-pulse.

We than have a situation in which laboratories in $\mathbf{T_1}$, $\mathbf{T_2}$ and $\mathbf{T_3}$ all report the same locally measured value for the (round-trip) speed of light.

Although this was calculated by assuming a fixed speed of light for $\mathbf{T_2}$, and assuming that clocks in $\mathbf{T_1}$ and $\mathbf{T_3}$ were supposedly running slower, in practice, if $\mathbf{T_1}$, $\mathbf{T_2}$ and $\mathbf{T_3}$ compare notes, they cannot see anything special about $\mathbf{T_2}$'s state of motion. If we'd instead decided that the speed of light was "really" fixed for $\mathbf{T_1}$, or for $\mathbf{T_3}$, the corresponding calculations would have ended up declaring that different trains were "really" time-dilated (with some nominal values being different), but we'd again have exactly the same physical end-result, that all three trains reported their local round-trip speeds of light to be constant and equal.

---

i Einstein, "**Relativity**", "8. On the idea of time in physics", " *'That light takes the same time to traverse the path A→M as for the path B→M is **neither a supposition nor a hypothesis**, but a **stipulation** which I can make of my own free will in order to arrive at a definition of simultaneity.'* "

ii A "pocket light-clock" will obviously be affected by lightspeed variations by the same amount as the larger light-clock formed by light passing back and forth along the body of the train. If the resonant frequencies of atoms are considered to be the result of internal light-clocks, then these will also slow by the same rate.
If we consider the round-trip actions and reactions of interatomic forces trying to stay in equilibrium to also be slowed, then the response of an inertial mass to an applied force will be more sluggish, and the mass will seem to be greater. A quartz crystal will resonate at a slower frequency (so that quartz clocks run slower), atomic clocks will run slower, and even if we have an antique Victorian pocket-watch, slowing the averaged electromagnetic signal speed will make the flywheel will take longer to accelerate and decelerate under the influence of the force applied by the watchspring, and the watch will tick more slowly, too. Similar arguments apply to the timings of the human nervous system and brain.

iii Essentially, the speed of light is the single factor that affects all other timing mechanisms – if we change the (round-trip) speed of light in a region, we change the rate at which all timing processes work in that region. Changing the speed of light effectively *changes the rate of timeflow*.

21.3. A more realistic description

In reality, the onlooker on bridge **B2** should see at least three flashes rather than one. Thanks to the Fizeau effect, the speed of light is *not* globally constant across the region, it is dragged by the different motions of the trains (including the motion of the transparent windows by which the light enters and exists the trains, and the motion of the contained air within the trains).

The viewer on the second bridge sees the flash from **T1** (which is dragging the light towards them), then the light from **T2** , and finally the light from **T3** (which is dragging the light away).

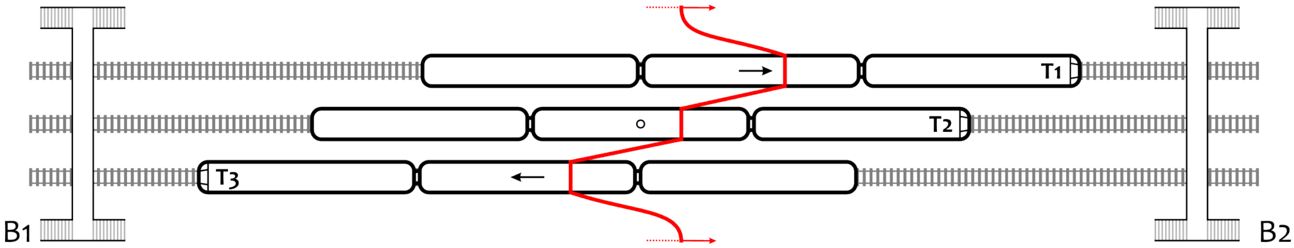

*Figure 6: **Three Trains (more realistic version). In this sketch, t**he motion of the trains' matter and contained air causes light to advance through the region at identifiably different rates (Fresnel/Fizeau effect). There is still no preferred frame, and the principle of relativity still works.*

The parallel light wavefronts progressing through the interiors of the three trains unambiguously do *not* advance at the same rate, and since the flash passing through **T1** advances fastest, Huygens' principle says that the overall wavefront veers to one side, in the direction of **T3** . The sideways deflection then shows that the lightbeam geometry of the region is not flat, and seems to show an apparent gravitational field pulling to one side.

Although an onlooker still can't unambiguously measure the absolute one-way velocity of any of these light-signals, they *can* measure the physical discrepancies between the differently-moving signals, and agree that the signal passing through **T1** reaches the bridge first, and that passing through **T3** reaches it last. All three observers agree that the one-way velocity of light in the direction of the flash is greatest for train **T1** , is less for **T2** , and is smaller again for **T3** , and that the pattern is reversed for signals moving in the opposite direction.

21.4. Relative lightspeed-asymmetries

In a relativistic model that incorporates the Fizeau effect, the principle of relativity is obeyed in a different manner to Einstein's descriptions: rather than explain that one-way velocity is unphysical, we accept that it is variable, and that the variability and **relative asymmetry** of lightspeeds has physical measurable consequences (the different consecutive arrival times of signals at **B2** moving through the three trains).

Technicians on all three trains are still entitled to believe (without anyone being able to prove otherwise) that the one-way speeds of light are symmetrical in both directions in their own trains.

A technician in **T3** can notice that the forward flash advances faster though the region in **T1** than on their own train and that the reflected flash moves more slowly in **T1**. They can declare,

> "*Light within **T1** shows a measurable anisotropy in the speed of light, biased towards **B2**, which is caused by the dragging effect of their moving train's matter on light. This makes their forward signal arrive more quickly and their rearward signal arrive late.*"

The technician in **T1** agrees that the difference is real. But **T1** is entitled to dispute **T3**'s interpretation of why the signals move at different rates. **T3** can argue,

> "***There is no detectable lightspeed asymmetry in our train.*** *It is* ***not*** *that our forward signal arrives* ***early****: it is rather that* ***T3****'s signal arrives* ***late****, because* ***T3****'s mass is moving against the light, and causes the light to propagate more slowly. Similarly, for the reflected flash, we refuse to accept that the one-way speed of light in our train is wrong: rather, the reflected flash in* ***T3*** *moves more quickly than it ought to because of* ***T3****'s matter now dragging the light along in the same direction.*"

The *relative* asymmetries in the one-way velocities of light *between* the three systems is physically real, experimentally identifiable, and agreed on by all observers. However, since no observer can identify an absolute lightspeed anisotropy in *their own* system, each observer is entitled to argue that their own system has velocities of light that are the same in both directions, and that the relative difference in one-way speeds in the other systems is due to the moving matter of those other systems. There is a physical *relative* anisotropy in the speed of light.

All observers agree that there are measurable offsets in the one-way velocities of light between the trains – the differences are real – but they disagree as to why, and as to whose fault it is.

21.5. Transverse deflections

For the transverse gravitomagnetic deflection in our exercise (predicted by Huygens' principle), the forward-aimed flash advancing in the gap *between* **T1** and **T2** will deflect a little towards **T2**, while the flash moving between **T2** and **T3** will deflect a little towards **T3** (*i.e.*, down the page).

If the flash reflects off the second bridge **B2** and tries to retrace its earlier path through the region inhabited by the trains, the reflected signal again deflects to one side but now in the opposite direction, from **T3** to **T2**, and from **T2** to **T1** (*i.e.*, up the page).

21.6. Summary

If a train's steel and glass structures and enclosed air drag light (similarly to how the water in the tubes in Fizeau's experiment drag light), then the paradox that Einstein presents ... in which we have to explain how a single signal passes through all three trains at the same rate while each observer sees global c to be constant ... never arises. Special relativity's solution is the answer to an artificially constructed question that Nature does not ask.

The principle of relativity is not maintained by experimenters in the three trains agreeing that light progresses throughout all three regions in the same way: it is maintained by them agreeing that light passes through each of the three individual regions ***differently***. The sense of apparently-constant lightspeed experienced for the inhabitants of each system is strictly local. It applies within each individual system but does not extend or extrapolate to a larger region containing the other moving systems. Each system is then in effect a local "island" of absolute c-constancy, [i] with the role of the principle of relativity giving the description of the relationships *between* the islands.

> Einstein's projection of absolute c-constancy across a region containing differently-moving masses disagrees with experience. In reality, we expect to see relative lightspeed *variations* across the region. Modelling this relativistically requires a different approach to Einstein's, and generates different equations.

i More accurately, *each fundamental massed particle* making up each train represents an "island" of local c.

22. SR Argument 22: "SR is correct because QM says that light-dragging effects don't exist"

22.1. QM against a flat-spacetime background

A more modern approach to the Fizeau effect within SR-based physics is to seize on special relativity's idea that the speed of light is globally fixed regardless of what happens along the signal path, and to take the "purist", "no compromises" position, that this rule must hold *even within particulate media.*

- In this interpretation, we agree with the extinction theorem (section 20) that atoms act as transponders, and that when the transponder array receives a signal, its response creates a new signal wavefront (perhaps aligned in a different direction), but we do *not* assign different behaviour to the regenerated light. Rather than saying that the original signal is *absorbed* (which suggests some form of modification of light by matter), we say that the array's signals are supplemented with a generated antiphase signal, which travels at c_{VACUUM} just like the original, and perfectly cancels it out.
- We also say that the component representing the "new" signal *also* travels at c_{VACUUM} ... but that it advances through the glass at a slower rate overall, because every time it encounters another atom, the new cancellation signal is emitted immediately, but the replacement signal is emitted with a further time-delay. It's not that the signal travels at *less* than *c*, it travels at *full c*, meets an atom, is delayed, is re-emitted and travels at full *c*, meets a second atom, is delayed again, and so on.

This interpretation has significantly greater explanatory power – it explains why packing more atoms into a region slows the speed of light and changes the region's refractive index, and also explains light-dragging: if the medium is moving, atoms will advance a little between absorbing a pulse and emitting its replacement, so if the medium is moving in the same direction as the light, the replacement signals will all be deposited at a slightly forward position each time, while if the medium moves against the light, the new signal will be deposited a little in the opposite direction. The more atoms we pack into a region, the slower the speed of light (more delays), and the stronger the proportional dragging effects.

We can also engineer the model to say that the timelag between a particle absorbing and re-emitting a signal is a function of signal frequency.

This is the first genuinely credible description that we have come across that successfully combines the Fizeau effect with special relativity's flat spacetime.

22.2. Equivalence to non-SR "acoustic metric" behaviour

There is, however, a "catch". The speed of light in particulate media is slower for smaller wavelengths, which means that the same signal transponder must somehow be able not only to store and regurgitate data with a timelag (meaning that an atom must somehow have an internal information storage capacity and be able to act as a temporary information buffer), but it must be able to remember and spit out different frequencies with different time-delays, and must be able to buffer a larger amount of information if the wavelengths are smaller.

How would we design a physical/mechanical system to do this? The obvious way to create a time-delay mechanism is to make use of the speed of light itself, and somehow pack extra distance into the atom's space. Could the atom be visualised as containing a coiled tube fitting more space into a given region, with a fractal characteristic that makes the effective distance longer for smaller wavelengths? This seems far too complicated. The simplest geometry that produces the effect that we need is a simple gravity-well, which would cause a Shapiro timelag as the signal traverses the additional distance represented by the well. We then get the frequency-dependent timelag by saying that the smaller-wavelength signals manage to penetrate deeper into the tiny gravity-wells, interact with more distance along their paths, [i] and emerge with a correspondingly stronger Shapiro time-delay. [ii]

The wavelength-dependent timelag also generates a wavelength-dependent dragging effect, as a well will move further before a shorter-wavelength signal re-emerges. Assigning gravity-wells to atoms creates the "timer" function described by QM and creates a region for temporary data-storage, complete with wavelength-dependence, in an incredibly simple way.

22.3. One metric, multiple virtual metrics

One of the main reasons that relativistic light-dragging models were abandoned in the Nineteenth Century (leaving the field open for Lorentzian theories and special relativity) was the observation that different colours of light had different speeds in a given refractive medium, and were deflected by different angles at a boundary, and (given their different speeds) seems to be dragged by different amounts. This seemed to suggest that we needed a *different aether* for each individual colour of light, with the different aethers having different speeds. This seemed unworkable. [iii]

A version of this argument survives: we say that we can't use light-metric arguments to model refractive index behaviour, because, since the internal dimensions of a glass block, as measured with light, are different depending on the colour of light used, so that we need a *different metric* for each colour of light.

In the "acoustic metric" description given in the last section this objection seems to disappear. For a given experiment we have a single agreed geometry, which is merely *experienced* differently by different colours of light, due to the ability of smaller wavelengths to interact with finer details of the metric's shape. We have an *general* metric, and multiple *effective* metrics, with the differences depending on the scale of the wavelengths that we choose to use to probe the region. [iv]

---

i This is similar to the fractal argument that the coastline of the British Isles has a different length depending on the size of our measuring equipment. It will yield one length value if we walk around the perimeter with a one-meter diameter surveyor's wheel, but a much larger length if we use a tiny one-centimeter-diameter wheel that also measures inside the small nooks and crannies between small rocks and individual pebbles.

ii It would be interesting to study the effect of wavelength dependency on the Shapiro effect with stellar-scale bodies, the principle being that wavelengths much larger than a feature's characteristic scale will tend to wash over the smaller-scale curvature and be less delayed, much as an ocean liner doesn't "feel" metre-scale waves that would be obvious to a small row-boat. However, for our Sun this would require signals with a range of wavelengths significantly longer than the Sun's diameter of 1.4 million km.

iii Møller (1955) [47] I§9, page 21: " *… one would have to introduce a separate ether for each colour of the light. This is, of course, an impossible assumption, …* "

iv In particulate media, measurements will show the speed of light in the particulate medium to be c/n rather than c. We might believe it to be c between atoms on the basis of the exercise, but we will never be able to get an experimental confirmation that we are right. If we try to insert sensors between atoms, to measure the "real" signal speeds between them, the presence of the particulate matter in our sensors will change the result.

22.4. Summary

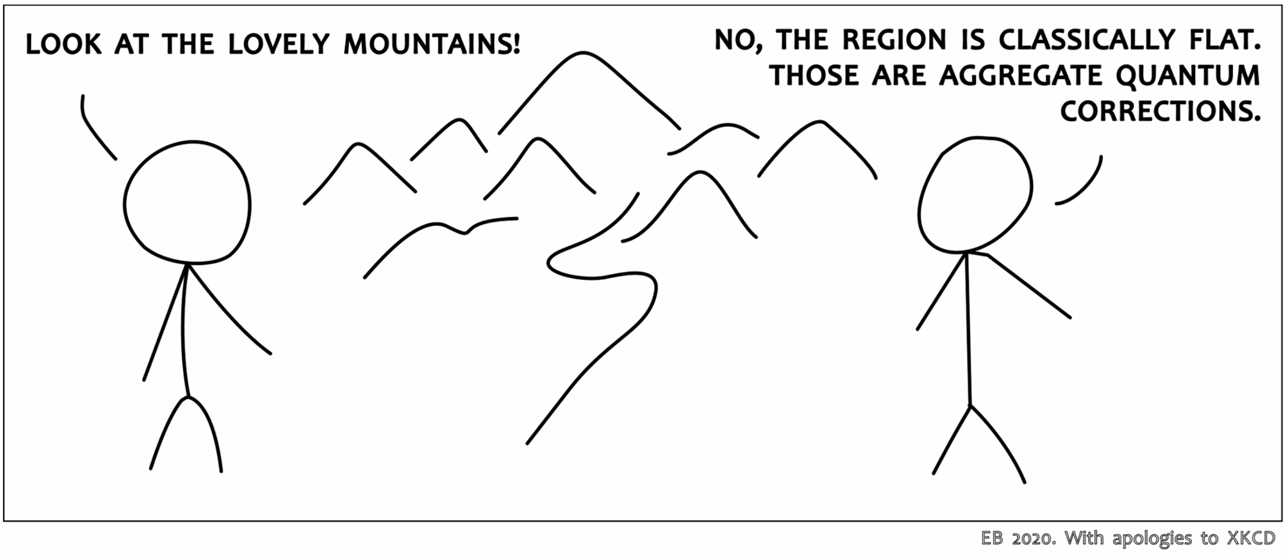

Figure 7: ***Insisting on flatness****, in the face of physical evidence to the contrary.*

Quantum mechanics can be used against a flat background to model refractive index and the Fizeau effect by simulating the physics of a "curved" relativistic acoustic metric. Since we cannot in principle tell the difference between a QM *simulation* of an acoustic metric and a *real* acoustic metric, [i] it would seem that rather than supporting the hypothesis of flat spacetime (and SR), this exercise supports the idea that classical physics behaves "as if" spacetime is curved by matter, and "as if" an acoustic metric applies rather than the Minkowski metric.

i We would seem to be able to use the QM toolset to create simulations of other classical effects, such as invoking the presence of virtual particle-pairs, more common near large masses, to recreate, say, the Shapiro effect, or gravitational light-bending. However, just because we *can* use QM tools to do something, it doesn't necessarily mean that it's a good idea.

23. SR Argument 23: "SR is correct because Newton's First Law says that light-dragging effects don't exist"

23.1. The deceleration problem under GR

An "empirical" defence of special relativity within a gravitational universe goes like this:

> "***It may well be*** *that the existence of dragging effects contradicts special relativity. Luckily, we know for a fact that no such effects exist. If a star, moving in a straight line at constant speed* **was** *able to drag light, then a traveller, moving through the universe at high speed, would experience a dragging effect from each and every one of the background stars, and would be made to decelerate until the stars appeared to them to be (on average) at rest. The only stable state would be one in which all matter was mutually stationary. This is at odds with Newton's First Law ("N1L"), which says that a body moving at constant speed in a straight line is known to continue at that speed until something intervenes. Empirical evidence therefore proves beyond any doubt that there is* ***NO*** *detectable velocity-dependent dragging effect, which mean that there is no effect to produce a deviation from SR, inertial physics is indeed a flat-spacetime problem, and special relativity is proven correct.*"

This argument amounts to saying: "*Although gravitational theory, and the GPoR, and basic principles of interaction and momentum-exchange all require dragging effects to exist, we will delete these effects from our theory, on the grounds that we know that they don't happen in real life*".

If we take this path, a general theory becomes a feeble creature, only able to avoid being crushed by experience with the help of "manual overrides" which artificially change the theory's predictions to agree with "what we know happens". We can avoid this by being more disciplined as theorists, and not rushing too quickly to insist that a theory agrees with our expectations:

23.2. The forward-acceleration problem under GR

The more cautious theorist will hold back judgement until they have acquired a more thorough understanding of the theoretical landscape in which the problem appears, at which point they will realise that there is not just *one* GR effect apparently in profound disagreement with N1L, but *two*:

Because relativistic aberration effects alter the angles of light-rays and change the apparent positions of stars, the traveller will see the background starfield to appear to be more concentrated ahead of them and more diluted behind them (Scott and Van Driel, 1970 [129]). If stars and galaxies cause an attraction towards their visibly-observed positions, this would be expected to cause the traveller to undergo a freefall acceleration forwards, towards the region of greatest apparent mass-density (forwards), with their increase in speed further distorting the apparent starfield geometry and further increasing their acceleration. The positive-feedback nature of the effect would make the universe a fundamentally unstable place – any time we bumped into an object, it would try to accelerate away from us at ever-increasing speed.

At first sight this is a puzzle: logic appears to *demand* that a star attracts towards its apparent position – the principle of relativity's apparent requirement that gravitational and optical signals have the same speed (section 8.2) suggests that the angular aberration of gravitational and optical signals should be the same, too. Further, if a star's different signals pointed to two different apparent origins, we would lose the ability to talk in theoretical terms about the star having such a thing as an "apparent position". With two different apparent positions to choose from, geometrical physics would suffer from a form of double-vision.

The usual mainstream response to gravitational aberration is again "pragmatic" rather than "theoretical": it is to repeat that, regardless of the logic, we know that it doesn't happen. Regardless of the fact that theory appears to make the effect unavoidable, because we know *empirically* that we don't see it, we are allowed to apply another manual override to the theory: We say that … without any obvious theoretical derivation or explanation why … the gravitational aberration effect must NOT exist, in order to bring the theory back into line with reality.

We may even try to pass off the "fudge" as a legitimate *definitional requirement*, due to the idea that general relativity must reduce to Newtonian physics as an approximation.

If we override general relativity *twice*, once to eliminate the braking problem, and again to eliminate the acceleration problem, then we have admitted *twice* that our gravitational theory doesn't work and fails to agree with the most basic observations of the world around us. We have not just "fudged" the theory to prevent it being invalidated: we have "double-fudged" it.

23.3. Cancellation

At this point some readers will already have realised the solution: the first problem causes a moving body to slow down, the second problem causes it to speed up, both problems are governed by the same apparent distortion of the outside universe, and both problems appear to have the same approximate strength, for any given velocity up to that of light.

In other words, the two effects pretty much cancel out (Carlip, 2000 [130]).

This is important: not only does the cancellation neatly eliminate what would otherwise be *two* catastrophic failures from our logical description of gravity, it makes the system significantly less arbitrary by removing the artificially imposed requirement that the system has to obey Newton's First Law. Instead, N1L emerges naturally from the physics, and the apparently flat spacetime background against which inertial physics plays out is explained as a natural emergent property of curved-spacetime physics.

Carlip has argued that not only does the dragging effect *not* show that general relativity is inconsistent, the effect is *required* to exist for N1L to hold (in a more enlightened view of general relativity, the dragging effect does not conflict with N1L, it is partly responsible for generating it).

> Instead of saying that curved spacetime is built on a flat-spacetime foundation, we can instead arrive at flat spacetime as an emergent outcome of underlying inherently curved-spacetime principles.

23.4. Dragging *vs.* SR

If the dragging of light by matter is universal and necessary part of geometrical gravitational physics, then this brings us back to the problem of reconciling it with special relativity, which assumes that no such effect exists. If the effect *does* exist, a region containing a pair of masses with relative motion will show a distortional effect increasing with the masses' relative velocity, the geometry will diverge from flat Minkowski spacetime as a function of velocity, and the equations of motion for the two particles will not correspond to the equations of special relativity.

> While the Carlip argument arguably rescues general relativity from the scrapheap, the form of general theory that emerges is not Einstein's: it is a different form of general relativity that does not incorporate SR physics and does not rest on SR foundational principles.

24. SR Argument 24: "SR is correct because velocity-dependent gravitomagnetism does not exist"

For special relativity to be a valid part of a general theory of relativity, we need moving gravitational bodies to not show any complicating dragging effects analogous to the Fizeau effect.

Unfortunately, they do.

24.1. Gravitomagnetism basics

Overview

Gravitomagnetism ("gm") can be characterised as being the additional effects that a moving gravitational field or gravity-well has on nearby light and matter due to its relative motion. The slightly awkward name is a reference to an approximate analogue with *electro*magnetism - when we move an electric charge, we get magnetic effects, when we move a gravitational charge, we get "gravitomagnetic" effects. Just as a static gravitational field can be considered as the result of smearing a body's mass out into the surrounding region, the body's *gravitomagnetic* field can be considered as the result of smearing the body's momentum or momenergy out into the surrounding region. [i]

In a geometrical theory, gravitomagnetism is the distortional effect on a gravity-well due to its motion relative to other masses – the throat of the well is tilted to align with the moving body's worldline, and the degree of tilt reduces as a function of distance. For a rotating body, the gravitomagnetic distortions show themselves as a twist in spacetime in the region between the relatively-rotating bodies, sometimes illustrated within current theory as a "tipping over" of Minkowski lightcones (MTW, [53] box 33.2).

In a field theory, the gravitomagnetic field appears as additional polarised field effects superimposed on the normal "static" fields

Categories

We can define three main categories of gravitomagnetic effect due to the relative motion of masses,

- effects due to relative **acceleration** ("a-gm"),
- effects due to relative **rotation** ("r-gm"), and
- effects due to relative **velocity** ("v-gm").

The first two of these three are "official", and were described by Einstein in 1921. [40] The third is logically and geometrically required to exist, but is usually omitted from textbook descriptions, for reasons that we will explain later.

Velocity-dependent gravitomagnetic effects

Although the absence of detectable "starfield" effects has been used to argue that v-gm is known not to exist, the Carlip argument (section 23.3) makes v-gm essential to a geometrical theory of gravity, and explains that when a body moves with respect to the background starfield, the v-gm

i Will (2006) [131] 4-4-1 "Search for Gravitomagnetism" " *According to GR, moving or rotating matter should produce a contribution to the gravitational field that is the analogue of the magnetic field of a moving charge or a magnetic dipole.* "

effect cancels with the starfield's gravitational aberration effects to generate an "arena" that then operates according to Newton's First Law. Velocity-gm effects *are* measurable, less directly, when the background distribution of mass is not perfectly uniform (*e.g.* when a passing planet tugs on us as it passes), or when varying velocity disrupts cancellation (gee-forces). Velocity-gm effects are (in effect) used by NASA to slingshot probes around the solar system (with the calculations done in the time domain, using NM). We can also derive the existence of particle-fields from the phenomenon of refractive index (section 20.5), after which the Fizeau effect becomes explicable as a *v*-gm effect.

24.2. Broad, "qualitative" time-domain arguments

Distance differentials

The finite speed of gravity means that an approaching object is sensed as it was when it was further away, and a receding body is sensed as it was when it was nearer. [i] This makes a body's gravitational field appear stronger if it is receding and weaker if it is approaching.

Dropping rocks onto the Moon

If we drop a rock onto the Moon from an agreed height, we can calculate the expected change in velocity of the rock during its fall (it's "delta-vee") caused by the Moon's gravitational field.

- If the Moon is moving *away* from us when we drop the rock, the velocity-change will be greater because gravity has more time to act on the rock before impact.
- If the Moon is *approaching*, the velocity-change will be smaller because gravity has less time to act on the rock.

If we define the Moon's effective gravitational field strength by the velocity-change that it produces on a body dropped by us onto its surface, then we will assign the Moon a stronger effective gravitational field if it recedes and a weaker effective gravitational field if it approaches.

24.3. Black holes

If we believe that black holes have absolute horizons, [ii] then outward-aimed light that is emitted by an object as it falls into the black hole should remain frozen into the horizon surface, and needs to be unable to escape, as seen by any external observer.

If we now move away from the hole at *v* m/s, we need the light in the original surface now receding from us at *v* m/s to still be trapped. This means that the light that was stationary when the hole was stationary, is now retreating backwards, away from us, as the hole moves away from us. The light cannot recede any *slower* than *v* m/s, because otherwise the horizon would move away faster than the light. The horizon of a black hole, receding at *v* m/s, must drag light at *v* m/s.

> A moving gravitational mass drags light, and a moving gravitational *horizon* drags light *completely.*

---

i A physicist educated in the Twentieth Century may say that this is wrong, and that gravity somehow "throws its voice" to make bodies seem to be at their instantaneous positions. This anomalous "gravitational ventriloquism" argument was a "fudge" added to general relativity that has since been replaced by the Carlip argument (2000).

ii Elsewhere we will argue that gravitational horizons should *not* be absolute … but we'll invoke a GR1916 "Wheeler" black hole for this exercise due to its simplicity and familiarity.

24.4. Rotational gravitomagnetism requires the v-gm effect

A general theory of relativity has to predict that a rotating mass drags nearby light and matter around with it (to some extent), with the amount of light-dragging being a function of proximity and the body's associated field intensity. [i]

Rotational gravitomagnetism includes an obvious velocity component − if we watch a rotating star from within its rotation plane, the receding part of the star drags light away from us, and the approaching part of the star drags light towards us, with the strength of the dragging effect again going to 100% if the star has a gravitational horizon.

We can then extrapolate from the "rotating-star" case to a range of other situations involving rotating bodies (for instance, the case of a rotating ring, or the case of a ring of identical stars orbiting a common centre of gravity) to derive how far the existence of the velocity effect depends on an identifiable coexisting rotation or acceleration. The result is that for local physics, there seems to be no requirement for curved spatial paths or perpendicular forces − the velocity effect is independent, and has to exist for moving masses even in the absence of identifiable rotation.

This conclusion is supported by the "moving black hole" exercise in section 24.3, which gives 100% dragging for a horizon moving in a straight line at constant velocity, in the absence of any relative rotation.

> The general principle of relativity lets us derive a dragging effect for masses that move in straight lines at constant velocity relative to other masses.

24.5. Velocity-dependent gm underlies the other gm effects

The velocity-based dragging effect is the "basic" gravitomagnetic effect without which the rest of gm doesn't make much sense. The acceleration effect is the higher-order version of the velocity effect, and rotational gravitomagnetism can be broken down into (and explained by) its perpendicular velocity and acceleration components.

> GR's dragging effects due to rotating bodies can be broken down into two effects: a velocity gravitomagnetic effect, and an accelerational gravitomagnetic effect (which is just a higher-order version of the same effect).

> By initially expanding the number of basic gravitomagnetic effects from two to three, we can eliminate all special cases and simplify to just one law: the relative motion of masses is always associated with gravitomagnetism.

i Wheeler (1999): [48] page 232: "*As electric charge, going round and round in a circle, produces magnetism, so mass, going round and round in a circle, must produce a new kind of force gravitomagnetism.*"

24.6. Basic principles of momentum exchange

Momentum exchange is a basic part of physics:

1. The law of reflection says that if we throw a tennisball at 20kph at the front of a locomotive approaching at 60kph, the driver should see the ball approach at 80kph and leave at the same 80kph, and we should then see the ball coming towards us at 80kph+60kph=140kph. From our point of view, the interaction between ball and locomotive due to the elastic collision has caused the two bodies to exchange momentum, and the loco's speed should be fractionally slowed by the effort used to accelerate the ball.

2. In a mechanical approximation of the gravitational case, we could fit the locomotive with a swivelling arm (or a rubber rope) with a cup at the end, that catches the ball, is caused by the impact to swing around in an arc, and then releases the ball again in our direction, with the forces in the arm's mount due to the interaction of the train's momentum with that of the ball whipping the ball back at us faster than it was received. [i]

3. Another possibility would be to surround the back of the loco with a frictionless semicircular or horseshoe-shaped surface, wider than the locomotive, that can catch the ball on one side and return it to us via the other. If we throw the ball a little inside the curved surface it will make a number of bounces around the surface before emerging back in our direction, while if we carefully throw the ball to come smoothly into contact with the surface, parallel with it, it will slide around the curve (number of effective bounces=infinity), undergoing a smooth continuous acceleration towards us.

4. Finally, we could replace the curved surface with an "open" gravitational orbit. If we assign the locomotive its own intense gravitational field, we could throw the ball slightly to one side of the cab and watch as it marked out a horseshoe-shaped arc around the cab to return to us. It should again reach us with an increased energy and momentum, this time from having *gravitationally slingshotted* around the approaching train.

Regardless of whether the ball is returned to us by interacting with the locomotive via a direct collision, a flexible mechanical or elastic coupling, a curved accelerating surface, or a path through curved space/spacetime, the interaction has to obey basic principles of momentum exchange.

24.7. Gravitational momentum exchange gives gravitomagnetism

When two gravitational masses pass by each other and their fields interact, the resulting momentum exchange can be considered to be due to the partial, indirect collisions of the two masses via their gravitational fields, which act as proxies to transfer momentum from one mass to the other. [ii] [iii]

The result of momentum-exchange is that a moving gravitational source should exert a dragging force on nearby matter (and therefore also on nearby light). The deflection behaviour, described as a field or a spacetime distortion, is velocity-dependent gravitomagnetism.

---

i The principle is also used by the "shepherd's sling".

ii As with conventional collisions, which mass is said to be the momentum donor and which is said to be the momentum recipient will be a function of the observer's state of motion … but external onlookers will agree that momentum *is* being exchanged.

iii If a body's gravitational field can be considered as a *spatial extension* of its mass, then the collision of gravitational fields is quite literally a partial collision of the associated masses.

24.8. Gravitational momentum exchange is incompatible with special relativity

If we apply the momentum exchange principle to light, then in the case of the previously-mentioned rotating star, light travelling in the rotation plane and skimming the star should arrive at us with more momentum if it skims the approaching side of the star (dragged towards us) than if it skims the side that recedes from us (dragged away from us).

The light’s altered momentum will be associated with an energy-change, giving a redshift for recession and a blueshift for approach.

What of the light is not skimming the visible edge of the star but is *emitted* by it?

- If we assume that the conventional motion shift on the starlight is described by the SR Doppler shift equations, then if the dragging effect is an *additional* effect, the total motion shift will disagree with SR.
- On the other hand we might argue that it is wrong to treat the dragging effect and the Doppler effect separately – just as the star’s static field can be considered a smearing of the star’s rest mass, its gravitomagnetic field can be considered a smearing of its momentum. With this way of looking a the problem, the total gravitomagnetic shift that the light encounters along its path to us *is* the star’s Doppler shift.
  The problem with this more advanced interpretation is that, since the Doppler shift now has to be compatible with the idea that relative motion is associated with curvature, the Doppler relationships can’t be those of flat spacetime and special relativity.

Either way, the Doppler relationships must disagree with the SR predictions.

> Since bodies with significant gravity must have significant gravitomagnetic curvature when they move, which is not dealt with by SR’s flat spacetime model, special relativity cannot correctly describe the Doppler relationships for strong-gravity bodies.

> Since all bodies must have the same Doppler relationships, if SR is not valid for strong-gravity bodies, it is not valid for anything.

> Since a viable gravitational theory HAS to produce gravitomagnetic effects, full gravitational theories cannot reduce to special relativity. To protect SR and GR1916, we set aside these issues by pretending that *v*-gm effects don’t exist.

24.9. Summary

Gravitomagnetism, is a comparatively simple subject where we should have been able to make a great deal of progress, comparatively easily, and very rapidly.

However, since gravitomagnetism (applied consistently) immediately destroys special relativity, progress on the subject has been stalled since around the 1960s by the need to avoid contradicting SR.

> The near-absence of further theoretical work into gravitomagnetism is an example of the damage that special relativity is doing (and has been doing for the last half-century) to theoretical physics, and to science in general.

25. SR Argument 25: Relativity's wonderful experimental testing regime

25.1. Test theory

It would be deeply annoying, after performing and publishing a test of a theory's predictions, to only be told afterwards that it was invalidated by some obscure possibility that had not been taken into account. To avoid this, we have standardised frameworks – **test theory** – that provide ground-rules and context for experiments. A test theory also allows for shorter experimental descriptions.

25.2. The range for testing relativity theories

We can divide SR testing into two categories: tests of the principle of relativity (which will tend by default to also support Newtonian physics), and specific tests of whether special relativity is the *correct implementation* of relativity theory, which are more difficult.

The predictions of these two main relativistic systems are:

		Special relativity	Newtonian theory
Recession Doppler effect (v=$v_{\text{RECESSION}}$)		E'/E= $\sqrt{\frac{c-v}{c+v}}$	E'/E= $\frac{c-v}{c}$
“Transverse” effect		E'/E= $\sqrt{1-\frac{v^2}{c^2}}$	E'/E= $1-\frac{v^2}{c^2}$

Note that the predictions in the right-hand column are redder than those of special relativity, by a Lorentz factor. The Newtonian recession redshift (which can also be written E'/E = 1- (v/c)), is a consequence of other Newtonian relationships such as the Newtonian momentum law, $p=mv$. It is also the result that we get by assuming that light is “thrown” as a speed of c with respect to the emitter (Ballistic Emission Theory, “BET”), which was the main way that people attempted to implement Newtonian principles in describing light-behaviour in the Nineteenth Century. The Newtonian “transverse” redshift effect in the bottom-right corner is the **aberration redshift** effect already encountered in section 5.2, due to the forward deflection of rays.

25.3. Range of relativistic theories

If we assume that Nature obeys the principle of relativity, then any divergence from the SR or Newtonian equations must *also* be relativistic, and take the Lorentzlike form E'/E = $(1-v^2/c^2)^{\exp}$, where the exponent exp (expressed as a divergence from special relativity) has a value somewhere in the range -0.5 to +0.5. A positive exponent gives us a range of intermediate relativistic solutions, which will all generate key results such as E=mc^2, the atmospheric muon result, and the particle accelerator limit for direct acceleration (sections 2.1, 2.2, and 2.3).

The negative range (the blank cells on the left of the table) can be considered “unphysical” as it generates net energy gains in complex systems, allowing “infinite energy” machines.

The positive range generates net energy losses in complex systems, and also represents a transition between totally flat spacetime with no gravity or gravitomagnetism (SR), and a physics with maximum gravitational curvature (horizons) and maximum effect of moving matter on light (which can be implemented as a purely local field proximity effect instead of using ballistic theory).

25.4. The range we actually tested

According to some Twentieth Century SR testing literature, the range that we were *expected* to test for, was instead:

	"Classical Theory"	**Special relativity**	
Recession Doppler effect (v=$v_{RECESSION}$)	$E'/E = \frac{c}{c+v}$	$E'/E = \sqrt{\frac{c-v}{c+v}}$	
"Transverse" effect	$E'/E = 1$	$E'/E = \sqrt{1-\frac{v^2}{c^2}}$	

The justification for this was that we "knew" that the speed of light was globally constant for the observer, so that our propagation-based predictions were the $c/(c+v)$ relationship for recession velocity, and no transverse redshift, giving column **A**, below. On top of this, special relativity predicted longitudinal and transverse frequencies that were both redder than the **A** predictions by the Lorentz factor, due to time dilation (the "non-classical" shift component under SR being split out as the "**transverse component**"). Since only SR and similar/equivalent theories seemed to predict time dilation, and none of them seemed to predict anything more than a Lorentz redshift, we could test SR by validating the existence of the transverse Lorentz component. [i]

According to this viewpoint, there was no need to test the region further to the right, as it corresponded to no known theory, and had no theoretical significance.

25.5. Context for testing

The full spectrum of potential relativistic theories is given by columns **A**-**C**, plus intermediate solutions. Extensions of the range outside **A**-**C** tend to go "weird", giving negative values.

	A **"Classical Theory"**	**B** **Special relativity**	**C** **Newtonian theory**
Recession Doppler effect (v=$v_{RECESSION}$)	$E'/E = \frac{c}{c+v}$	$E'/E = \sqrt{\frac{c-v}{c+v}}$	$E'/E = \frac{c-v}{c}$
"Transverse" effect	$E'/E = 1$	$E'/E = \sqrt{1-\frac{v^2}{c^2}}$	$E'/E = 1-\frac{v^2}{c^2}$

Each member of this continuous spectrum of solutions can be identified by defining an initial reference solution (typically **A**, **B**, or **C**), and then a Lorentzlike deviation away from it, of the form $(1-v^2/c^2)^{exp}$. In the case of special relativity, we typically use **A** as our starting reference and say that the SR predictions are then "redder and shorter" than **A** by a full Lorentz factor. **C** is in turn even "redder and shorter" than SR, by an additional full Lorentz factor.

Once we see the full range, various patterns start to emerge: **A** gives the predictions for an absolute fixed aether stationary in the *observer's* frame, **C** gives the shift predictions for a speed of light fixed in the *emitter's* frame, and **B**, the predictions for special relativity, are exactly intermediate. Rather than predicting brand new classes of effect, special relativity's physical predictions are a simple "geometric-mean" average of the earlier relationships. [8]

i … either by writing $\boldsymbol{k}(1-v^2/c^2)^{0.5}$, and evaluating the value of ***k*** between 0 and 1, or writing $(1-v^2/c^2)^{exp}$, and evaluating the exponent in the range between 0 and 0.5

We also see that:

- Solutions become progressively "redder and shorter" towards the right-hand side, and "bluer and longer" to the left
- **B** is the set in which matter does not drag light.
  C is the set in which light-dragging at the surface of a fundamental particle is 100% (the "horizon-dragging solution").
- **B** is the set for particles with no gravitational fields and no gravitomagnetic effects.
  C is the prediction for bodies with extremal gravitational fields and extremal gravitomagnetism.
- **B** is the unique set in which the energy put into a system is the same as the energy released.
  Redder solutions to the right give a progressively stronger energy-loss, while those to the left give a progressively stronger energy-gain.
- (**B**-**C**)>**B** represents a range in which energy-losses and redshifts create stronger thermodynamic biases towards exothermicity, a stronger "arrow of time", and a more strongly evolving universe. **B** implies a static, unchanging universe.
- **B** is the unique set in which spacetime curvature remains zero regardless of the energy of particles. Solutions to the right associate positive recoverable KE with positive curvature, solutions to the left associate positive recoverable KE with negative curvature.

Since deviations from SR in the direction of **A** associate *negative* curvature with *positive* energy, these can be ruled out on the grounds that positive energy shouldn't be associated with anything but positive curvature. [i] This range, which gives energy-gains, can also be ruled out by adding the additional condition that infinite energy machines should not be possible.

If we believe in relativity and strict traditional energy-conservation, then everything but **B** is ruled out, special relativity is compulsory, and we have no compelling reason to do testing (except for the purpose of keeping up appearances). If we weaken the condition of energy conservation to "no infinite energy machines", then we lose the range **A**-**B**, and the appropriate range for testing is then **B**-**C**.

The existence of Lorentzlike deviations from special relativity, towards the red, is potentially of critical importance not only to inertial physics but also to gravitational physics and cosmology – both these subjects might be simplified if this deviation existed. Unfortunately, this deviation is in a range that our experimenters have been told to ignore, and to "calibrate away" if they find it.

25.6. Expectations overriding data: the Hasselkamp experiment

It is natural to want to disbelieve that experimenters could be eliminating potentially critical deviations from SR in their data – sceptics are cordially invited to read and analyse the write-up of the 1979 "true transverse" redshift experiment (Hasselkamp, Mondry and Scharmann, 1979 [107]), which gives an unusual insight into the way that our beliefs regarding what constitutes a "legal" result can influence final reported results.

---

i … although, since we find (section 25.2) that the leftmost range corresponds to the condition of reversed timeflow, some definitions that we project onto this range may need reanalysis to check whether the everyday definitions still hold under the condition of time-reversal. Crudely, the range **B**→**C** has an increasingly strong thermodynamic arrow of time, the range **B**→**A** has an increasingly strong *negative* arrow of time, and **B** itself (special relativity) represents fully time-symmetric physics (no arrow of time).

Analysis

The 1979 paper says that the experiment used velocities up to 9.28 x 10^8 cm/s , between ~0.8% and ~3.1% of the speed of light. It also *defines* the actual angle of the detector according to a "best fit" with special relativity. Although the detector was intended to be aimed at 90 degrees, the experimenters found that the angle had to be interpreted as being 90.5 degrees to get theoretical agreement between the data and SR. The cosine of 90.5° is ~−0.0087… multiplying this by the velocity ~0.031c gives a recession velocity component of v=~0.0002705c, and applying the longitudinal Doppler formula then gives again effectively E'/E=~0.0002705… [i] Given that the transverse Doppler effect at 3.1% of lightspeed is only ~0.00048, it would seem that the shift the equipment actually reported would have been perhaps ~0.00048+~0.00027=~0.00075, about one-and-a-half times the SR value, suggesting a value for the exponent of the Lorentz factor (for a detector *actually* aimed at 90°) of maybe about 0.75, rather than the 0.52 ±0.03 quoted in the paper.

The belief that the results *couldn't* disagree with special relativity as much as they seemed to let the experimenters argue that the detector must have been unexpectedly badly aligned, and justified redefining the supposed detector angle retrospectively to bring its results into agreement with the theory being tested.

In other words: belief that the right-hand range was invalid allowed us to start with a ~50% overshoot in our data, and still end up announcing an agreement with SR within a few percent. [ii]

25.7. Implications of the Hasselkamp experiment

Regardless of whether or not we believe that the 1979 experiment has anything reliable to say about *physics*, the test does say something compelling about *physics testing procedures*. The Hasselkamp experiment is a "proof of concept" that testing protocols can allow (or require) experimenters to take data that appears to disagree strongly with SR and transform it into an apparently "pro-SR" result, as long as the disagreement is too much redshift rather than too little.

More alarmingly, the Hasselkamp test is unusual in that the only reason we know about the issue is because it was dealt with in the "analysis" phase of the experiment. If the experimenters had been able to adjust the angle of the optics and do a further experimental run, the assumption that only the range **A**-**B** was legitimate would have entitled them to *physically* correct for the excess redshift by tilting the detector by a half-degree, treat the first run as a calibration exercise, and report an excellent agreement with SR from the second run without being obliged to mention that anything untoward had happened.

This makes it difficult for us to judge whether automatic correction of "impossible" excess redshifts in SR tests is likely to be a rare occurrence, or whether it might be common practice.

> SR testing tells us with a pretty high confidence that the real Doppler equations are not significantly *bluer* than those of SR (because *that* possibility is taken very seriously indeed by experimeters), but it's difficult to know whether or not they might be redder.

i … to a limited number of decimal places, this *looks* like the same number. Due to the tiny value of the recession velocity component, the Lorentz factor difference between the SR recession shift prediction and 1- v/c is tiny-tiny.

ii The obvious problem with the 1979 test is that we have an experiment that is very sensitive to small angular errors. There is no obvious independent way afterwards to tell whether the angle "*really was*" 90°, or 90.5°, or 89.5°, or something else. What we would normally want to do in this situation is mount the "problematic" section of the optics (or the entire optics bed) on a stage that can be rotated very exactly between two different orientations separated by 180 degrees. This will then give us two sets of transverse readings, one at (90°+ e°) and one at (90°- e°). The divergence in the two sets of values then shows that an angular error e° actually exists (if it does), gives us the magnitude, and allows us to confidently derive a set of averaged figures for our transverse data.

25.8. Why is SR testing so bad?

Why would we set up a testing regime for special relativity that excludes Newtonian mechanics?

A brutal answer is: because bad analysis allows us to report more impressive results for our experiments than good analysis.

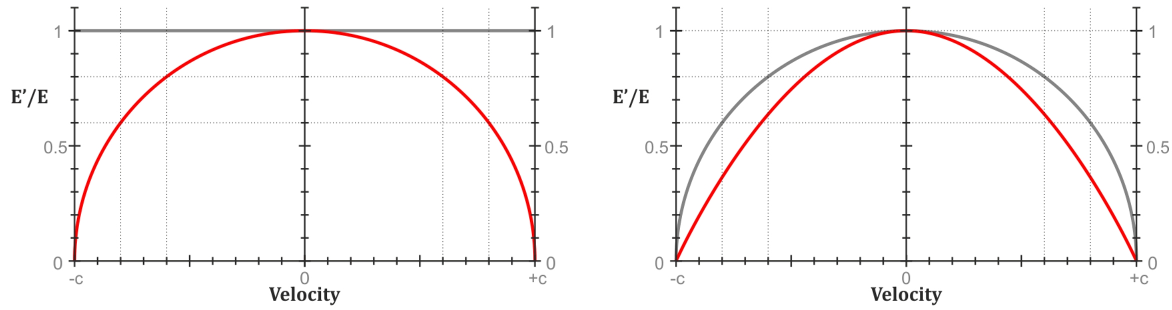

Figure 8: ***Transverse redshifts.***
" 'Classical Theory' and SR" (left), and "SR and NM" (right). The "Classical" predictions are a flat horizontal line through 1,1,1 (=no effect), the SR and NM plots both go through 0, 1, 0

Testing the difference between "CT" and SR (left-hand graph) is comparatively easy: testing the difference between SR and the Newtonian relationships (right-hand graph) is much harder. If we look at Figure 6, it is easier to prove the existence of an approximately SR-like transverse redshift effect that is at least as strong as the Lorentz factor (compared to "no effect at all"), than it is to show whether the actual curve is Lorentz or Lorentz-squared … especially given special relativity's habit of redefining distances, times, and velocities by a Lorentz factor.

In many situations, once the theory-specific elements of our definitions have been taken into account, there will be no physical difference between the SR and Newtonian predictions (see **muon test**, **E=*mc*²**, sections 2.2, 2.1). In others, there will be a difference, but one that is difficult to isolate.

A technical answer is: *even if special relativity is exactly correct*, we can expect the "centre of gravity" for the raw data that we collect *not* to be centred around the SR prediction, but somewhat over to the right (in the direction of Newtonian theory), due to complicating recoil redshift effects in the emitting atoms, recoil redshift in the detector material, and recoil redshifts in the intervening optics.

The role played by recoil effects is awkward to derive from first principles, and rather than include a further layer of difficult data-analysis, it is simpler for experimenters to ignore it, and choose a testing framework and belief system that classifies any deviation from SR to the red as being due to this sort of general complication. This allows the experimenter to eliminate any significant excess redshifts in the range (**B**-**C**)>**B** , without necessarily feeling the need to go into a lot of detail in their experimental write-up.

A more sympathetic answer is: "*because we can only test what we understand*". Options **A** and **B** are both well-understood. But if the data is really in the range **B**-**C**, then in order to be able to set up our experiment properly for that range, we will want to know more about how a consistent theory in this range might operate. If the textbooks and peer-reviewed literature can't tell us, we may decide that we cannot produce a provably valid test of **B** *vs.* **C**, and decide that it is safer (and more professional) to simply leave the range alone, and not even attempt to evaluate it.

25.9. Exclusions

It might be argued that we don't need to test the range that includes **C**, because **C** is historically associated with C19th Newtonian theory, which is already known to be wrong in some ways, and incomplete in others. But if we go down this path, we *also* shouldn't take a range that includes **A** seriously, as $c/(c+v)$ is historically associated with a Ptolemaic geocentric universe, which we know to be *even more* wrong.

Similarly, we could argue that we don't need to test **B**-**C** because anything redder than SR involves energy-losses (section 45), which breaks conservation laws. But in that case, we *certainly* shouldn't be testing the range **B**-**A**, because this not only violates energy conservation, it lets us build infinite energy machines.

What we *seem* to be doing here is basing SR testing on a perverse set of beliefs whose only obvious "purpose" is to make special relativity's verification test results look as good as possible. [108]

25.10. The market for lemons

By now we should realise that there are circumstances in which an inferior, incomplete, or faulty SR testing framework can have Darwinian advantages over a better and "more scientific" one. The experimenter using a standard "bad" test framework or a conveniently bad set of assumptions can report better results, more quickly, with less complex analysis, and using familiar references. In a research community where prestige and citability (and subsequent career success) is often a matter of "first past the post, wins", a team that has a more "casual" approach to their choice of testing frameworks and default assumptions has a competitive advantage over their more conscientious colleagues. There are distinct professional advantages to genuinely believing that a transverse redshift or a "transverse component" can only possibly appear under SR-based physics, despite of the mathematics insisting otherwise.

This is an example of "adverse selection", an idea usually based on information-asymmetry (Akerlof, 1970 [109]), in which lower-quality product ("lemons") drives out the good. If the end-user cannot distinguish between experiments conducted under "good" and "bad" assumptions, the experimenter is free to choose an inferior evaluation system that allows them to obtain better reported results, more easily. If there is no "market premium" for carrying out the experiment using more difficult and exacting criteria, the successful experimenter tends to select the "cheapest" and most effective test theory in terms of the ability to generate impressive figures. Eventually the market "fails", as end-users find they have no trustworthy way to tell whether a given experiment is likely to be good or bad, other than by how well it agrees with the established theory.

Once a body of experimental results has been published using "problematic" testing assumptions, a more scrupulous experimenter may find that they simply cannot compete unless they follow the same procedures as their competitors – otherwise their results may be more convoluted, their analysis less clear and less standardised, and their final results may be less emphatic than previous authors' papers. If their *more rigorous* analysis gives less conclusive results, the work might not even be considered worth publishing. The temptation is for an experimenter to either "*choose to believe*" in bad (but convenient) information, or to not involve themselves with tests of special relativity.

If we are also being told that special relativity is "fact, not theory", making further and better experimental tests seem redundant, then this also undermines the perceived worth and "market value" of expending significant effort and time to develop newer and more careful and reliable testing frameworks.

25.11. A modern GR example

These problems are not merely historical. Consider this high-profile experimental paper from 2019:

> Do et. al., 2019: [112] "… We detect the combination of special relativistic- and gravitational-redshift, quantified using a redshift parameter, ϒ. Our result, ϒ = 0.88 ± 0.17, is consistent with General Relativity (ϒ=1) and excludes a Newtonian model (ϒ=0) with a statistical significance of 5 σ."

The team measured the redshifts on a star rapidly orbiting what seems to be a black hole at the centre of our galaxy, and reported that they had validated general relativity's superiority over Newtonian theory to a five sigma confidence level, suggesting that the chances of this being a fluke if Einstein's theories *weren't* right might be one in three and a half million. Five-sigma confidence is generally held to be good enough to credibly claim the discovery of a new effect.

This sounds great until we look at the analysis. The redshift parameter that they were evaluating, was ϒ($\text{TransverseRedshift}_{SR}$ + $\text{GravitationalRedshift}_{GR}$), where both redshift effects were assumed to exist fully under GR ("ϒ=1"), and both were assumed to be *totally absent* under Newtonian theory ("ϒ=0"). When assessing the amount of redshift they found a value for ϒ that was significantly nearer to one than to zero. Success!

But we know full well that Newtonian equations *do* predict a transverse effect (section 5), and we've known since 1784 (Michell [11]) that Newtonian theory (and, apparently pretty much any other system that incorporates the principle of equivalence) also predicts an energy-loss in light climbing out of a gravitational gradient (a fact also usefully mentioned by Einstein in 1911 [12]).

To assign ϒ=0 to non-Einstein theory and imply that both effects are unique to SR/GR is simply wrong. It allows us to quote *magnificently* better figures, and obtain "*Einstein proved right!*" headlines, but that precious sigma rating depends on a serious misrepresentation of what other theories would have predicted in the same situation.

The paper demonstrates that in modern relativity theory, the lack of understanding of basic context, even amongst world-class researchers and experimenters, means that even a published peer-reviewed claim with an associated *five-sigma significance* should not be taken seriously.

25.12. Summary

Twentieth-Century SR testing seems to have been a remarkably unscientific process … or rather it has been a careful and exacting process based on an unscientific and easily-disprovable belief-set. We *chose to believe*, in the face of everything that mathematics told us to the contrary, that special relativity's predictions were in a separate range, that SR was always redder than pre-SR theory, and that that other theories didn't predict "transverse" redshifts or a transverse redshift component. This allowed us to believe that the evidence for special relativity was compelling.

A widespread belief in things that are not true can also undermine the credibility – perhaps unfairly – of other pro-SR results that do *not* have an obvious asymmetrical bias, as we cannot tell how far other experimenters' procedures may have been compromised by exposure to bad information.

> The range of potential relativistic theories corresponds to the redder range **B**-**C**.
>
> If massed particles have associated curvature, there should be a deviation from SR somewhere in this range. This gives the persistent mischaracterisation of "transverse" effects as being exclusive to SR, and the resulting focus instead on the **A**-**B** range, the potential to be one of the most disastrous mistakes in modern scientific testing.

26. SR Argument 26: “We know that SR’s geometry for empty space applies to moving matter”

26.1. Vacuum as a founding principle

Einstein’s 1905 paper “On the Electrodynamics of Moving Bodies” [1] bases the behaviour of moving matter on the assumed relativistic geometry of lightbeams in empty space.

This is obviously slightly problematic. Einstein states the theory’s second postulate as:

> Einstein (1905): [1] “ *… that the speed of light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body. …* ”

A logician may see this as a troublesome statement, as we start with a statement about empty space, and then part-way though start referring to a body. Is the region that we are considering to be regarded as “empty space”, or does it include an emitting body? Is the body considered to be *outside* the space? Is a second body (or a set of measuring apparatus) that receives the light, *also* deemed to be outside the empty space?

If the particles are deemed to be *inside* the region being modelled, then we may still treat the region as being empty, if we can produce some sort of supporting argument (or explicitly state as an additional postulate), that the presence of matter doesn’t affect the lightbeam geometry. ... but this additional postulate might be disprovable from the available evidence, which would then disprove the theory.

If the particles are deemed to be *outside* the region being modelled, then we can have an effectively pure vacuum and special relativity ... but our theory will be a theory of the relativity of the empty spaces *between* bodies, it will not be a theory that includes how the bodies themselves and any associated surrounding fields or distortions interact with the light. It will not be a theory of *how matter interacts with matter* via light, unless we add further hypotheses, providing additional opportunities for falsification.

Alternatively, we could agree that the presence of matter affects light, but argue that in simple cases it won’t affect light by very much. But this then demands a further set of theory or calculations, quantifying “how much is not very much”.

> Special relativity is a theory of how matter interacts with other matter and light, derived from the simplifying assumption that no matter is present.

26.2. Extending empty geometry to real physics

In order to be considered as a *physical* theory of moving bodies, special relativity requires the “flat” relationships of empty space to still apply when we add (as a minimum) a pair of atoms with significant relative motion, exchanging signals, to the previously empty region. In other words, it requires the motion of matter in the signal path (our two atoms) not to affect the propagation of light, compared to the situation that would exist if the matter was not present.

Einstein was aware that if one added *significant* quantities of matter to a region, one obtained a *significant*, obvious, measurable change in the region’s light geometry (refractive index, Fizeau effect), which is presumably why he specified “empty space”. He does not appear to have provided a rationale or supporting argument for why the introduction of matter to a region would not be expected to change its light-propagation properties.

Minimal relativistic physics

We might consider the minimum requirement of special relativity being allowed to be considered a physical theory might be that the addition of merely one or two atoms to an otherwise empty region (in order to be able to carry out some physics) does not disrupt the region's properties or invalidate its status as "empty".

Does a pair of atoms exchanging signals in *otherwise* empty space act as a region of *completely* empty space (SR), or does it act like a small section of a larger particulate medium (non-SR)?

- If we draw a box around a pair of atoms in otherwise empty space, and say that this region in isolation, behaves in exactly the same way regardless of whether it is neighboured by other similar regions, containing other pairs of atoms, then the initial two-atom region will behave as if it is part of a larger particulate medium (*see:* section 20.5 "*Does a single atom count as a particulate medium*"), and SR's derived lightbeam geometry will not apply. Special relativity will then not be valid even for the simplest possible case of two atoms exchanging signals in an (otherwise) empty region.
- On the other hand, we might argue that one or two atoms in otherwise empty space will behave as if the space is still empty. We could argue that the non-SR "particulate matter" behaviours are *bulk* behaviours, that only applies to larger groups of more than two atoms. However, if we tile space with multiple regions each containing a pair of atoms, and the behaviour within each region is now different because the regions now form a *group*, then this requires some form of intercommunication between regions, so that a smaller region can know whether or not it is part of a group. This requires either signalling or fields. Since SR-compliant signals presumably can't create non-SR behaviour, this leaves the possibility that regions can tell whether or not they are in a group by their proximity, as a field effect. ... But if the fields of surrounding atoms (electric, magnetic, gravitational, whatever) are intruding into a two-atom region and affecting the physics to give non-SR "particulate medium" behaviour, then presumably even an isolated two-atom region will also contain the same SR-invalidating fields due to *its own* two contained atoms.

With either of these two options, we end up with non-SR behaviour. Alternatively, if we start with the idea of refractive index and we try to apply the principle that particles have *no* fields (to try to make the best possible case for special relativity), we end up deriving the opposite result (again, argued in section 20.5), that a region has a refractive index *even if it contains only a single particle*, with the refractive index value then varying with a signal's proximity to the lone particle. This amounts to a derivation that a particle capable of being part of a particulate medium has an effective gravitational field, and that the SR description cannot apply in a vacuum *even if that vacuum is only "contaminated" by a single atom.*

Given his experience with particle arguments by 1905 (Brownian motion, [105] photons [106]), one might have expected Einstein, at *some* point in his subsequent fifty-year career, to have gone back and constructed the missing section of interlinking theory that connects special relativity to particulate matter physics, and that allows us to justify treating a space empty apart from one or two atoms as an "effective" vacuum in which SR applies.

> The fact that Einstein doesn't ever seem to have done this exercise suggests that perhaps it simply cannot be done. Perhaps it is not possible to study the interrelationship of refractive index and SR without concluding that SR cannot be a correct description, even for a minimal two-particle problem.

26.3. The "Taylor and Wheeler" position

An extreme version of the "vacuum" argument appears in the second edition of Taylor and Wheeler's "**Spacetime Physics**", whose back cover describes it as having become a standard for modern physics and relativity courses.

> Taylor and Wheeler, **Spacetime Physics: 2nd edition** (1992) [38]
> "BOX 3.1 THE PRINCIPLE OF RELATIVITY RESTS ON EMPTINESS!"
>
> *" What lies behind the Principle of Relativity? This is a philosophical question, not a scientific one. You will have your own opinion; here is ours. We think the Principle of Relativity as used in special relativity rests on one word: emptiness. "*
>
> *" But is space really empty? "Definitely not!" says modern quantum physics. "Space is a boiling cauldron of virtual particles." "*
>
> *" In the realm of classical (nonquantum) physics is space really empty? "Of course not!" says modern cosmology. "Space is full of stars and dust and radiation and neutrinos and white dwarfs and neutron stars and (many believe) black holes." "*
>
> *" Notice that for the very small and also for the very large, the "regions" described span both space and time – they are regions of spacetime . "Emptiness" refers to spacetime. Therefore we should have said from the beginning, "Spacetime is empty – except for us and our apparatus – with limitations described above." "*
>
> *" In brief, we can find "effectively empty" regions of spacetime of spatial extent quite a few orders of magnitude larger and smaller than dimensions of our bodies and of time spread quite a few orders of magnitude longer and shorter than times that describe our reflexes. In spacetime regions of this general size, empty spacetime can be found. In empty spacetime the Principle of Relativity applies. Where the Principle of Relativity applies, special relativity correctly describes Nature. "*

... in other words, in order to remove the scenarios in which special relativity is entitled to fail, we have to contract the domain of applicability for the PoR so that it only applies to vacuum.

26.4. Objections

We are entitled to take exception to the idea that the principle of relativity (in physics) is as limited as Taylor and Wheeler say. Historically, the principle of relativity has been about the physical observation that *complex systems of matter* appeared to operate in exactly the same way independently of any (simple) relative motion between them.

> J.B. Stallo, "*The primary concepts of modern physical science*" (1873) [115]:
> *" The essential relativity of all physical reality implies the persistence both of force and of matter … the principle of relativity. "*
>
> *" As there is no Unconditional in subjective thought, so there is no Absolute in objective reality. There is no absolute system of coordinates in space to which the positions of bodies and their changes can be referred; and there is neither an absolute measure of quantity, nor an absolute standard of quality. There is no physical constant. "*

It has been about the absence of a preferred frame for the behaviour of *physical systems*, as in Newton's famous "ship" example. We can invoke the principle of relativity to explain why, when we use the "Crazy Golf" course on the deck of a luxury ocean liner, the local physics behaves just as we would expect if we were stationary, despite the fact that the liner is moving relative to the Earth (and the Earth is spinning relative to its centre, and is also hurtling around the Sun, which is in turn circling our galaxy's centre). This principle of relativity does not specify that we need to remove the air, and the ship, and the putting green, and the planet – it requires that a physical system including *all* of these elements behaves in exactly the same way regardless of how the system moves with respect to other remote external bodies and systems. It does not require that the messy system of macroscopic, molecular, atomic and subatomic components making up the ocean liner *does not exist*, it requires that every atom, every molecular bond, every force, every field, every dimensional kink or curve making up the structure, and any other structure, exists and *geometrically translates.*

26.5. *Totally* empty space

Where T-W say that we need to specify that "*Spacetime is empty - except for us and our apparatus*", the requirement for empty space ***precludes the existence*** of us and our apparatus. If a theory of physics derives fundamental and universal laws from the results of measurement processes involving observers and their equipment in communication, then we cannot make the observers and their equipment somehow exempt from the laws that are being derived.

We are not Gods, and our apparatus is not supernatural, with some sort of diplomatic immunity to the petty rules governing ordinary matter. And if we *were* somehow able to adopt a ghostly form and persuade our apparatus to take on a similarly unphysical state, then the laws of physics that we would derive from using our unphysical apparatus would themselves be unphysical. What we *want* is the laws that affect real-life matter, and to derive *those*, we and our equipment have to be mundanely corporeal, and subject to the same laws.

Where T-W present special relativity as a success, it's literally an "empty" success: they say that "*Where the Principle of Relativity applies, special relativity correctly describes Nature*"), but the version of Nature that they refer to has no trees or clouds, or stars or planets, or interstellar dust, or people. It has no atoms or molecules, or gravitational fields, or any identifiable objects with which to carry out physics.

It may be more accurate to say that in the T-W worldview, special relativity correctly describes the *absence* of Nature.

> If Taylor and Wheeler's attempt to make relativity theory rigorous while incorporating special relativity results in the contraction of the domain of the principle of relativity until it only applies to a vacuum (with guest observer and equipment), then since we cannot rigorously extend the vacuum case to the case of a region including measuring equipment, the domain of special relativity shrinks to just *unobserved* vacuum.
> In other words, to nothing.

> Taylor and Wheeler also commit the cardinal sin in physics of claiming that a theory successfully obeys a principle, when this has only been achieved by *retrospectively redefining the principle to match the theory*. This makes it seem that the theory represents a perfect implementation of the principle, when in fact the *need* for redefinition implies that the theory is a failure.

26.6. Taylor and Wheeler *vs.* the principle of relativity

The principle of relativity has been one of the most powerful exclusionary principles in the history of physics. Its value lies in its power to unforgivingly destroy and dismiss non-compliant theories, acting like a logical scythe that in one sweep eliminates an entire field of long grass, leaving just a single stalk of wheat. The principle of relativity eliminates the impossible, and once you have eliminated the impossible, whatever remains, however improbable, must be the truth.

The problem that parts of the community started to have with the principle of relativity in the mid-Twentieth Century was that the principle started to work *too well*, and started to eliminate theories and systems that we were fond of and did not want to lose. It started to tell us things about our universe that we preferred not to hear, and as a result we started to put our own limits onto where the principle ought to be applied, to prevent the invalidation of things that we liked. If the principle of relativity applied to rotation and acceleration resulted in the invalidation of special relativity, then since we did not *want* SR to be invalidated we "moved the goalposts" and declared that the principle simply didn't apply in those cases (Schild 1960 [46]).

Special relativity has turned out not to be geometrically valid for moving objects with gravitational fields: and in a universe in which the principle of equivalence is correct, *all* objects have gravitational fields. If we want to maintain a belief in special relativity as a perfect theory, all objects need to be *excluded*, which is essentially what Taylor and Wheeler do.

Since the contraction of special relativity's range leaves relativistic physics open as a subject where relativity holds but special relativity doesn't ... suggesting an opportunity to construct a new theory of relativity, which might then endanger SR ... Taylor and Wheeler eliminate this potential existential threat to SR by further contracting the application of the *principle of relativity itself* to only cover the one case whether SR is rigorously provable (vacuum).

Finally, to avoid the possible return of the principle of relativity to overthrow SR, they banish questions as to the fundamental nature of the relativity principle from science altogether, to the realm of philosophy – this is declared to be "*... a philosophical question, not a scientific one*".

26.7. Taylor and Wheeler *vs.* physics?

A common response by physics people confronted with the Taylor/Wheeler viewpoint is that this is crazy. To physicists, the relativity principle was always about objects, bodies, systems and matter, and after all, Einstein's 1905 paper was titled "*On the electrodynamics of moving* ***bodies***". If the theory had been presented as only working in the absence of matter, we would never have taken it seriously as physics theory.

However, Taylor and Wheeler's position is entirely logical. If we truly believe that special relativity is correct, and that competing theories must be provably impossible, then, since the principle of relativity applied to more realistic matter insists on undermining the theory, and suggesting alternative systems, the principle must, in effect, be gotten rid of.

> If the Taylor-Wheeler position ("the principle of relativity only holds in a vacuum"), and the deletion of the general principle of relativity and the principle of relativity applied to matter are the price of saving special relativity, then special relativity is not worth saving.

26.8. Summary

> If the relativistic geometry of moving field-sources (moving matter) is different to the geometry of empty space, then the equations for empty space will not carry over. If light in the *presence* of matter behaves differently to light in the *absence* of matter, then, if we want to derive the relativistic laws that hold for real matter interacting with light, we have to deliberately apply the principle of relativity to non-vacuum behaviour.
>
> Since the resulting geometry will be different to vacuum geometry, a relativistic theory of moving matter will not correspond to special relativity or Minkowski spacetime, and will not incorporate SR as a physical limiting case. It will be a different class of solution.

27. SR Argument 14: "There is no experimental evidence against SR"

27.1. Perfect scores are untrustworthy

Cutting-edge experimental physics is, by its very nature, experimental. "World-first" experiments may be using new and untried techniques, where a supporting body of knowledge is not available, and we should expect a certain failure rate. Even with more mundane experiments, there should be a certain failure rate due to equipment malfunctions and the unforeseen, even if a theory is totally correct.

In order to evaluate the likelihood of a theory being right, it is useful to analyse this "outlier" data ("edge cases") to assess likely probabilities. If a class of experiment has 98% success rate, then we might decide that the theory being tested is likely to be right, and that the underlying agreement is likely to be ~100%.

However, if the reported agreement is *already* 100%, then we know that we are not seeing the full dataset. We cannot now *independently* assess whether or not the theory is any good, because we know (or can make a good guess) that the available data has already been helpfully edited by somebody, applying some sort of editorial rules to eliminate outcomes that they do not believe to be right. Without access to this "outlier" data, we cannot assess whether the raw data is likely to agree with the theory 90% of the time, 50% of the time, or 10% of the time. If a journal rejects results that disagree with a theory, or the experimenter self-censors by not submitting rogue results, and perhaps applies bias by accepting or rejecting individual experimental runs based on the data's agreement with expectations ("this was a bad run, let's recalibrate the gear and try again") … and their experiment is conducted under a test theory that requires them to either "adjust" for aberrant data or discard the experiment as a failure … then it is difficult for an outsider to judge the real level of agreement between theory and experiment.

Our rule of thumb in other situations is to regard suspiciously good scores as suspect. If a student scores 98% on a deliberately-extreme test, they may well be brilliant – if they score 100% we will tend to suspect them of cheating. The difficulty with the current system is that we cannot – unfortunately – tell the difference between "cheating" and genuinely excellent science.

27.2. Filters

We are sometimes told that there is not (and never has been) any data that contradicts SR. If we point out that during the early years of the theory, the main published relevant experimental evidence (by Kaufmann, between 1901 and 1905) concluded that the Lorentz-Einstein predictions were less accurate than other theories, [110] then we are told, okay, but we now consider those early experiments to *not be credible* – there is no known *credible* evidence against SR.

But what counts as "credible"? A journal may say that an experimental result that conflicts with known accepted theory and has no known explanation is not credible, and probably not worth publishing. Another journal may say that a *theory* that conflicts with the known experimental data, and whose predicted divergences have no known experimental support is also not worth putting into print.

But this leaves open the possibility that a unpublishable theory and an unpublishable experimental outcome may actually support and explain each other. While filtering research according to its agreement with current beliefs (as "quality control") is certainly *useful*, it can also create a self-perpetuating "echo chamber" in which our current beliefs are constantly being reinforced regardless of whether or not they are actually correct. [113], [108], [114], [111]

27.3. Historical perspective

Students of physics history will know that we have been in this situation before.

Newton presented a unified advanced system of physics in *Principia* [60] and *Opticks* [15] that merged together optical and gravitational principles. The speed of light was a function of local mass-density, which affected the density of an underlying aetheric medium, with this density variation deflecting light and matter towards the region of most concentrated mass, "*... with all that power which we call Gravity*" (Newton, *Opticks*, Qu.21 [15]).

An unfortunate inversion in Newton's logic meant that his system assigned a greater speed of light to denser regions, with light then being deflected towards regions where light had an *increased* speed. Huygens' principle disagreed, and by treating light as a wave, argued (correctly) that light must be deflected to the region of *slowest* lightspeed.

To the English physics community, Newton's system was supreme and couldn't possibly be wrong: in Joseph Priestley's *History of Optics* (1772 [117]), we are told that Newton's system, in which lightspeed is faster in glass than air, is in accordance with *all* known experimental evidence, and Huygens' principle is referred to in the past tense. John Michell's letter (1783, published 1784), also awarded Newton's faulty relationship a "perfect score":

> Michell (1784), [11] page 51: " *For let us suppose with Sir Isaac Newton (see his Optics; prop, vi paragraphs 4 and 5) that the refraction of light is occasioned by a certain force impelling it toward the refracting medium, an hypothesis which perfectly accounts for all the appearances.* "

Since we "knew" that Newton's system was right, we knew that Huygens' alternative description (which disagreed) had to be wrong.

Around thirty years after Priestley's book (and ~fifteen years after Michell's letter), Newton's "perfect" description was overthrown in ways that the community should have been able to anticipate if they had not been so convinced that the Newton system was invulnerable. **I. Bernard Cohen**'s preface to modern editions of *Opticks* repeats a comment that the apparent invulnerability of Newton's system held back the acceptance of wave theory for a century.

27.4. Perfect but wrong

It is difficult to improve on Eugene Wigner's characterisation of the problem:

> Eugene Wigner, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" (1960): [140] " *Considered from this point of view, the fact that some of the theories which we know to be false give such amazingly accurate results is an adverse factor. Had we somewhat less knowledge, the group of phenomena which these "false" theories explain would appear to us to be large enough to "prove" these theories. However, these theories are considered to be "false" by us just for the reason that they are, in ultimate analysis, incompatible with more encompassing pictures and, if sufficiently many such false theories are discovered, they are bound to prove also to be in conflict with each other. Similarly, it is possible that the theories, which we consider to be "proved" by a number of numerical agreements which appears to be large enough for us, are false because they are in conflict with a possible more encompassing theory which is beyond our means of discovery. If this were true, we would have to expect conflicts between our theories as soon as their number grows beyond a certain point and as soon as they cover a sufficiently large number of groups of phenomena. In contrast to the article of faith of the theoretical physicist mentioned before, this is the nightmare of the theorist.* "

This is pretty much the situation we now find ourselves in. Special relativity is *supposed* to have no counter-evidence, but is incompatible with the general principle of relativity, which is more all-encompassing and therefore (if implemented properly) more credible than the theoretical basis of the special theory. When a “shotgun wedding” between SR and the general principle gives GR1916, then this, in turn, is incompatible with quantum mechanics.

What appears to be a compelling case for special relativity when we look at a very narrow range of effects becmes less convincing when we consider a wider range of phenomena.

27.5. Context

Argument based on physical evidence can be worse than worthless if the correct context is not properly understood.

For an example we only have to consider the case of gravitational theory before the Carlip argument (section 23). Before this, we “knew” *for an experimental fact* that there was no such thing as gravitational aberration, and we similarly knew for an experimental fact that there was no such thing as gravitational dragging, because either effect would break Newton’s First Law. One could hardly imagine a clearer and more unambiguous disproof of the existence of dragging effects than the observation that a freely-moving mass, opposed by the combined dragging effects *of all the matter in the the entire outside universe* still obstinately refused to slow down. Faced with this observation, one would surely have to be delusional to still cling to the belief that dragging effects were still real.

And yet … once we realised in the 1990s that the predicted aberration and dragging effects *cancelled* for a uniform distribution of background matter, it destroyed our previous interpretation of the evidence. Once we realised that dragging effects were theoretically necessary, the empirical evidence *required* the opposing aberration effect to exist in order to cancel it out, and *vice versa.* Once we understood cancellation, the failure of cancellation when objects had *non*-constant velocities showed up in the description as gee-forces. The two unambiguous empirical disproofs magically transmuted into proofs.

Without adequate context, we cannot always understand what we are seeing.

27.6. Summary

An apparently perfect agreement between theory and facts does not necessarily make a theory correct, and does not mean that we are justified in dismissing or not bothering to research alternatives. New data overthrowing current beliefs may appear at any time.

Theories with perfect track records have been wrong before and will doubtless be wrong again, and part of the point of science (and more generally, of human intelligence) is to be able to be proactive rather than reactive, anticipate these potential failings, and plan accordingly.

28. SR Argument 28: "If there *was* any credible physical evidence against special relativity, we'd have noticed"

28.1. Non-SR energy-loss and thermal redshifts

While most of the effects that might be considered to undermine special relativity have *ad hoc* explanations (refractive index, Fizeau effect, QM retrofits), one has slipped through the net: thermal redshifts.

As previously mentioned, special relativity's solution is uniquely tailored to the idea that when two particles exchange signals, the signals behave as if the particles were not there – light is supposed to propagate through the region as if its was still empty space. This behaviour requires a unique Doppler solution – that of Lorentzian electrodynamics and Minkowski spacetime.

We can demonstrate this by a thought-experiment in which we send a low-power laser signal between two opposing walls of a room, and insert a moving glass bead into the signal path. If the bead moves along the beam (in either direction) at v m/s, the signal will be received at the far wall after undergoing two Doppler shifts with equal and opposite velocities, a redshift and a blueshift. With special relativity this gives a final frequency of

$$E'/E = \sqrt{\frac{c-v}{c+v}} = \sqrt{\frac{c-(-v)}{c+(-v)}} = 1$$

... the motion of the bead has zero effect on the final frequency, as expected. We can add as many additional objects in the signal path as we like, and as long as each individual shift obeys special relativity, the results will always perfectly cancel. The calculation for objects moving *across* the path is more complicated (because of angle-changes), but again, we require total cancellation.

By contrast, if Doppler shifts obey the Newtonian relationships we get

$$E'/E = \frac{c-v}{c} = \frac{c-(-v)}{c} = 1-v^2/c^2$$

In the Newtonian version of the exercise, the light emerges with a net redshift, and the greater the number of objects with different velocities that we put in the signal path, the greater the final cumulative effect, and the redder the final frequency.

Applied to gravitational physics, the result is a round-trip redshift for signals sent through a single gravity well, and in a universe containing many gravitational sources, cumulative distance-dependent redshifts over cosmological distances (Hubble redshift).

Applied to lab-scale physics, the result is a round-trip redshift for signals sent via a moving transponder, and a thermal redshift effect for signals sent through a transparent medium.

In lab experiments, *recoil* redshift effects will also generate a redshift regardless of which equations are correct, (which partly explains our "idiosyncratic" tendency to ignore excess redshifts in SR testing), but in 1958 **Rudolf Mössbauer** discovered that when a crystalline material absorbs or emits radiation, the recoil forces are spread through the whole lattice, making the recoil velocity (and the resulting recoil redshift) "effectively" zero. This rapidly led to the Harwell and Harvard groups using the new Mössbauer effect to measure gravitational-acceleration redshifts.

When the Harvard group [43] attempted to measure the tiny gravitational shift across a few floors of a university building, they found that their "recoilless" hardware insisted on reporting a thermal redshift effect that – if SR was right – shouldn't have existed.

28.2. Historical context

If this effect had been found as part of a separate experiment examining the properties of crystalline materials, the "impossible" thermal redshift (which appeared to be evidence that the SR equations were wrong) could have been rejected by peer review on the grounds of being incompatible with known physics, and we might never have known about it. However, since it was part of a larger experiment whose headline was that it claimed to *be proving Einstein right* (existence of gravitational redshifts), it seemed to get through the peer review filter.

1960/1961 was an odd and awkward time in relativity theory: The 1925 experiment by Adams, which had supposedly "*definitely*" proved the existence of the gravitational redshift in light from Sirius B (winning Einstein a medal from the Royal Astronomical Society) was quietly known by 1950 to be junk: [118], [119] Eddington had calculated the expected shift using a bad model and accidentally made an impossible prediction, which Adams had then confirmed experimentally, and – to make matters worse, the bad result had then been independently experimentally confirmed in 1928, before the mistake had been realised. This was embarrassing.
James Terrell had just gotten his paper though peer review in 1959 saying that part of standard SR teaching was obviously wrong, [26] and in 1960 the UK Harwell group's apparently pro-Einstein verification of accelerational redshifts turned out to inadvertently cause the community to realise that SR was incompatible with the GPoR and the principle of equivalence, making GR1916 an impossible theory, and apparently bringing down both of Einstein's relativity theories. [46] The community needed some good news, and if the Harvard group were producing a *legitimate* proof of gravitational redshifts, then perhaps the thing to do was to focus on this and try to ignore the fact that the experiment also seemed to have accidentally provided an unwanted experimental disproof of special relativity.

The importance of the previously-unpredicted thermal redshift effect is evident from its invalidation of earlier attempts to detect gravitational shifts in starlight. Our knowledge of stellar atmospheres was patchy, and since stars tend to be rather hot, we couldn't safely tell how much of a given shift might be "gravitational" and how much "thermal". These attempts had gone ahead partly because SR had supposedly proved that there were no additional "thermal" complications.

If the UK Harwell group's verification of GR couldn't be used because it broke GR1916, then the community seemed to decide that it would embrace the US Harvard experiment as the first proper validation of gravitational shifts, and hope that nobody would notice that *this* seemed to break SR, too. And they seemed to get away with it, as, for years, there seemed to be no peer-reviewed papers discussing the unpredicted and apparently SR-breaking effect (despite R.V. Pound of the Harvard group producing a two-part paper to remind people of the experiment and the "SOD" (for "Second Order Doppler") anomaly, [45] in 2000.

28.3. The Rindler defence

Although most texts avoid the subject, Rindler's book bravely tries to "spin" the effect as being explained by special relativity:

> Rindler (2006) [34] "4.3 The Doppler Effect": " *A canceling of the first-order contribution also occurs in the so-called thermal Doppler effect. Radioactive nuclei bound in a hot crystal move thermally in a rapid and random way. Because of this randomness, their first-order (classical) Doppler effects average out, but not the second-order (relativistic) time dilation effects. The former cause a mere broadening of the spectral lines, the latter a shift of the entire spectrum. This shift was observed, once again by use of Mössbauer resonance, in 1960 by Rebka and Pound.* "

This characterisation does not seem to bear up to any form of serious analysis.

If every particle sees the same Doppler relationships, then we only get cancellation with the full SR relativistic Doppler equation, which has time dilation already built-in. There is then no *additional* residual time dilation effect to be added, and we have E'/E = 1.

If we take the propagation shifts to be based on the c/(c+v) predictions for lightspeed fixed for each individual observer-particle, then these do *not* cancel. Looking at a pair of shifts, since $E'/E = c/(c+v) \times c/(c+(-v)) = 1/(1-v^2/c^2)$, a Lorentz-squared *blue*shift, when we then include the two missing SR Lorentz *red*shifts, we get cancellation and, as before, E'/E = 1.

If we declare a single overall frame for the propagation of light, then we can calculate each propagation shift independently by assuming that same (arbitrary) preferred frame. For our two equal and opposite velocities, we *do* then get cancellation of the propagation shifts as we alternate between using different Doppler equations that assume lightspeed fixed for a particle-source, and lightspeed fixed for a particle-observer: $(c-v)/c \times c/(c+(-v)) = 1$. But if we are switching between $c=c_{EMITTER}$ and $c=c_{OBSERVER}$ when calculating individual propagation shifts, then we must use the same frame references when calculating the corresponding Lorentz components, alternating between Lorentz blueshifts (moving observer is time dilated) and Lorentz redshifts (moving observed particle is time dilated). The end result is – predictably – no total redshift, E'/E = 1.

Rindler's argument seems bizarre. Given that the SR prediction in section 28.1 is unambiguously a "no net shift" result – which is the required correct result for SR – we cannot use mathematics to *change the results* of a deterministic calculation within a fixed global (Minkowski) geometry by breaking the SR Doppler equation into components and recalculating the components separately! This is not respectable mathematics.

Might there instead be an *additional* redshift due to acceleration? Quite possibly, but a separate acceleration-based redshift would contradict the SR clock hypothesis, and therefore once again end up invalidating SR. Rindler also suggests that the experiment *supports* the clock hypothesis: " *... it also yielded some evidence for the existence of approximately ideal clocks: in spite of proper accelerations up to 10^{16} g, these nuclear 'clocks' were slowed only by the velocity factor $(1 - v^2/c^2)^{1/2}$.* " Unfortunately, if we look back at section 28.1, the equation for velocity that gives a cumulative Lorentz-squared redshift after two opposing frame transitions – equivalent overall to the result of a cumulative Lorentz redshift per transition – is the Newtonian Doppler equation. Assuming that Rindler is correct about the absence of acceleration effects, and that his reasoning for this *second* calculation is correct (unlike the other one!), and that the accuracy of the data is significant, then what he has accidentally done is to show that the experimental result can be explained by saying that the SR predictions are wrong, and that it is the NM Doppler equations that are correct.

At this point, the only way to rescue special relativity is to suggest that perhaps the experiment might not be recoilless after all, which would mean admitting that we didn't really understand the "instrumentation physics" that was used to carry out a famous experiment.

28.4. Summary

The "SOD" thermal redshift effect represents an apparent (inadvertent!) experimental disproof of special relativity, produced by respected researchers, as part of a high-profile peer-reviewed and published experiment that we consider to be important and trustworthy.

It appears not to have been widely noticed or commented on.

29. SR Argument 29: "We cannot use external theory (particle, gravitational, cosmological, quantum) to disprove SR"

Anyone trying to use gravitational arguments to invalidate SR is liable to be told that it is inappropriate to use gravitational logic to attack a theory that is explicitly *not* trying or claiming to be gravitational. As special relativity's geometry explicitly doesn't attempt to deal with gravitation, invoking gravitation to try to undermine it is a futile exercise.

This is not altogether true.

If special relativity (rather than some other system) *is* correct for our universe, then the relationships of SR will then intrude into calculations for a range of other theories including the gravitational, particulate, cosmological, and quantum. We will require gravitational theory to agree that the motion shift of a strong-gravity body or a particulate mass agrees with SR, and to agree that the gravitational shift equation, as a function of velocity differential, is the SR version (giving "absolute" horizons). We will require cosmological theory to say that light emitted in the early dense universe and being received here-and-now should have a redshift that follows the SR/GR1916 gravitational shift equation, and we will require the statistics of quantum mechanics to conspire to produce the SR behaviours.

If these things don't work, and gravity, cosmology and quantum mechanics require a *different* set of equations, then while this doesn't disprove special relativity in the context of its own limited universe, the theory will be in trouble in a larger context. If the SR equations should turn out to be unworkable as the basis of gravitational theory, then, if gravity is a real "thing" in our universe, special relativity cannot apply within it.

29.1. Two horizon-based disproofs of special relativity (and GR1916)

A gravitomagnetic disproof of SR based on energies:

Consider the case of a moving black hole. According to SR-based physics, the horizon of a gravitationally-censored body located at *r*=2M [i]is an absolute horizon through which nothing can escape, by any means – the interior of the horizon is causally disconnected from the outside world in the sense that what happens inside the horizon stays inside the horizon. An outward-aimed signal emitted *at* the *r*=2M horizon (by an infalling light-source) is supposed to remain frozen into the horizon surface in perpetuity.

If this light is invisible to *all* external observers, then if the hole recedes at *v* m/s, the light is required to also recede at no less than *v* m/s,[ii] or else the horizon would recede faster than the light, and the light would be exposed (and could escape). The outward-aimed light, which would be said to be stationary of the hole was stationary is said to recede at *v* if the hole recedes at *v*.

This means that a moving black hole's horizon is exerting a 100% dragging effect on light. If we assume that the moving star's gravitomagnetic effect is "classical" (=continuous), and that the dragging effect extends beyond the horizon and does not suddenly stop abruptly at *r*=2M, then the momentum exchange between the star and light emitted just above *r*=2M should give the light a dragging shift, making the light appear redder of the star recedes than if it approaches. The

i The radius "*r*=2M" is shorthand for $r=2GM/c^2$. Since G is the gravitational constant and c is the speed of light, the G/c^2 part is a fixed constant, and is often omitted for brevity.

ii For simplicity, we are assuming that the emission point of the light *ep* is at the closest part of the receding horizon to the onlooker (who is situated at point *op*), and that *op*, *ep* and the hole's presumed centre *hp* all lie on a straight line.

dragging shift combined with the conventional SR motion shift gives a non-SR result. Alternatively, if the dragging effect is treated as being "dual" with the conventional motion shift (as in a gravitomagnetic theory of inertial physics), the Doppler relationships must be compatible with a geometry that changes shape with relative velocity, and therefore can't conform to Minkowski spacetime or special relativity.

In either case, the Doppler relationships of a moving strong-gravity body cannot be those of special relativity, and since metric-compatibility and the principle of relativity require a single set of Doppler relationships to hold for *all* moving bodies in the universe, we then have a general, gravitationally-based disproof that holds for all relatively-moving masses whatsoever.

A gravitomagnetic disproof of based on horizon behaviour:

As a variation on the argument, we can define the effective horizon of a gravitational body as being the critical surface at which a light-signal aimed directly at a given observer just fails to reach them (with the signal wavefront described as "frozen", neither approaching or receding). If outward-aimed light is emitted at the part of the r=2M horizon nearest to an onlooker, and the hole is moving away from the onlooker at v m/s, then the light will be said to be also moving away at v m/s/

But this means that the relevant section of r=2M horizon is no longer at the effective horizon – since it is *moving away*, the critical surface at which light would neither approach or recede must be somewhere *outside* r=2M, forming a further, *effective* horizon surface between the onlooker and the black hole. The hole's additional pull on light due to its recession changes the critical surface,

This new horizon ***is not a Wheeler horizon***: it is observer-specific and light that cannot reach the observer directly from the "twilight zone" between the horizons can still affect the region outside the effective horizon in ways that the onlooker can see: the light can reach dust co-moving with the hole outside the secondary horizon, and this dust can then be seen by our onlooker to be illuminated by light that wouldn't otherwise have been expected to be there.

This second horizon is an *acoustic* horizon, radiates indirectly and obeys the general behaviours predicted for a black hole by quantum mechanics (we can say that the presence of the dust converts "virtual" light from behind the second horizon into "real" light.) This complex nonlinear and often nonintuitive behaviour is the hallmark of an acoustic metric, and the only way that such a horizon can radiate indirectly (while still obeying the principle of relativity) is if its Doppler equations are redder then those of SR, by exactly one additional Lorentz factor. [23]

To summarise: the SR-based Wheeler black hole *self-invalidates.* We start by assuming that SR is correct, and use this to prove that gravitational horizons must be absolute … but to remain absolute even when they move, holes must drag light, and the Doppler relationship must be non-SR. There is then a further (relative) horizon in the region outside the hole whose behaviour is not absolute but acoustic, and whose local physics is non-SR. Since the non-SR behaviour changes what is seen even by distant observers, everyone then sees at least part of their universe to be obeying acoustic laws and logic corresponding to the $(c\text{-}v)/c$ equation-set. Since the same set of equations needs to apply everywhere, we live in a non-SR universe, and applying the altered equations to gravitational physics then tells us that our initial assumption that the r=2m horizon was absolute was not actually true. Assuming SR disproves SR.

This can be considered an "honourable death" for Einstein's general theory, in that the nature of its demise defines the characteristics of the theory that must replace it. But it is a death nevertheless , both for the 1916 theory and the 1905 theory.

29.2. Two cosmology-based disproofs of special relativity (and GR1916)

A cosmological disproof of SR based on energies:

Some sources on cosmology are adamant that cosmological redshifts (as a function of recession velocity), do not obey the SR Doppler relationship, but a simpler "first-order" Doppler law. The two candidates are then $E'/E = c/(c+v)$, and $E'/E = (c-v)/c$. Of these two only the second generates a horizon, making the default cosmological shift (if we *want* there to be a cosmological horizon) the "$(c-v)/c$" version. [i]

However, in order to be geometrically consistent, a geometrical theory of cosmology has to make the cosmological Doppler relationship identical to the relationship for redshifts due to gravitational curvature. [ii] [121] This forces the gravitational shift relationship to be "$(c-v)/c$" which in turn makes the conventional Doppler relationships $(c-v)/c$ as well. We then have a relativistic acoustic metric. We could try to bring this situation back into line with SR physics by running the argument in reverse, requiring cosmological shifts to agree with the SR equations used by GR1916 for gravity, but this would involve modifying current cosmological theory to (somehow) be compatible with flat-spacetime geometry, which might not be possible.

A cosmological disproof of SR based on horizon behaviour:

Alternatively, we could concentrate on just the horizon physics. If the cosmological shift is $(c-v)/c$, the cosmological horizon becomes **causally acoustic** – a signal generated behind the horizon, which is unable to reach us directly, is able to cross the line and hit a body in front of the horizon, which we can then see. We then have a cosmological counterpart of Hawking radiation, and the region intersected by the horizon operates according to acoustic metric rules rather then those of Minkowski spacetime. If the intersected region obeys non-SR acoustic physics, and any point in the universe including our Earth can be considered as straddling a cosmological horizon for some hypothetical distant future observer, then the physics of the entire universe becomes non-SR and "acoustic".

29.3. QM-based disproofs of special relativity (and GR1916)

Misner, Thorne, and Wheeler ("MTW")'s iconic textbook "Gravitation" (1973, [53] p.1066-) lists three criteria that any credible competitor to general relativity needs to be able to meet in order to be even considered worth testing: Criterion number (iii) is that the theory must mesh with other fundamental systems, including quantum mechanics.

Unfortunately for MTW, almost immediately after the book was published, we realised that, thanks to Hawking radiation effect (Hawking, 1974 [20]), Einstein's general theory *itself* didn't pass the essential test of "meshing" with QM. If we hold GR1916 to the same standards that MTW gave for GR1916's possible competitors, than the 1916 theory gets classified by MTW as "not sufficiently right to even be worth testing". Classical Hawking radiation plus relativity, requires the non-SR "$(c-v)/c$" relationship. [23]

i Suppose that (purely for the sake of simplicity), we treat the assumed dense state of the early universe as a point-singularity. We then expect there to be a censoring horizon between the singularity and us. This requires the equations NOT to be the $c/(c+v)$ set, since these do not generate a horizon until the relative recession velocity is *infinite*.

ii Light reaching us from an earlier, denser universe has in a sense climbed an uphill density gradient, and should reach us with a redshift. If the same curvature along a signal's path can be described either as due to cosmology or gravity (geometry doesn't care about causes), their equations must be the same. If the two effects are *not* dual, and the gravitational redshift needs to be added as a separate effect, then a lot of our cosmological calculations will be wrong.

29.4. Rescue attempt: Quantum gravity

Justification for quantum gravity

Since the GR community are somewhat reluctant to lose both Einstein's classical theories in favour of a new system that they haven't yet devised, it has become popular to argue that, rather than Einstein's general theory being dismissed as wrong for conflicting with QM, GR1916 and QM might be totally-included subsets of some hypothetical larger theory, referred to as **quantum gravity** (**QG**). The concept of QG is analogous to saying that we can reconcile the incompatible two-dimensional shapes of a square and a circle by adding a dimension and making both shapes the shadows of a single more complex shape, a cylinder. Similarly, even though the "outlines" of classical (GR1916) and quantum (QM) descriptions of physics disagree, they may represent "projections" into the classical and quantum domains of some larger, more complex structure.

While the optimism of this argument cannot be faulted, the reasoning appears to be somewhat "faith-based": in "the parable of the cylinder" we achieve two different silhouette shapes by shining a light at the cylinder *in two different directions*. In the case of the black hole information paradox, [149] where Einstein's general theory and quantum mechanics make two different physical predictions *for the same situation*, this is the equivalent of aiming a light at the cylinder *at a single angle* and getting two different incompatible predictions for the resulting shadow. Either the black hole gives off radiation that registers on a given defined detector or it doesn't. Classical and quantum arguments can give different explanations for *why* the detector should report a result, and these explanations can be *apparently* irreconcilable, as long as the final answer is the same … but that final answer really does have to agree, or else the whole exercise becomes pathological, and somewhat pointless.

Quantum gravity in a non SR system

The problem with invoking a hypothetical theory of quantum gravity to argue that current GR is *not* flawed (other than that no such theory yet exists [150]), is that any such theory, while it would be allowed to produce alternative (apparently incompatible) explanations of why agreed physical inputs into an experiment generate particular outputs, would still require the relationships between inputs and outputs has to be identical in both descriptions. If the quantum description says that radiation appears outside a gravitational horizon, then so must the classical description. The quantum description may say that the radiation arises through a pair production effects (Hawking radiation), and the classical description may say that it emerges through the horizon as a result of multiple accelerations distorting the path-geometry, but both must agree that the final radiation effect registers on a detector in exactly the same way, regardless of how it got there.

Effects analogous to Hawking radiation cannot be replicated under Einstein's general theory, or under any other classical gravitational theory based on the SR Doppler relationships. [23] A theory of quantum gravity reconciling classical gravitational theory with QM *does* seem to be possible, but only if the version of GR that is reconciled is not Einstein's. The current SR-compliant general theory, as it stands, is genuinely beyond rescue, rehabilitation or redemption.

29.5. Rescue attempt: The holographic principle

An attempt to rescue the 1916 theory from failure was made in the 1990s by invoking the **holographic principle**. This principle [151], [152], [153] notes that if we surround an arbitrary three-dimensional volume with a two-dimensional surface, then since all our measurements of that region are taken *though* that surface, we can't be sure whether the region really exists, or whether it is a non-region bounded by a surface that cleverly *simulates* a non-existent interior physics.

Suppose that we watch light from a distant galaxy being lensed by a second galaxy in its path. If we delete the region of spacetime containing the intermediate galaxy and replace its "bleeding edges" with a holographic surface, we can describe the light as hitting and being absorbed by the far side of the surface, and the resulting information then propagating around the surface and reconverging on our side, and spitting out a modified version of the same light. As the information spreads around the surface, it encounters and interacts with all the data that belongs to the supposed interior physics, and when it emerges, it shows us a modified image of the central galaxy, complete with the expected lensing effects. [i]

The attempted application of the holographic principle to the black hole information paradox was based on the idea that if we identified the horizon with a holographic surface, it would allow massenergy and information to *seem* to enter and exit the horizon without anything actually passing outward through *r*=2M. [ii] [iii]

The flaw in the argument was that for the holographic surface to be emitting (rather than silent), then the interior physics that the surface would be simulating would have to be something other than that of Einstein's general theory! Assuming that holographic surfaces are not allowed to decide *which* physics they emulate, and that all holographic surfaces must mimic volumes that operate according to the same physical laws, we would then have a situation in which our entire universe must appears to behave according to rules that are not those of the 1916 theory, or special relativity. If every measurement we take appears to be obeying non-Einstein laws, then **(a)** we will want to derive what those laws *are*, and **(b)** this is functionally equivalent to saying that Einstein's theories don't correspond to apparent reality – and are simply wrong.

Although useful in other situations, the holographic principle can't be used to change definite physical predictions.

29.6. Summary

If special relativity is considered to be fundamental foundation theory, then its relationships must also appear in other external theories (*e.g.* whichever Doppler shift relationship applies for moving bodies must also apply within gravitational theory, for gravitational shifts).

The SR relationships, in these other contexts, produce logical clashes.

It is legitimate to regard these failures as possible disproofs.

---

i An alternative variant of the holographic principle uses a surface to enclose a region of *time* rather than space. If we define a "now" surface that intersects the entire universe, and say that information is neither created or destroyed over time, then the "now" surface will contain all information necessary to reconstruct the volume of spacetime that we call "the past"and also the other volume we call "the future".

ii For black holes, the holographic principle can be considered as a descendent of the 1980s black hole **Membrane Paradigm**. [154], [155] The Membrane Paradigm treats black hole physics as it is *seen* to behave by a distant observer. Since the far observer never sees any matter crossing the horizon, and instead sees infalling material to be moving ever-closer to *r*=2M without ever actually reaching it, the Membrane Paradigm treats all infallen matter as forming an apparent two-dimensional film at (or at a Planck distance above) the *r*=2M surface. Within observerspace, infalling matter deposits its information at this membrane, which can then re-emit the same information as Hawking radiation. The requirement that the membrane appears to present all of the hole's information then tells us that the black hole horizon's minimum surface area needs to be proportional to the quantity of contained information.

iii A difficulty with making horizon area proportional to information-content is that topologically speaking, when seen from the other side, the same surface also "encloses" the entire outside universe, which will normally contain substantially more data. It may be more reasonable to invoke a connection theory, and to relate the horizon area to the number of parallel *connections* between information inside and outside the hole.

30. SR Argument 30: "There is no disagreement between quantum mechanics and special relativity"

Unlike classical field theory, in which spacetime adopts a shape that is smoothly continuous, and where causal relationships are transmitted smoothly through the classical surface, quantum mechanics is typified by discontinuity. The subject is partly rooted in another of Einstein's 1905 papers, [106] in which a statistical analysis of the spectra of hot bodies suggested that light couldn't just "seep" out of atoms, but required a definite threshold energy to escape.

30.1. Philosophy

Under the usual interpretation of QM, light has both "wave-like" and "particle-like" aspects

Under most of the common approaches, we say:

> " Light exists as particles (photons). Light is always generated as discrete quanta of energy, and is always received as discrete quanta of energy. Any attempt to intercept light between its emitter and receiver with a piece of measuring equipment to assess the nature of light will again only ever report discrete amounts of energy. It is therefore philosophically wrong to hypothesise a different wavelike nature of light in *transmission* that cannot on principle ever be directly verified. Light is emitted, absorbed, *and travels* as photons. However, these photons show weird counter-intuitive effects when we are not directly looking at them, they show interference patterns (wavelike behaviour), and sometimes the number of photons detected is less than or more than the number emitted.
>
> Photons obey statistical laws rather than classical laws, These wierdnesses are simply part and parcel of quantum theory, in which we cannot expect normal rules or intuition to apply. "

Under a "sampling theory" interpretation of QM, we say,

> " Light propagates as waves, but is always unavoidably *measured* in discrete quanta due to the inherently quantised nature of our measuring equipment. Since this forces us to see light as particulate regardless of what its underlying nature may be, there is no deeper significance to the observation that we always measure light as quanta. The photon exists at generation and detection – it does not necessarily exist during transmission.
>
> Since light can only be absorbed in quantised amounts, the excess fraction of energy of an incoming signal that cannot be absorbed and detected remains in the environment, bouncing around as background noise until a convergence of noise-and-noise, or noise-and-signal with matter is sufficiently intense to trigger another absorption event. If the environment is inherently noisy, it can be impossible to predict for a "quiet" signal whether or not a detector triggers and reports a photon (noise dominates over signal). However, over time, the consistency of the signal and the inconsistency of the noise (the noise tends to self-cancel, the signal doesn't) will allow the signal to dominate over noise, and the signal can emerge from a large number of apparently individually-random events.
>
> System noise allows a detector to trigger when the signal is otherwise below the nominal threshold, or to fail to trigger when the signal is otherwise above the threshold (or sometimes to trigger when there's no official signal at all). But although we do not expect to be able to predict an individual event with certainty at the quantum scale, the noise is not *truly* random, it consists of the leftovers of all previous detection events. While many things are *not predictable*, there is nothing *truly* random in the system (preservation of microcausality). "

While the second interpretation is attractive, in that it demystifies quantum theory and eliminates the "spooky" aspects, the reason it has not been taken up is that it doesn't work with current classical theory. Our current SR-based classical theory, plus quantisation, does not yield the correct QM behaviour.

30.2. Conversions, and quantum gravity

Ideally, we would like to be able to convert a QM description into a classical description and *vice versa*:

- The classical system plus sampling, should give the QM statistics
- The QM statistics, averaged over time, should yield the classical system

In practice, something about our current system of classical physics stops this working. If we model the simplest conceivable hypothetical particle, a point (or pointlike) mass, cloaked by a horizon, with no properties at all other than mass and state of motion, our classical theory of the particle's external curvature (general relativity) and our quantum description (QM) are in fundamental disagreement. With SR-based GR, the horizon temperature is zero and the particle can never emit a signal, with QM the horizon temperature is nonzero and the particle must eventually radiate *something* (even if it destroys itself in the process). GR1916 physics, plus quantisation, does not give QM.

In the case of black holes, we also have the problem that, if Hawking radiation emitted outside the black hole is *truly* random, we lose **microcausality**, and (because very large events can be triggered by very small events) also lose larger-scale causality. If Hawking radiation is *not* truly random, then the information encoded in it would seem to belong to systems inside the hole's horizon – when a black hole shrinks by throwing off massenergy in the form of Hawking radiation, the radiated information (as well as the energy), originally belonged to the hole. [i] Quantum mechanics therefore conspires to make it *look* as if general relativity has the wrong description of classical physics.

30.3. "Copenhagen" and "hidden variable" interpretations

Two major interpretations of how to *think* of quantum mechanics are loosely referred to as the Copenhagen interpretation (Bohr/Heisenberg), and the hidden variable interpretation.

In the Copenhagen Interpretation, ("CI"), the statistical descriptions given by quantum mechanics are reality. The statistics do not refer to the statistical behaviour of some underlying system, they are self-contained.

In the Hidden Variable Interpretation, ("HVI"), the statistical descriptions apply to the behaviours of more conventional physics below the quantisation threshold.

In the CI, things seem to happen at random, and events are fundamentally unpredictable. Einstein famously took a strong dislike to this idea, and argued that "I cannot believe that God plays dice". Einstein produced a series of arguments trying to show that the CI could not work, and the idea's supporters responded with a series of counter-arguments showing that, yes it could. Eventually the exchanges seemed to show that the Copenhagen interpretation was solid, and held up even in the face of a determined attack.

---

i Hawking did explore for a while with the idea of abandoning microcausality to bring QM more into line with GR1916, and protect the SR-based GR1916 idea that information could not possibly travel outwards through r=2M. The general verdict on this work was that it was worthy, but unconvincing.

However, with renewed studies of the microcausality problem in the context of Hawking radiation, the mainstream view (circa 2020) seems to be that microcausality *is* preserved under QM, meaning that even if God *does* play dice, the outcome of each roll of the dice, while not predictable, is also not truly random. If one had knowledge of every physical property of every atom involved in the dice-roll, to arbitrarily fine resolution, and and programmed them into a computer simulation, we could predict how a dice roll would come out, every single time. This means that *information is preserved* – if we put a tiny amount of information into a system, it might not be enough to produce an immediate visible effect, but eventually the system will produce an event or events that reflect that information. The information is not erased or obliterated by any truly random processes while it waits to be rediscovered, by another dice-roll.

The *strict* Copenhagen interpretation was increasingly undermined by the emerging subject of Hawking radiation. Hawking radiation is not limited to the emission of single particles – if the HR description said that there were objects on the macroscopic scale, whose presence was real for some observers but that other more distant observers could only sense indirectly, then QM introduced a class of macroscopic behaviour in which elements of reality existed that were only detectable indirectly. This undermined the principle that if a thing could not be measured directly (by a given observer), it did not exist for them and only had a statistical existence.

Utility of a non-Copenhagen interpretation

The argument in favour of a CI-style interpretation is that if it *really does* provides the most complete description of physical behaviour possible, there will be nothing left for a "hidden variable" theory to do.

However, we have seen in the last few decades that applying a more classical approach to QM is genuinely useful. In classical physics, we can usually use our physical intuition about how an experiment ought to play out to notice when something has gone badly wrong with our calculations, but in the case of purely "statistical" QM it can be difficult to know whether a "surprising" prediction is genuine or is the result of user-error. In the case of the initial work towards predicting Hawking radiation, a number of theorists initially dismissed the prototypical arguments simply because (what we now know to be) the correct calculations disagreed with their expectations (Thorne 1994, §12 [22]). A theory that cannot make proper predictions has limited use – but a theory that *does* make proper predictions can also be problematic if the theory is so complex that we cannot be sure what those proper predictions *are*.

In the case of Hawking radiation ("HR"), our willingness to accept that the result *was* a genuine prediction was helped by the growing realisation that analogous effects (involving radiation through horizons) did appear in some classical models, and that the idea was therefore not as definitionally contradictory as it first appeared. *Understanding* the theory was useful – it was all very well to say "shut up and calculate", but if we did not understand what we were doing, we could not be sure that we were calculating *correctly*.

If we invoke a classical acoustic metric as a classical "toy model" for quantum behaviour, many of the weirder behaviours of Hawking radiation become obvious. We can say,

> "***Of course*** *gravitational horizons radiate and are hotter the smaller they are … **of course** the region around a horizon can be considered as populated by particles that cannot be directly seen by an arbitrarily distant observer but are physically real for us, and **of course** the supposed position along an artificial particle trajectory where the particle and antiparticle are supposed to be created is a function of observer position … once you understand the underlying mechanism, it's **obvious**!*"

30.4. Hawking radiation

Hawking radiation under quantum mechanics

Under the originally presented "naive" description of HR, we say that the radiation effect appears outside the horizon as the result of particle-pair-production effects. We say that small-scale randomness makes small regions of spacetime fluctuate, with momentary energy-peaks creating particle-pairs (*e.g.* electron-positron pairs) which usually then mutually self-annihilate and pay back their borrowed energy before we've had a chance to even realise that they were ever there.

If this process occurs on a gravitational gradient, and the pair is aligned with this gradient, then the associated tidal forces can pull the pair apart before they have a chance to destroy each other. One half of the pair is swallowed by the hole, and the other escapes. Because tidal forces are strongest around smaller holes and get very weak around larger holes, smaller holes give off more Hawking radiation.

Hawking radiation in dark star models

In an Eighteenth Century "dark star" model if the sort suggested by John Michell, ballistic emission theory says that, since the escape velocity at a distance of r=2M equals the speed of light, particles travelling at or at any less than the speed of light, and emitted below r=2M (below the gravitational horizon) cannot completely escape from the body's gravity along a ballistic trajectory: they will always be turned back. [i]

However, while they are visiting the region outside r=2M, chance collisions with each other (or with passing matter) can knock a particle free along an accelerated path, to be seen by a distant observer.

In a dark star model, the region just outside the horizon is illuminated by visiting light that a distant onlooker cannot see: however if a spaceship passes through the region, light can be reflected off the ship, which the distant observer then sees being illuminated by light that they may feel shouldn't be there.

The analogue of this under quantum mechanics is to say that the visiting light is *virtual* – its existence cannot be verified directly, but has visible consequences, and the interaction with the spaceship converts a virtual light0corpuscle into a real particle.

Every freaky detail of how QM says that observers see Hawking radiation seems to have a counterpart in the dark star model.

Hawking radiation in acoustic metrics

The dark star model historically used ballistic emission theory, which doesn't work with wave theory (or with sensible metrics). If we force wave-compatibility by making a particle's influence on the speed of light purely a field effect, dependent on mass and proximity, then the ballistic theory, smoothed, and supporting local lightspeed constancy everywhere, turns into an acoustic metric theory. The Hawking effect, whereby particles can be indirectly radiated through a horizon along accelerated path, remains.

Hawking radiation under GR1916

Under Einstein's general theory, the Hawking effect does not exist. Light emitted lower than

i If we define a horizon as being the threshold between visibility and non-visibility, GR1916 and Newtonian theory agree that the horizon is at r=2M. Under GR1916, r=2M represents the point ***at*** which light ceases to be visible to a distant observer, under NM, it is the point ***beyond*** which light ceases to be ***directly*** visible to a distant observer.

r=2M not only can't completely escape on its own, it never ventures outside the horizon at all. A spaceship that has just fallen through *r*=2M, and which tries to fire its engines to escape, is doomed to fail: the causal consequences of its acceleration cannot move outwards to influence the geometry of the horizon.

Quantum mechanics from classical theory

Suppose that we assume that the escape mechanism in **(b)** and **(c)** is real, but that we believe (wrongly) that we live in a GR1916 universe. What is the result?

We aim a detector at a distant "bare" black hole that has no accretion disc or external complicating physics, and eventually receive a particle (perhaps an electron) that has migrated out of the hole along an accelerated path, thanks to a tortuous and unlikely set of collisions with its fellow-prisoners.

Receiving the electron, we explain the particle's arrival by extrapolating a simple ballistic trajectory for it, from its final velocity. Realising that it came from the direction of the hole, and remembering that nothing can escape from r=2M along a simple trajectory if it starts out moving at any less then the speed of light, we describe the first part of the projected trajectory as being superluminal, and since particles approaching at more than lightspeed should be seen as being time-reversed, the artificial description then has the first part of the path time-reversed, describing as a *positron* that moves *away* from us. The junction of these two path sections then describes the production of an electron-positron particle pair outside the horizon, with the electron escaping and the positron being swallowed by the surface.

At this point we have successfully converted a classical description of Hawking radiation, with local causality and smooth local geometry at all points, in which particles genuinely escape outward through the (effective) horizon, into an artificial statistical description in which particles are generated *outside* the horizon, discontinuously, as particle-pairs. We have successfully connected the "naive" 1970s quantum-mechanical description with a classical explanation. …

Incompatibility with special relativity

… but this "integrated" description does not work with SR-based physics.

Under the NM-based description we can drop off a stationary spaceship at the *r*=2M horizon, and it can fire its engines and try to escape. With the Newtonian "$(c-v)/c$" relationship applied to gravitational shifts, using $v=(-c)$, the spaceship initially sees the incoming gravitational blueshift on light to be a mere doubling. As it fires its engines to escape, its physical acceleration makes the blueshift worsen (acceleration blueshift), but the situation does not seem *utterly* hopeless.

Under SR-based GR, the incoming blueshift calculated for the horizon using the SR Doppler law, $\sqrt{(c-v) / (c+v)}$, $v=(-c)$, is *infinite*. The stationary ship cannot even *exist* at r=2M, where it would experience an infinite inward radiation pressure, and an infinite temperature, and firing its engines would only make things worse … and if it did somehow survive and escape, an infinite amount of outsider-time would already have elapsed by the time it got free.

> Quantum mechanics cannot mesh with classical physics if classical physics is based on special relativity.
>
> A theory of quantum gravity, merging QM with GR, requires a general theory whose basic equations are redder than those of special relativity.

This is in broad agreement with the result we've already obtained from curved-spacetime principles, that, if massed particles have associated gravitational fields, the equations for relativistic physics must be redder than the SR set.

All we have to do now is to work out exactly *how much* redder they need to be.

Newtonian relationships as a precondition for quantum gravity

What if we "perturb" special relativity? Can we make some tiny adjustment or correction to SR that will fix things? Treating the infinite horizon blueshift as the immediate defining feature that makes SR incompatible with QM, we can graph how the blueshift seen at the horizon varies as a function of a Lorentzlike deviation from NM, to find the range of possible solutions that have a fighting chance of working. [23]

This graph shows that the NM result (E'/E=2) is a "cliff-edge" solution – anything redder than NM yields E'/E=0 (which defies analysis), and solutions bluer than NM, down to SR, all give E'/E=∞.

> Classical Hawking radiation cannot be implemented with a "small" correction to special relativity. The merging of classical and quantum theory requires the SR equations to be reddened by precisely one full additional Lorentz redshift.
>
> Solving the problem of quantum gravity requires a reversion to the Doppler relationships of Nineteenth Century Newtonian theory.

30.5. Namsrai and stochastic quantum mechanics

If we try to measure the position of a fundamental massed particle, the Heisenberg uncertainty principle says that there will be a certain amount of uncertainty in the measurement (when we get to very small scales, the difficulty of measuring certain quantities without our measurement attempt changing the data becomes not just inconvenient, but a fundamental limitation of the physics. [i] [ii]). If we measure the particle's position in a thought-experiment, and repeat the thought experiment over and over, the resulting probabilistic scattering of different positions will build up to produce a density map, which represents the apparent classical distribution of mass and momentum that, with noisy quantisation, would give QM statistics.

This method lets us derive the hypothetical shape of a smooth classical spacetime around a particle that would need to be correct in order for that classical theory to mesh with QM (Namsrai, 1984 [160]).

The shape sketched by Namsrai is essentially a tilted gravitational well, with the depth of the well giving the mass, and the tilt of its throat giving the relative direction and speed of motion. The particle's mass is smeared out into the surrounding region as a field or as a curvature distortion, and the momentum is smeared out as a gravitomagnetic field or as a geometrical tilt of the rest-field distortion. We then have a description in which particles have mass-fields and

i For large objects such as buildings, we can see them and measure them without their being destroyed in the process. For very small objects, the wavelength of light we have to use to see them can be in the x-ray or gamma-ray range, and the attempt to measure small structures can end up destroying them.

ii **Conjugate variables:** the bulk sales of Mars Bars across North America can be considered classical physics, a continuously variable system where sales are affected by weather, season, time of day, sporting events, national mood, and so on. As we zoom in on individual quantum "sale" events (nobody buys half a Mars Bar), quantum mechanics starts to apply. We can declare with confidence that someone in California will buy a Mars Bar at mid-day (give or take a minute) as long as we do not specify *where*. We can also say with confidence that a specific supermarket checkout will sell a Mars Bar, if we do not specify *when*. An individual sale can be predicted statistically to high accuracy with regard to either position or time, but not both together. In this situation, the "time" and "position" of individual sales events are conjugate variables.

gravitomagnetic fields, and we once again end up with a gravitomagnetic theory, using relativistic acoustic metric instead of SR's Minkowski metric.

Although the problem of reconciling classical field theory with quantum mechanics is considered unsolved, we can derive a classical field theory *from* QM that quantises to give QM statistics. However, that form of field theory is not compatible with special relativity.

30.6. Direct and indirect observation

Quantum mechanics and SR-based physics differ philosophically over the subject of indirect observation. With SR physics, the entire universe is visible and directly accessible, while with quantum mechanics we sometimes have to deal with information whose existence has to be inferred indirectly.

> Heisenberg (1989), [156] " *To my astonishment, Einstein was not at all satisfied with this argument. He thought that every theory in fact contains unobservable quantities. The principle of employing only observable quantities simply cannot be consistently carried out. And when I objected that in this I had merely been carrying out the type of philosophy that he, too, had made the basis of his special theory of relativity, he answered simply: 'Perhaps I did use such philosophy earlier, and also wrote it, but it is nonsense all the same.'*
>
> *… the very concept of observation was itself problematic. Every observation … presupposes that there is an unambiguous connection known to us, between the phenomenon to be observed, and the sensation which eventually penetrates into our consciousness. But we can only be sure of this connection if we already know the natural laws by which it is determined. If … these laws have to be called into question, then even the concept of 'observation' loses its clear meaning. In that case it is theory which first determines what can be observed.* "

The concept of indirect observation was initially a difficult subject to do with our ability (or inability) to interact with sensitive atomic structures without disturbing them: since the introduction of Hawking radiation in the 1970s it has had another application, in the subject of horizon behaviour. [i]

Within "core" special relativity, the difference in approach is not immediately obvious, since in empty flat spacetime there's no obvious way to "hide" information from view. However, when we try to embed special relativity in a *gravitational* physics (which includes curvature horizons), the difference becomes critical.

- Einstein's 1916 general theory inherits special relativity's shift relationships and philosophical approach, and as a result, a GR1916 gravitational horizon is **absolute**. Under GR1916, if we cannot directly "see" something, then it *does not exist* for us – it does not exist in our universe, and cannot – by definition – affect us either directly or indirectly. Under GR1916, once an object is behind a gravitational horizon, its subsequent events are forever inaccessible to the outside world – the horizon is an **event horizon**, and the idea of trying to escape or signal through the horizon surface is as futile as trying to travel faster than (or send a signal faster than) the speed of light under special relativity. What

i **Indirect observation and acoustic horizons:** If we stand at a location on Planet Earth, we may not be able to see a person if they are standing beyond the planetary horizon as calculated from our position. However if we ask someone else to stand half-way between us, and have that person hold up a mirror, we can see the first person *indirectly*. In the language of quantum mechanics, we can describe the first person as "***virtual***" – we cannot detect them directly, but we can infer their existence from indirect information available to us (for instance, we could ask the central person to look and tell us what they see, or they could aim a video camera at the first person, and we and we could aim a telescope at the device's screen).
This behaviour does not exist in SR-based physics, under which curvature horizons are *absolute* horizons.

can be observed by a remote observer decides the physics, which in turn decides what can be seen by all other external observers that exist in the first observer's universe:
"If I cannot see a thing, then nobody else I know (and can talk with) can see it either."

- A more QM-centric gravitational horizon allows objects that cannot be seen directly to still have effects that can allow their existence to be inferred indirectly. With an *acoustic* horizon, the horizon is a secondary, projective consequence of geometry, and an event occurring behind a surface that marks out a horizon for us, and whose signals cannot be seen by us *directly*, can nevertheless be intercepted by a nearer observer, who can then relay its information on to us.
The position of a horizon then depends on where the observer is and how they are moving: it is a relative, projected, limit – it is instead the theory which decides what can be observed.

30.7. Direct and indirect causality

While the concept of observation may sometimes seem a little on the "philosophical" side of natural philosophy, its inverse is the concept of *causality*.

- In an SR-based system, everything that exists in a given universe, for a given observer, is able to influence that observer directly. If we draw a causal network of objects and their interrelations, every object is directly connected to every other (subject to signal timelags), and the existence of this primary network lets us define a causal map, consisting of nothing but spacetime events and their separations, which can be considered to be the basis of Minkowski spacetime. The Minkowski geometry defines causality in an SR universe, and nothing the observer can do alters the underlying geometry of those relationships
- In a system based on NM and an acoustic metric, the causal network is more flexible and fluid. Gravitomagnetic curvature alters light-times between spacetime points, and if we define the separation between two points, and then realise that we have forgotten to take into account the motion of some intermediate object, the separation needs to be recalculated. These recalculations can even affect whether or not sections of two objects' timelines are said to be separated by a horizon, and the system supports horizon-spanning **causal chains** (for both gravitational and cosmological horizons), which generate QM-like behaviour.

The "philosophical" difference between special relativity (and GR1916) and quantum mechanics regarding indirect observation appears as a difference in the physical predictions of GR1916 and QM regarding trans-horizon physics, which is the basis of the black hole information paradox.

30.8. The correspondence principle

Niels Bohr (1885-1962) defined the correspondence principle in the ~1920s, [158] as being the idea that every independent property in quantum mechanics should (wherever possible) have a counterpart in classical physics, so that QM behaviour at small scales could, via statistics, yield classical theory at larger scales.

An example listing of the sort of larger-scale classical principles that quantum theory needed to support (Faye 2019 [159]) included: the principles of physical objects and their identities, the principle of separated properties, and the principles of value determinateness, causality, determination, continuity, and the conservation of energy.

Unfortunately, thanks to Eddington's 1928 misstatement about all classical theories being time-symmetric, [162] Bohr, Heisenberg, and the rest of the QM community were misinformed about *which* properties of classical theory a "nonrelativistic" (meaning "relativistic but Newtonian") quantum theory needed to reproduce, as traditional energy conservation is not a feature of a consistent implementation of NM – an NM system is "lossy" with respect to energy (section 45.5).

30.9. Energy-loss via gravitational waves, no isolated systems

Even under textbook theory there should have been clues that a system with moving parts should not be expected to preserve its initial energy, due to the unavoidable continuous generation of small gravitational waves. While this energy-loss may be considered insignificant in magnitude for most practical purposes, it establishes the principle that a complete physical theory has to describe a system as being lossy in practice with respect to energy.

Since gravitational waves radiate massenergy and momentum out of a system, [i] an NM-centric system of physics has no *truly* isolated systems. Given that space is disinclined to being bent (otherwise the universe would have been happy to collapse into a scrunched-up ball by now), we can expect a "free" gravitational wave (a freely-moving wave rather than a moving distortion anchored to matter) to show a tendency to straighten out as it propagates, so that the associated energy-loss is associated with an expansion of space in the surrounding region. We then have a causal relationship between atomic-scale energy losses and cosmological expansion, and between thermodynamic and cosmological arrows of time. [ii]

30.10. Updating quantum mechanics

> It is correct to insist on traditional energy-conservation as a law under QM, invoking the correspondence principle, if QM is to correspond to the physics of special relativity.
> It is not correct to use it in exercises where QM is supposed to "correspond" to NM.

In the application of the correspondence principle to "nonrelativistic" (*i.e.* "relativistic but Newtonian") QM, QM needs to be able to recreate NM's energy-loss behaviour, or else the quantum and classical predictions will not mesh. Once these energy-losses are "programmed into" QM, it will violate **T-symmetry** (will generate laws of physics that are different in forward and reversed time), and will allow the theory to generate a proper thermodynamic "arrow of time" that does not rely on interpretations of the wave function. This will also have implications for the broader subject of **CPT symmetry** (Charge, Parity, Time-reversal).

30.11. Summary

> Although we may not see an immediate incompatibility between SR and QM in the absence of gravity (or in the absence of particles), a general theory of relativity ***based*** on special relativity (Einstein's 1916 theory) is known to be fundamentally incompatible with QM.

> Eddington's reassurance that SR's lack of an arrow of time was general to *all* classical theories may also have caused quantum mechanics to include an incorrect assumption.

> Hawking's 2014 solution to the BHIP, "*gravitational collapse produces apparent horizons but no event horizons behind which information is lost*" [161] is valid … but non-SR and non-GR1916, as relativistic apparent horizons require the non-Einstein $(c\text{-}v)/c$ relationships. [23]

---

i In an SR-based system we can try to contain gravitational waves with a horizon. In an NM-based system, horizons are acoustic and *even this* does not work. Gravitational waves would seem to have to have "acoustic" characteristics, and therefore also imply an NM-based acoustic metric approach rather than SR's flat spacetime

ii This is not a quantitative calculation – however it establishes the expected existence of small-scale energy losses as a *principle*, and also explains how the energy leaves a system, and what happens to it.

31. SR Argument 31: "There is no alternative or competing theory to SR"

31.1. Relativity theory for particles with curvature

A reasonable response to the suggestion that the special theory may be the wrong theory of relativity is "well, what's the alternative?".

We are entitled to throw the same question back at the community.

It is a basic principle of theoretical physics that if one builds a theory on an idealisation, one is expected either to show that the idealisation does not alter the theory, or to explore the consequences of moving away from that idealisation. If special relativity's derivations depend on flat empty spacetime and the absence of gravitational fields, then it does not obviously apply to moving transparent media (like blocks of glass), or to bodies with measurable gravitational fields (like the Earth and the Sun).

But these bodies are not exempt from the principle of relativity.

What we need to know, and what we can *reasonably insist that the mainstream relativity community tell us*, is whether the equations derived for particulate matter or moving gravitational masses differ from those of SR, and if so, by how much.

- If the equations of motion derived for bodies with curvature turn out to agree with those for flat spacetime, then we will expect the SR community to champion the result, as a way of showing that the SR equations are *genuinely* solid, and can be safely applied within general relativity and other gravitational theory, for (*e.g.*) the calculation of gravitational shifts.
- If the equations of motion derived for bodies with curvature turn out NOT to agree with those for flat spacetime, then we have two different relativistic theories that can be (and need to be) compared.

If the alternative approach is *reasonably* good but somehow disappointing, we expect to see its failings usefully documented. If it is thoroughly *bad*, we expect to see the reasons *why* it is bad gleefully exposed by the SR community as a convincing reason to believe in SR.

What we do *not* expect is for the community to somehow have "forgotten" to carry out a basic sanity-check on the possible side-effects of SR's founding assumption of flatness, despite having had over a century to do get around to it.

> The absence of any obvious *published* relativistic competitor to special relativity's flat-spacetime approach does not mean that we know that SR has no logical competition: there logically *must exist* a parallel derivation of relativity theory within curved spacetime, for bodies whose curvature is not zero. It just isn't found in the mainstream literature.

The lack of a curved-spacetime derivation provided by the mainstream relativity community tells us that something is wrong: either the community is *not capable* of coming up with a curved spacetime derivation (in which case their expertise on such things is suspect), or any attempts at producing one are being abandoned every time the researchers find that the curved-spacetime exercise does not generate special relativity.

Either way, the implications for special relativity are not good.

31.2. Relativity in curved spacetime

Some of the previous sections have shown how we can produce a very simple [i] relativistic model based on the principle of relativity and the observation that moving bodies drag light. If a body moves with respect to its neighbours at v m/s, and this motion is associated with a gravitomagnetic field or distortion whose associated velocity-differential is also v m/s then we immediately have the basis of a theory that supports relativity, local lightspeed constancy, and also gravitation, and which, with the Carlip argument (section 23.3), also generates Newton’s First Law.

Because this approach associates the relative velocity of particles with curvature, its equations are automatically not the SR versions.

This is not a difficult idea, and if there was something obviously wrong with it, we would expect to be able to find peer-reviewed research telling us why it could not work. If the idea is being rejected without further analysis simply because it inevitably disagrees with SR, then we have an obvious competing system to SR, that has not yet been shown to be wrong, and which is absent from the literature simply because we aren't interested in anything that’s not SR-compliant.

31.3. Acoustic metrics (1997-)

As luck would have it, it would seem that, despite the determined efforts of some of the classical relativity community NOT to develop any form of potential relativistic competitor theory to SR, one has, regardless, eventually emerged in the shape of the mathematical theory of acoustic metrics.

The modern subject of acoustic metrics seems to have its origins in the **black hole information paradox** [149] that was identified in the 1970s.

Quantum mechanics insisted that horizons had to radiate (Hawking radiation), while SR-based theories such as GR1916 allowed one to prove that outward radiation through r=2M radiation was utterly impossible, which resulted in Hawking radiation initially being classified as a “non-classical” effect. Some researchers then noticed that some non-SR classical systems included a trans-horizon radiation effect whose results seemed “analogous” to Hawking radiation (Unruh, 1981 [182], 2016 [183]). In acoustics, we might naively expect that nothing can cross an acoustic horizon the wrong way, but it turns out that acoustic horizons leak information indirectly.

A mathematical description of the resulting acoustic metrics was then produced by **Matt Visser** in the late 1990s, [181] prompting a burst of research into the “acoustic analogues” of how QM effects played out against a classical curved background (“analogue gravity”). [184], [185]

Acoustic metrics and “analog gravity”

Our new abilities with respect to modelling acoustic metrics generated some excitement in the late 1990s and early 2000’s. We were keen to find a way for a hypothetical theory of quantum gravity to be able to “project” equivalent descriptions into the classical and quantum realms, and if the classical projection was to include Hawking radiation, then it seemed that we could explore at least *some* of the properties of a theory of quantum gravity by exploring the corresponding properties of an acoustic metric … with the advantage that at least with an acoustic metric, we know that it had at least some connection to real physics, and we could use our intuition to tell the difference between a genuine “crazy” outcome that was a valid prediction, and a false “crazy” outcome that was due to user-error.

i At least, *conceptually* simple.

31.4. Properties of acoustic metrics

Supersonic airflow and horizons

Suppose that we have a testing chamber for a jet engine whose exhaust jet is supersonic. We can argue that since the one-way velocity of the jet is greater than s, the speed of sound, a disturbance downstream of the jet cannot possibly be transmitted upstream, on principle. The exhaust jet can be considered to be intersected by a horizon, which allows signals to easily pass through in one direction, but not the other.

However the acoustic horizon is not like a GR1916 horizon: If we place an obstruction in the path of the jet, the resulting pressure increase pressure will "back up" along the jet until the presence of the obstruction is felt inside the engine itself. In the GR1916 description the horizon is "absolute" in a similar sense to the Minkowski spacetime relationships that generate it: in some respects it acts as a thing that acts but cannot be acted upon – (section 10.5), as far as the interior physics is concerned, an *absolute* horizon dictates that interior events can never affect the outside universe in any way, and interior physics cannot cause any variation in the shape of spacetime that can modify the horizon from the inside.

In the acoustic description, the *relative* horizon is more like a cosmological horizon, or the optical horizon of the Earth – it is a *projected*, effective boundary that is different for different observers, and is the result of the interplay between the physics and a specific observer's circumstances.

Causality

In SR, the absence of curvature means that the whole universe is visible, and we can base our theory on directly observable quantities. Since we declare that we know the exact behaviour of light everywhere (Minkowski's geometry), we extrapolate from what we see to what we know exists. Some of this approach carries over to GR1916, in which all variables are directly visible apart from those lost inside a black hole. Once a body passes through the horizon, it is permanently cut off from the outside world, the subsequent section of its worldline does not exist for us. Events in the subsequent section of worldline cannot in any way affect what we see (the GR1916 horizon is an **event horizon**).

Quantum mechanics and acoustic metrics introduce the concept of regions of spacetime that – analogous to the far side of the Moon, or regions of the Earth's surface that cannot be seen from a given location – are not *directly* accessible to the experience of a given observer, but whose existence can still be inferred, *indirectly*. With an **acoustic horizon**, an explorer just behind the horizon can choose to accelerate towards us, causing a gravitomagnetic field effect that increases the speed of light towards us in the region (Einstein 1921 [40]), and the region's modified curvature can then cause the horizon limit that we project onto the region to jump discontinuously from in front of the explorer to behind them. If we attempted to use this fluctuating horizon as a fixed reference, a piece of spacetime would seem to suddenly pop into existence in front of the horizon, apparently acausally, in response to events that we would say do not exist for us. We therefore get discontinuous geometrical *descriptions* (quantum geometry) as artefacts, even though the local physics is totally smooth and local causality is totally classical.

Thee is an obvious correpondence here between Einstein's definition of the difference in approach between SR and QM (section 30.6) – under SR-based theory, what we can observe defines the physics (hence absolute Wheeler horizons of GR1916), whereas under QM, it is the physics that defines what can be observed (relative, acoustic horizons).

Acoustic horizons are also not always complete closed surfaces. With a GR1916 horizon we say that the horizon is absolute, it is a boundary between two distinctly different regions, and that this boundary is always closed, with no gaps or "raw edges" (as Wheeler's mantra has it: "*a boundary has no boundary*" [48]). By contrast, acoustic horizons can be incomplete, with ragged or indistinct edges. If a speck of dust in our supersonic exhaust wants to communicate with the front of the jet engine, all it has to do is slip sideways out of the jet, and then meander though "normal" air back towards the front of the machinery (perhaps to be sucked in a second time).

Supersonic aircraft

If we inadvisedly chose to treat the speed of sound, "***s***", as a fundamental property, we could "prove" that an airplane with a flat tip to its nose could never exceed the speed of sound – in order for the 'plane to travel supersonically, it would have to be able to bat the air away from its blunt nose, but the deflection signal, moving at $s_{BACKGROUND}$, could not advance faster than the plane. For signals travelling forwards at ***s*** from the front of the plane, their wavelengths would compact towards zero as the aircraft reached the speed of sound, implying an infinite energy requirement.

In reality, a supersonic aircraft (or a supersonic thrown housebrick!) is not impossible, and does not break any real laws of physics, as the speed of sound is not a fixed, immutable property, and is affected by the local physics. The motion of the aircraft alters the air's properties (temperature, pressure, mean offset velocity, *etc.*), which in turn alters the velocity at which soundwaves propagate.

If we wanted to model this without referring to modified signal speeds, we could use an abstract QM-style statistical approach, with sound considered to be particulate (phonons), and say that the speed of sound represents a barrier that phonons "quantum tunnel" across to precondition the air ahead of the aircraft, and allow it to receive the 'plane. Or, we could suggest that no information actually moves forwards superluminally, but that the transitional region represents a horizon that cannot be crossed by anything that moves at less than the speed of sound, but that phonon-antiphonon pairs are generated ahead of the aircraft, with one member of the pair carrying information forwards, and its anti-phonon being absorbed by the aircraft.

31.5. Acoustic metrics outside of acoustics

The case for acoustic metrics

Although we refer to acoustic metrics as "acoustic", they do not require a conventional particulate medium, and can be described in terms of abstract classical fields, and/or curved geometry. We then have a classical field description, complete with a classical geometrical interpretation, that generates a counterpart of Hawking radiation.

This ability to model a Hawking radiation analogue in the classical domain was the obvious reason we were drawn to acoustic metrics – Hawking radiation was considered counter-intuitive and its existence seemed to contradict some of our most deeply held beliefs about gravitaitonal behaviour. If an analogue existed in classical physics, then by studying the analogue we could extend our general knowledge, and train our intuition so that these effects could be in our vocabulary of situations for which we felt we understood the difference between "right" and "wrong" physics. A relativistic acoustic metric could act as a "**toy model**" of quantum gravity, helping us to understand the rules and behaviours that might apply in a future theory of quantum gravity would need to obey, even if we didn't consider it to be literally correct physics.

A further minority position was that acoustic metrics *might* be able to be more than that – they might turn out to provide the geometrical basis of quantum gravity itself.

> Barcelo, Liberati and Visser (2005): [184] " *Secondary reasons* [for developing these analogies] *include the rather speculative suggestion that there may be more going on than just analogy – it is conceivable (though perhaps unlikely) that one or more of these analogue models could suggest a relatively simple and useful way of quantizing gravity that side-steps much of the technical machinery currently employed in such efforts.* "

Nonlinearity

The defining feature of an acoustic metric is not the existence of a particulate medium, but the existence of extreme nonlinearity in the behaviour of fields and the metric's associated geometry.

Normally in acoustics we would tend to describe the behaviour of signals in a region by simply superimposing or overlaying them on a common background. In an acoustic metric, signals interfere with each other, and also with themselves: if we have a stage speaker system designed for large concert venues producing a large-amplitude very low-frequency signal (say, a subsonic sub-bass signal used in some dance music), then the speaker cone will effectively be causing volumes of air to rock backwards and forwards, so that in one half of the wavecycle, there will be a net forward velocity in a region's air-molecules, while in the backcycle, there will be a net rearward velocity. This means that the forward speed of sound will be increased in the first half-cycle and reduced in the second, as if the low-frequency signal was a series of alternating gusts of wind, carrying other audio along with them. The speed of sound is partly a function of air density, so by creating regions where air density is different, the low-frequency signal creates a second effect that distorts the effective sound-metric in the region. If we then send a much higher-frequency lower-amplitude test signal through the same region, its propagation will be affected by the presence of the first signal, and may end up being frequency-modulated.

- **In a linear addition approach** to audio (which is usually adequate), we can calculate the expected sound-pressure at any moment and spatial position due to each of two signals individually, and when both signals are present at the same time, simply add the two results together.
- **In the more advanced nonlinear description**, the two signals alter each others' propagation behaviour, and even with a single signal, the presence of the signal (especially at high amplitudes) can invalidate the assumptions that we'd want to use for calculating how signals ought to move through the region.

In other words, sending a signal through a region to measure it's properties can end up altering the very properties that we were trying to evaluate – a very QM-style situation.

Gravity-waves are nonlinear

Some aspects of our audio problem may seem familiar. While the varying motion of gravitational sources, combined with a finite speed of gravitation, gives the *creation* of gravitational waves under almost any system, the question of how the waves *propagate* is more difficult, as (again) a gravitational wave is a signal that is expected to travel at the speed of light, but the signal itself represents a *modification* of the speed of light.

Are gravity-waves "acoustic"?

This might lead us to expect that the proper modelling of gravitational waves requires an acoustic metric. If the "fast" half of a gravitational wave-cycle is not allowed to progress through the background field any faster than background c_g (as the signal hasn't yet reached the region ahead of itself in order to modify its speed of light) then we expect severe waveshaping distortions and perhaps paradoxical energy-gains. On the other hand, if the fast half-cycle manages to propagate at *its own* speed of gravity, then we are back to the apparently-paradoxical case of the supersonic aircraft, that somehow manages to "communicate ahead" and precondition the medium to accept it. The propagation of gravitational waves may require a QM-style description, or a classical description that corresponds to QM … which would mean an acoustic metric.

Acoustic metrics also seem to have aesthetic "resonances" with the properties of a physics that includes gravitomagnetism. We have the same circular-looking associations … a moving mass cannot travel at more than the speed of light, *c*, but if the moving mass is associated with a gravitomagnetic field, the field modifies the local velocities of light in such a way that the various "*c*'s" in the region do not have the same values that they'd have had if the moving mass was not there. A **metrodynamic** theory appears to require an acoustic metric.

A potentially important aspect of "acoustic" wave behaviour is that it might have an influence on how easy and how likely it is that we can detect gravitational waves from very distant sources. Given that gravitational-wave detection is now an active experimental field, it would be odd not to bother doing more theoretical analysis of the expected propagation behaviour of the waves that experiments are trying to detect. [i] However, this might be considered politically delicate, in that, since acoustic metric properties don't make a match with special relativity, a proper "acoustic" analysis of gravitational wave behaviour might suggest that we were considering the possibility of SR being wrong.

31.6. Acoustic metrics *vs.* the Minkowski metric

Given that the Minkowski metric is the only relativistic geometry in which particles have no curvature, the only alternative relativistic approach would seem to be one in which particles *do* have curvature, giving us a relativistic acoustic metric. The late arrival of this subject (1990s) means that for most of the lifetime of SR and GR1916, the logical alternative that both theories needed to be tested against was not yet established, and the success of most SR-based research was conducted against a contextual "blank".

31.7. Dissolving gravitational singularities

If an observer dives into the acoustic metric counterpart of a black hole, the effective horizon retreats before them, and the faster and deeper they fall, the more the effective censoring horizon separating them from the assumed central singularity shrinks. As they see the horizon shrink, its Hawking radiation temperature increases, and the closer they get to r=0, the higher the ambient temperature becomes. If they hope to see the singularity at r=0 they will be disappointed – as they approach r=0, the Hawking radiation pressure increases towards infinity, and they see the shrinking hole radiating away all of its remaining energy in an explosion [20] while r is still fractionally larger than zero. Before they reach r=0, they will reach a height at which the total amount of massenergy corresponding to the Hawking radiation in all the layers above them will equal the total massenergy of the hole, and there will be nothing left to see at r=0.

Although this description is a simplification, because it ignores the gravitational effects of matter

i It would be especially awkward if we only managed to realise that a particular gw-detection configuration should *not* be able to register waves from distant sources, after a project had already been built and reported positive results.

above the observer's position, this matter would (if anything) only further weaken matter's expected attraction towards the centre, and the tendency to form a singularity.

Under Einstein's SR-based general theory, total collapse to a singularity is inevitable once a body has contracted to less than *r*=2M, while an "acoustic metric"-based general theory does not appear to allow singularities to form.

> Einstein's general relativity suffers from a known defect in that it is supposed to be a classical theory without infinities or geometrical discontinuities, and yet it generates illegal singularities when dense stars collapse.
>
> Since switching to an acoustic metric appears to eliminate the singularity problem, this defect appears to be the fault of GR1916's inclusion of special relativity.

31.8. Occam's razor

In the realm of special relativity, the Minkowski metric is obviously an awful lot simpler than the "acoustic" alternative, so if we had to make a choice between the two just for simple inertial motion, on current evidence, the Minkowski metric would be the winner. However, since acoustic metrics are "real physics" in other areas (*e.g.* Bose-Einstein condensate), the choice is not between a universe supporting SR *or* one supporting acoustic behaviour – it is between a universe that supports both acoustic *and* SR behaviour, or a universe in which only acoustic metric principles apply, everywhere.

Given that the principle of equivalence requires that particles have curvature, it would seem that the metric is required to be acoustic whenever matter is involved.

31.9. All roads lead to Rome

Acoustic metrics appear to be the end-point that we arrive at by trying to solve the shortcomings of a range of popular theories produced within the last two hundred years (the exception being Lorentzian electrodynamics).

Stephen Hawking suggested in 2014 that we might be able to solve the black hole information paradox by replacing GR's absolute horizons with relative horizons [161] – these are acoustic horizons, and in order to make this switch, we would have to change to an acoustic metric and the NM equations. Updating Fresnel's early C19th relativistic dragged-light model gives an acoustic metric. Forcing wave-compatibility onto Newtonian emission theory (section 14.5) generates an acoustic metric while preserving the dark star indirect radiation effect. Invoking the principle of equivalence to give particles curvature generates an acoustic metric. Treating cosmological horizons as leaky gives an acoustic cosmological horizon, intersecting an acoustic metric. The general principle of relativity requires gravitomagnetic behaviour, which generates an acoustic metric. The dragging effect of a receding black hole creates an additional pull on light, and a secondary horizon outside *r*=2m, which is an *acoustic* horizon, intersecting an acoustic metric. Working backwards from stochastic QM to derive a QM-compatible classical metric gives an acoustic metric. Assuming curvature-regulated local *c* for every mass generates an acoustic metric. Gravitational waves seem to require an acoustic metric.

Given all of this, we might wonder why it is that we have not already updated general relativity to use an acoustic metric, and announced quantum gravity to be a solved problem.

It is because of special relativity.

While the principle of relativity plus flat spacetime requires special relativity’s Doppler relationships, the principle of relativity plus Hawking radiation requires the Newtonian relationships. [23] This means that although we *can have* a Twenty-First Century theory of quantum gravity that unites QM and general relativity via an acoustic metric, the relationships of that theory would have to be those of Nineteenth-Century Newtonian theory. In a railway analogy, it is as if our current physical theory is a train on an “SR” branch line running parallel to the main express line … but in order to switch *onto* that parallel set of tracks, we have to put the train into reverse, back it up all the way to *circa* ~1900, change the rail points, and then proceed forwards again along the other set of tracks. We cannot go forwards without (briefly) going backwards.

To any physicists who have spent their careers arguing that special relativity is indisputable, that the experimental evidence is overwhelming, and that Einstein’s general theory has no problems, the idea of admitting that all of these things are wrong may be thought too high a price to pay for scientific advance. We have already waited half a century for someone to work out how to reconcile QM and GR *without* losing special relativity, and we would rather wait a little longer.

And since such a thing is geometrically impossible, we remain stalled in our current state, for the foreseeable future. The obstacles preventing us from having a theory of quantum gravity are not not technical, but social. [134]

31.10. Summary

The apparent absence of a well-known major competing system to SR/GR1916 should not be taken as suggesting that no such competing system exists, or that we do not know how to construct one – the logical alternative is a system built on a relativistic acoustic metric rather than the Minkowski metric.

In other words … we *do know* how to assemble a credible alternative system to compete with SR/GR1916 … we appear to know the specifications and have all the necessary tools … but we do not consider the idea to be “respectable physics” since it doesn’t agree with SR.

> The absence of a studied alternative to SR/GR1916 is not because such a thing is not available if we want it … it is because we are not willing to investigate any solution to gravitational theory or quantum gravity that does not incorporate special relativity.

32. SR Argument 32: "SR is unavoidable" / "Relativistic alternatives to SR are impossible"

32.1. Defining SR-alternatives out of existence

The easiest way to eliminate potential competitor theories to SR is to convince ourselves that no alternatives are possible. [i] This belief can be strengthened by repetition, by claiming "ownership" for SR or GR1916 of results that we know are right (so that other theories are not allowed those results, making them wrong), and by tailoring and redefining the definitions of common words and technical terms to reflect how they behave under the current system, which can then make explaining the workings of these principles under alternative systems impossible. [ii] To some people who think linguistically (which may include any mathematicians who place too much emphasis on notation), a corruption of definitions can result in an inability to think. [iii]

- "**Relativity means SR-based**" – it has become common practice for researchers to follow Einstein in using the word "relativistic" to mean "something based on special relativity". This creates a fog of confusion over attempts to produce non-SR relativistic theories, as one could be forgiven for thinking that no such thing can exist by definition. It also creates confusion over dependencies between theories: "The relativistic aberration formula" is common to all relativistic theories but is assumed by some people to be exclusive to SR, while "the relativistic Doppler equation", which *is* exclusive to SR and "SR-alikes", risks being wrongly considered as the *only* relativistic Doppler equation, preventing us from considering the possibility of others. [iv]

- **Metric theories** – one might reasonably expect the literal definition of a "metric theory" to be "a theory that has or relies on a metric". GR texts can add a further condition: a metric theory must also reduce the physics of special relativity.

> - MTW, [53] page 1067: "(1) **Spacetime possesses a metric; and** (2) **That metric satisfies the equivalence principle** (the standard special relativistic laws of physics are valid in each local Lorentz frame). **Theories of gravity that incorporate these two principles are called metric theories.** "

---

i This is efficient, work-wise, as it avoids having to do the actual research.

ii "The art of the proof is knowing how to hide your propositions." Einstein was especially good at creating narratives where previous theory was supposed to predict *X*, a result that was then shown to be untenable, at which point the new theory would rescue the situation by explaining that the correct outcome was instead *Y*. As we've seen, it was often the case that previous theories predicted *Y* as well.

iii This idea was explored in George Orwell's "1984", [135] in which the fictitious "Ministry of Truth" uses redefinitions of words to make some undesirable concepts literally "unthinkable": "*The purpose of Newspeak was not only to provide a medium of expression for the world-view and mental habits proper to the devotees of Ingsoc, but to make all other modes of thought impossible. It was intended that when Newspeak had been adopted once and for all and Oldspeak forgotten, a heretical thought — that is, a thought diverging from the principles of Ingsoc — should be literally unthinkable, at least so far as thought is dependent on words. Its vocabulary was so constructed as to give exact and often very subtle expression to every meaning that a Party member could properly wish to express, while excluding all other meanings and also the possibility of arriving at them by indirect methods.*"

iv One of the most untrustworthy words in theoretical physics is "the" – it creates an implicit understanding that there is only *one* of something, without the reader always being aware that they have just accepted a potentially dubious premise. If we start with the SR Doppler formula, pretty much everything else in the special theory follows. So if we were to accept this equation as being "*the* relativistic Doppler formula" (the terminology used by Einstein), we would be unable to produce a relativistic system that was not equivalent to SR. If we say "the only possible relativistic Doppler equation is the one used by SR", then this invites an exploration of whether the statement is true. If we make the statement *implicit*, by starting a sentence "According to the relativistic Doppler formula …", the assumption is much more likely to escape analysis.

If we wish to study *non-SR* metric theories as potential competitors to general relativity (such as theories based on *acoustic* metrics), MTW's default definitions make this impossible – the subject is already defined out of existence.
It is difficult not to suspect a form of foul play here, as, if researchers *genuinely* believed that there were no other conceivable forms of metric theory, the second part of the definition would be redundant. Condition **(2)** only seems to have a purpose if we suspect that other possibilities *may* be available, but we wish not to be bothered by them.

- **Gravitational theory** – Since SR legitimately "owns" relativity in flat spacetime, any alternative classical relativistic system to SR needs to be a curved spacetime theory, and therefore "gravitational". Perhaps we can develop our non-SR metric theory without explicitly *calling* it a metric theory, by basing it only on general gravitational principles and only using field-theory language? Here, again, we are confounded by the textbooks. Will's list of "*Basic criteria for the viability of a gravitation theory*" (Theory and Experiment in Gravitational Physics (1993) [136] 2.2) declares the third condition of viability to be that "***(iii) It must be relativistic**, i.e., in the limit as gravity is "turned off" compared to other physical interactions, the non gravitational laws of physics must reduce to the laws of special relativity.*" In other words, even if we successfully manage to construct a viable relativistic theory of gravity that does not rely on SR, it risks officially being classified as both non-viable and non-relativistic, "by definition". [i]

- **Redefining the Principle of Equivalence** – The principle of equivalence of inertial and gravitational mass is one of the cornerstones of theoretical physics, and any theory that violates the principle is liable to be seen as an automatic failure. After the 1960 crisis, we began to appreciate that the principle of equivalence was incompatible with the SR approach of treating inertial physics as a flat-spacetime problem. Special relativity and the PoE were not logically capable of coexisting as exact entities within the same structure (Schild 1960 [46]) – theories based on special relativity violate the principle of equivalence. How can we avoid admitting this awkward failure to students? In true Orwellian style, if the theory fails to obey the principle, we can redefine the principle to fit the theory. We can define a new replacement principle whose definitions do not mention the troublesome concepts of inertial mass, gravitational mass or equivalence, tailor it to the behaviour of the current theory, and teach students that *this* is to be accepted as the "modern" implementation of the principle of equivalence – the "**Einstein Equivalence Principle**", or "**EEP**".

> MTW [53]: "*Of all the principles at work in gravitation, none is more central than the equivalence principle. As enunciated in §.16.2, it states:* ***'In any and every local Lorentz frame, anywhere and any time in the universe, all the (nongravitational) laws of physics must take on their familiar special-relativistic forms.'*** "

In other words, when our theory fails badly, by refusing to conform to one of the most fundamental principles in physics, we respond by taking the unsuccessful way that our theory *tried* to implement the principle, and defining that failed implementation *as being the principle itself.* [ii]

i The very recent 2018 edition of "Theory and Experiment in Gravitational Physics" [137] softens this position a little, to "*… laws of physics must reduce to the laws of special relativity, either perfectly or to a high degree of precision.*" (2.1: "The Dicke Framework"). Since we've known since 1960 that a full general theory *cannot on principle* reduce exactly to SR physics, the belated concession that reduction needn't be absolutely *perfect* is rather late.

ii Since the PoE makes SR invalid, and the EEP makes SR compulsory, the *Einstein* equivalence principle violates the (traditional) principle of equivalence. This situation allows statements such as "*the modern principle of equivalence*

This not only gives the appearance that the theory is compliant (without actually fixing the problem), it also means that if further problems are found that again lead to the rejection of special relativity, these can now be dismissed as invalid, for violating "the (modern) principle of equivalence". [i]

- **The PPN System** – Shortly after the appearance of Einstein's general theory and a cluster of broadly similar other geometrical theories of gravitation, it became clear that it would be useful to have some sort of objective system for labelling and cataloguing theories according to their predictions (a test theory, or "theory of theories"). The two main systems that emerged were the **Dicke** framework, and the **Parameterised Post Newtonian** ("**PPN**") system, which was developed by a series of researchers from the 1920s to the present day. Can the PPN system be used to fairly assess a non-SR theory? According to MTW, no ... because it assumes that all theories to be assessed are metric theories, and its definition of a metric theory requires a reduction to special relativity.

The promotion of the SR-specific EEP as being the improved, "modern" version of the equivalence principle is particularly invidious as it not only supports failure, but makes failure compulsory. [ii]

If all "competitor theories" to Einstein's general theory are required to obey the EEP and reduce to SR physics, then they will all fail in exactly the same way as the 1916 theory. Of the range of gravitational theories currently "on the books" that pass the test of EEP-compliance, every one that has a natural, *unforced* agreement with the EEP is pretty much guaranteed to be a structural failure and either explicitly incompatible with the general principle of relativity, or pathological. Of all the theories that fail in the same way as GR1916, GR1916 will be the most efficient, and will therefore the "best" theory available. [iii] If a genuinely better theory comes along, that is *genuinely* compliant with the GPoR, it can be discarded for not being compliant with the EEP.

Similarly with quantum gravity. Our search for a way of uniting classical gravity and and quantum physics to date has been a failure, not because there is not a fairly obvious way ahead, but because that way necessarily conflicts with special relativity, and therefore violates the assumed conditions for a credible theory.

32.2. Summary

> Where we have found that SR is incompatible with a basic physics principle (such as the principle of equivalence), we have tended to quietly redefine the principle to agree with the theory. The number of incongruous definitions making SR compulsory suggests that the apparent success of the SR-based approach may have been achieved artificially.

---

violates the principle of equivalence". This is the consequence of a defensive "political" position, that we adopt in order to avoid admitting that a major theory has failed one of its founding principles. [138]

i One can compare the rewriting of "the equivalence principle" to the proclamation of principle in George Orwell's earlier book, *Animal Farm* (1945), "*All animals are equal*", which one night becomes mysteriously amended to "*All animals are equal, but some animals are more equal than others.*" The case of replacing the original equivalence principle with the EEP is more extreme than this, in the context of the Orwell book it would be more similar to rewriting the principle of equality by removing the words "equal" and "animals", so that the modernised "equality principle" reads, "*The pigs are necessarily always in charge of everything, everywhere*".

ii Leveraging control over industry standards and definitions to make competing products "noncompliant" is a well-known form of corporate misbehaviour. [139]

iii The institution of the EEP is a little like the owner of a racehorse horse realising that the animal that they have been paying to have maintained and trained only has three legs, but that, as the owner of the racecourse, they have the ability to set the rules of the race, and then exercising this ability to declare that in order to be eligible to take part in the race, a horse is not allowed to have *any more* than three legs.

33. SR Argument 33: "We can't just update Newtonian theory because we feel like it"

33.1. What we mean by "Newtonian"

Some of the behaviour described here as "Newtonian" does not correspond to textbook Newtonian physics, and it is reasonable to object that one should not go around changing the meanings of understood words and phrases on a whim. However, if we have two possible competing systems of relativistic physics, one based on the Lorentz-Einstein relationships for energy and momentum, and one based on the redder relationships required for Newtonian theory, then it seems natural to refer to the the second set as "Newtonian". This seems more exact than the relativity community's habit of comparing SR to the predictions of "classical theory", which usually seems to mean an incompatible hand-picked selection of the worst aspect of pre-SR theory taken from either Newtonian or aether theory arguments, apparently selected to make special relaitvity look as good as possible.

Some may find it more convenient to refer to the arguments given here as "updated Newtonian", to refer to the C19th arguments as "low-velocity Newtonian" and the C21st versions as "high-velocity Newtonian", or to use some other terminology.

33.2. Newtonian theory has already been changed

What we think of as Newtonian theory has already undergone at least one major iterative change. The system that Newton described in Principia and Opticks is not the same system taught since around 1800, prior to ~1800 Newtonian theory used an inverted relationship between energy and wavelength, and between lightspeeds and light-deflection, so technically, Newton's system (as published) was invalidated back in the early C19th. What was taught in the later C19th as "Newtonian" was a cut-down (and corrected) version, with the Doppler relationships that the theory *should* have had. Since the project of updating Newtonian theory was never completed (acoustic metrics not becoming mainstream until the end of the Twentieth Century), C19th Newtonian theory remained an unfinished project with respect to the behaviour of light.

33.3. Some updates to Newtonian theory are obvious

With the benefit of hindsight, we can see that the "SR" $E=mc^2$ result is also an exact consequence of the NM energy and momentum laws. We can also use C19th Newtonian theory to quantify the frequency and wavelength predictions of Michell's 1783 predictions for gravitational shifts in energy to get predictions that nobody would seriously object to being called "Newtonian". As well as the established Newtonian gravitational shift relationship, we also already have the established Newtonian counterpart of a black hole, known as a "dark star". Einstein's 1911 paper rederived the Newtonian gravitational shift and argued that the consequence was time dilation, so (thanks to Einstein) we can also talk about "Newtonian" gravitational time dilation. Since the 1783 dark star model was sufficient to describe lightcorpuscle leakage through a horizon along accelerated paths due to momentum exchange, we also arguably have "Newtonian Hawking radiation", although the phrase "classical Hawking radiation" might be sufficient.

Since the time dilation effect arguably changes the Newtonian predictions for lightbending, we should probably update these too, but we will need to be much more careful about how we refer to the updated predictions, because we do not want to make existing texts more difficult to read that refer to the "Newtonian *vs.* Einsteinian" lightbending predictions. "Updated" Newtonian light-bending predictions should probably be presented with some sort of qualifier.

33.4. Some existing textbook arguments are wrong

In some cases, it is legitimate to override accepted textbook definitions of Newtonian theory and of what it predicted, because they are simply not correct, or seem to represent an "artificial" divergence between Einstein's arguments and those of earlier theory.

In the case of transverse redshifts (the shifts that we see when we aim a detector directly across the path of a moving object, at what *we* see to be 90 degrees), a range of reference sources define this effects as being unique to special relativity, whereas practically any C19th theory will predict some sort of redshift in this situation. Rather than accept the historical (but mathematically and geometrically wrong!) characterisation of how SR compares to earlier theory, it is legitimate to *correct* the existing narratives. As scientists we are not obliged to perpetuate bad math and bad geometry just to spare the blushes of earlier researchers. If a paper *defines* pre-Einstein theory as not predicting transverse redshifts or gravitational redshifts, we are entitled to disagree.

We are not obliged to cooperate with misleading comparisons designed to make a current theory look good. If a comparison would be considered too misleading to be used legally in advertising for consumer electronics or motor cars, then it should not be used in science. Our standards ought to be higher than those of wider human society, not lower.

33.5. Grey areas

The difficulty in deciding how far we should "correct" older theory is similar to the problem of how far we should go in "restoring" old paintings. When the author of a theory, and the wider community accidentally "gets the theory wrong", then should the theory to be defined by what it is *believed* to say, or by what it logically *ought* to say? In the case of special relativity, we have no problem in saying that, regardless of how many sources declared that SR predicted approaching objects as being seen to be contracted (Gamow, *et.al.*), the "true" theory was different to its published characterisation. The predictions of SR and Minkowski spacetime for simple situations are mathematically unambiguous, and if books (and perhaps even Einstein) got some of these predictions wrong, then we can use correct math to dismiss the "historical" view as faulty.

In the case of velocity-addition, where modern textbooks define the Newtonian relationship as $v_3=v_1+v_2$, then this might well be how Newtonian calculations were done in the C19th by default, but were C19th students ever actually *taught* that this relationship was still to be considered exact for relativistic velocities? Or would it just have been generally assumed, until someone actually studied the problem? We can see in section 18.3, that if we calculate the "composite" results of the Newtonian Doppler relationships, the result disagrees with the textbook version. It seems fair to refer to the correct *derived* result of the Newtonian relationships as being "Newtonian", rather than deferring to retrospective definitions in textbooks that may have been partly motivated by a desire to show som other theory in the best possible light.

The situation with retrospective redefinitions with regard to Einstein's general theory is more difficult. In this case, the specifications for the theory (general principle of relativity, principle of equivalence, and SR physics as a limit) are now known to be unachievable in a single theory. While Einstein had specifically defined GR1916 as a principle theory, the community responded to the 1960 breakdown by treating Einstein's *definitions* as user-error, and re-cast the mathematical structure as being the result of a non-principle theory that was allowed to violate the GPoR and the PoE. This does not appear to be an honest move to correct a misunderstanding – it seems to be a defensive reaction to avoid admitting publicly that Einstein's theory had failed. Correcting the Gamow mistake (or changing Einstein's 1905 use of the phrase "velocity of light" to "speed of light") is a legitimate return to what the math really says, but making major *logical-structural* changes to a theory (without announcing a breakdown) is more dubious.

We can now technically distinguish between at least three different meanings for "general relativity": "**GR1916**" – the 1916 theory (principle-based, GPoR, SR, logically inconsistent), "**GR1960**" – the 1960 "reboot" (constructive theory inheriting a body of math from GR1916, SR has priority, not actually a general theory of relativity), and an ***actual*** general theory of relativity, which *cannot* include SR as a component, but which does not seem to be explored in GR textbooks or the GR literature, due to our insistence that gravitational theories be SR-compliant.

Standard definitions that make SR-compliance compulsory, mean that we cannot develop a working general theory of relativity, and (since SR is the reason for the black hole information paradox) also cannot produce a theory of quantum gravity. If strict textbook definitions are preventing us from being able to do science, then we should be prepared to ignore them.

33.6. Summary

> When this paper talks about "Newtonian" results, it is referring to results derived from Newtonian energy/momentum relationships, and is not pretending that all these results were actually known to Newton or to C19th physicists.

If our goal is to try to discover the fundamental nature of physical reality (for instance, whether motion shifts really obey the SR or C19th Newtonian relationships), then we have to be guided by mathematics, logic and geometry. Where standard historical definitions violate these disciplines, then we may acknowledge these definitions as historically interesting, but we cannot allow them to dictate to us what is to be considered logically possible, or the form of the debate.

A framework using the C19th Newtonian relationships is the logical competitor theory to both Einstein's special and general theories. Within the realm of purely inertial/gravitational physics, it appears to be the *only possible* logical competitor theory that still supports the principle of relativity.

While it is understandable that some may object to using the term "Newtonian" to refer to aspects of such a thing, we can ask, "Well, what then is the existing *correct* term for this class of theory?" If the class has no name, and the use of standard definitions has led to our failing to *even realise that this region of theory-space exists*, then the current scheme is not working.

34.SR Argument 34: "Frame-based arguments prove SR"

34.1. Frame logic

We can simplify descriptions of the interactions between *arrays* of simply-moving systems with relative motion by discarding the location information of the individual bodies in each array, and replacing each array with a "**frame**" that represents all possible objects that have a particular relative direction and speed. Then instead of saying "motion relative to the galaxy", or "motion relative to the Earth", we can say, "motion relative to the galaxy's frame", or "motion relative to the Earth's frame".

By assuming that the speed of light appears to be globally fixed across each valid inertial frame, we can applying the principle of relativity to the relationships between frames, and prove that these relationships are those of special relativity.

However, the "frame" abstraction eliminates some of the potentially interesting properties of particles from the description. If we smear a particle and its fields out over an arbitrarily-large region, without proximity-dependence, to produce a "flat", field-free frame, we have effectively deleted the particle's fields by deleting the particle's location (or deleting the particle itself). We have changed the physics.

34.2. Flat spacetime by the back door

While frame-frame interrelationships are excellent geometry, they not automatically physics – we have never seen an inertial frame, or accidentally stubbed our toe on one. [i]

The "frame" arguments succeed in deriving special relativity by abstracting away as unimportant the locations (and therefore the field effects) of particles, simplifying the exercise and generating SR … but the initial decision *to treat field effects as unimportant to the physics* is *a decision and not a derivation*. The approach simplifies by eliminating physical properties from the description that would otherwise have lead to non-SR physics, and leaves us with just one possible solution. But geometry does not tell us whether this simplification is "proper", and it certainly does not tell us that this choice is free from long-term consequences. [ii]

If we want to believe that the inertial frame arguments generate real physics, we have to believe either that particle curvature does not exist, or that it plays no part in how particles interact. Since the "frame" approach implicitly applies the same assumptions as special relativity, we should not be surprised that they generate the same outcome … the approach is *a valid geometrical derivation* of special relativity, but should not be taken as an *independent* supporting verification of SR, as it is essentially the same argument.

> While "inertial frame" arguments do lead to special relativity, this is because flat spacetime plus the principle of relativity will always lead to SR. However, if we introduce particles with curvature-fields, since the result of particles-with-curvature must then be a *different* geometry, our proof of SR turns into a disproof.

---

i Eddington (1927) [{162}] "*Nature … is not enthusiastic about frames of space. They are a method of partition which we have found useful for reckoning, but they play no part in the architecture of the universe. … She herself has paid no attention to them ...* "

ii Idealising the Earth as a gravitational point-mass is useful for calculating orbits but not for calculating landings. Assuming the absence of atmosphere simplifies some problems but is not helpful when designing aeroplanes. Assuming the absence of fields (or assuming that fields are irrelevant) makes the result incompatible with a larger gravitational theory that requires inertia to be associated with gravitation.

35. SR Argument 35: "Coordinate-system logic generates SR physics"

35.1. SR coordinate systems

Coordinate-system arguments generate special relativity *if they are based on the same assumptions as special relativity*. Some aspects of SR coordinate systems are also seem to be more about imposing human values onto physics, than finding physical law.

For instance, special relativity allows us to assign distance and time coordinates to distant events, in order to be able to label those events and decide whether two distant events in different directions are really simultaneous or not, by deciding that we know that the speed of light is fixed with respect to us, over astronomical distances.

There is no obvious application of this to fundamental physics. An atom, receiving two signals at the same moment does not care about how far away they were generated, or whether they were nominally created at the same time or different times. It makes no difference to the atom. Which is just as well, because our distance and time labels for those events are different under SR for another atom passing by.

35.2. SR coordinate observer-dependence

In addition to this, SR coordinate systems break down over interstellar distances if the observer has even a very mild acceleration.

If we have a particular inertial state of motion, we can assign SR coordinates to the creation-events of the light that we are now seeing, based on the idea that the light has all crossed interstellar space at a velocity of c_{US}. We may decide that the light we're seeing from two stars $\mathbf{S_1}$ and $\mathbf{S_2}$, seen in opposite directions, has taken the same time to reach us, and that the two origin-events are therefore simultaneous.

If we now give ourselves a velocity towards $\mathbf{S_1}$ and repeat the exercise, we will calculate that the approaching $\mathbf{S_1}$ was further away when the light was generated, and therefore must have been generated earlier than in the first exercise. Similarly, light from receding $\mathbf{S_2}$ must originally have been closer to us in the past, and therefore in order for its signals to be reaching us now, the $\mathbf{S_2}$ origin events must have been generated later. By changing velocity, the set of events in spacetime that we declare are happening "now" – our **plane of simultaneity** – tilts, aligning differently with the space and time coordinates that we used before the acceleration. While we accelerate, the plane advances along our worldline, and also continually changes angle, in such a way that the plane *pivots* around a point in space somewhere ahead of us (Baird, 2007 [5] fig.15.2, p.202).

35.3. Coordinate breakdowns

For worldlines between us and the pivot-point, assigned time coordinates advance more slowly during our acceleration, and for an event *at* the pivot-point, we are continuously assigning different values to exactly the same event, for as long as we accelerate. For worldlines further away than the pivot-point, the plane of simultaneity sweeps *backward* when we accelerate, with the result that a sequence of events in the distant worldline gets "covered" three times, once forwards, once backwards (during our acceleration), and then (when the acceleration has finished) forwards again.

35.4. The SR breakdown distance

When we calculate the breakdown distance for a given rate of change of velocity, the "pivot" distance turns out to be,

BreakdownDistance (lightyears) = 1/acceleration (gees)

By a rather improbable coincidence of units, the distance measured in lightyears gives the right answer if the acceleration is measured in Earth-gravities. If a region is any more than one lightyear away, it's assigned SR time coordinates run backwards and become nonsensical if we accelerate it toward it at any more than about ten m/s^2 (one "gee"). If we accelerate towards it at *ten* gees, the breakdown limit is about one tenth of a lightyear. Since the nearest star, Sirius, is over four lightyears away, an astronomer can scramble the star's SR-generated coordinates just by getting up out of their observation chair too quickly.

35.5. Responses

There are four main responses to the breakdown:

1. **This is obviously wrong**, if SR coordinates really failed like this, we'd all know about it.
2. **Special relativity is only a local theory.** By trying to apply it, with acceleration, over interstellar distances, we have used it outside its domain of applicability.
3. **Under Einstein's general theory, physical acceleration is associated with spacetime curvature**. To model these problems properly requires "full GR".
4. **Special relativity's assigned coordinates are purely a matter of cartographic convenience** (as we should be able to tell by the fact that they are different for differently-moving people).

To the first response we can point out that the breakdown ***is*** documented (complete with "lightyear" and "Earth-gravity" units) in Misner, Thorne, and Wheeler's "**Gravitation**" [53] (on page 165). It is a genuine, known breakdown. The fact that it's not better known suggests that some SR textbook authors might never have tried the problem, or prefer not to mention it in SR textbooks because it spoils the narrative.

To the second response, that SR is only a local theory, we can point out that SR was *derived* as a global theory, making use of global lightspeeds – Einstein didn't specify in 1905 that the theory was to only apply over small regions. SR textbooks are littered with examples of the "Twin Astronaut" problem, with astronauts travelling several lightyears to a nearby star, turning around *very* quickly, and returning to be found to have aged less. If special relativity is not to be considered valid in these situations, then we should stop using these examples to teach it.

The third response, about complicating acceleration-related curvature effects is correct, but also destroys the SR clock hypothesis (section 7.6), which is necessary to save SR from being invalidated, Without the SR clock hypothesis, SR probably has to be considered to be "wrong physics" (making any issues with SR coordinate systems somewhat moot).

The fourth answer, that we shouldn't be putting so much stock in SR coordinate systems is probably the best answer. [i] [ii]

---

i (Einstein, "Relativity…", §8, "*That light requires the same time to traverse the path A→M as for the path B→M is in reality neither a **supposition nor a hypothesis**, about the physical nature of light, but a **stipulation** that I make of my own freewill in order to arrive at a definition of simultaneity*").

ii Rindler [34] " *Even in GR, the clock hypothesis is but a hypothesis* ".

36. SR Argument 36: “Special relativity can coexist with gravitational theory”

36.1. Compatibility

Although many physicists would probably agree that special relativity “doesn’t do gravity” (because “that’s general relativity’s job”), they will tend to believe that SR can at least *coexist* with gravitational theory. If it can’t, then since our universe includes gravity, it can’t be physics.

36.2. Behaviour at the strong-gravity limit

For the case of a rotating black hole, the dragging effect of matter on light is supposed to be “total” at the horizon, [i] and when a non-rotating black hole moves away from us, the GR1916 condition that outward-aimed light emitted at the horizon remains trapped in the horizon requires that when the hole recedes at v m/s, light “frozen” into the nearest part of the horizon also recedes at v m/s (if the light receded any slower than the horizon, the light would be exposed and would be able to escape).

When a patch of horizon moves, because its owning mass is either moving or revolving, it is expected to drag light completely.

If we now take a pair of black holes whose relative velocity is v, and send a signal between the two horizons, [ii] the light will start its journey with a velocity-dependent offset of c_{EMITTER}, will ride a gravitomagnetic gradient equal to the relative velocity of the two holes, and will then arrive at its destination with a velocity-dependent offset of c_{RECEIVER}.

36.3. Numerical results (for complete dragging)

The default gravitomagnetic shift due to gravitational dragging would then be expected to change the light-energy of signals by the same ratio as the gravitomagnetic change in the speed of light, $\pm v/c$, giving a default Doppler-like gravitomagnetic recession shift of $E'/E=(c-v)/c$ for recession.

By assuming total lightdragging at a moving surface, we end up (perhaps to our surprise) back with the Newtonian Doppler relationships, even if we don’t particularly regard C19th Newtonian theory as being either valid or credible.

36.4. Doppler shift “extinction”

For the extremal case of a gravitational mass with a horizon, this gravitomagnetic shift appears to mimic a (non-SR) Doppler shift in sign and magnitude. If we try to model this shift in an SR-based theory, as a separate gravitomagnetic effect to be superimposed on top of the “normal” SR Doppler motion-shift prediction, then “simple” recession/approach optical Doppler shifts would be about twice as strong as we currently assume. Being “off” by a factor of two for simple colinear Doppler effects would be a difficult thing to miss.

The alternative is to suggest that, since the signal has already changed its velocity by v to match the relative velocity of its target by the time that it arrives, there is no longer any justification for there being a conventional propagation-based Doppler effect – the traditional Doppler shift effect has effectively already been ***erased*** or ***extinguished***, and the gravitomagnetic shift effect has replaced it.

---

i For a diagrams of dragging around a rotating black hole, see Thorne 1994, [22] figs 7.7 and 7.8, pages 291 and 292. For the resulting tilting of Minkowski lightcones around a rotating black hole, *see* MTW [53] Box 33.2, page 881).

ii (or perhaps a vanishingly-small distance above the surface of the “emitting” hole)

A full-blown general theory of relativity therefore does not even use the flat-spacetime *concept* of Doppler shifts as it exists under SR (except as a rough approximation).

36.5. Universality of the Doppler equations

In the extremal strong-gravity case, we have local lightspeed constancy everywhere, and the (c-v)/c relationships. Can we smoothly transition from SR for weak-gravity physics to this result for strong-gravity physics, depending on the gravitational properties of the mass being examined?

No ... because the principle of relativity and the principles of wave-compatibility and metric-compatibility all require that all objects in the universe obey identical Doppler relationships. If a moving black hole exchanges signals with an atom, and the moving black hole's Doppler shift relationship is $(c\text{-}v)/c$, then if the hole is said to be stationary and the electron is moving, we need the Doppler relationship to *still* be $(c\text{-}v)/c$, otherwise we could tell who was "really" moving. Once we have worked out that the black hole case needs the $(c\text{-}v)/c$ relationships, every massed particle in the universe also has to obey $(c\text{-}v)/c$, and has show total lightspeed dragging (at a sufficiently small radius), and act as if it is horizon-bounded.

> Not only is there *no such thing as non-gravitational physics* (sections 11, 42.2), there is also *no such thing as weak-gravity physics*, or medium-gravity physics. In a relativistic universe with gravity, the extremal strong-gravity solution must act everywhere, and the SR equations must not just be *not quite exact*, they must "miss" the correct target equations by an entire Lorentz factor.

Once we have realised this, we are on our way to creating an alternative general theory, based on proper general relativistic principles, that meshes properly with quantum mechanics.

We can now see why the velocity-dependent gravitomagnetic effect gets sidelined and ignored wherever possible in general relativity textbooks: Although it has to exist (see also the Carlip argument, section 23.3 [130]), it invalidates, replaces and supersedes both special relativity, and Einstein's 1916 general theory.

36.6. Summary

The system of relativistic physics that we get by assuming gravity, and the system that we get by assuming the *absence* of gravity, do not have a superset/subset relationship. These are two separate antagonistically-incompatible and immiscible systems with discretely different laws and different fundamental relationships and behaviours, occupying different logical universes. We cannot "mix and match" them.

> If we have special relativity, we cannot have a working general theory of relativity.
>
> If we want a working general theory, we cannot have special relativity.

37. SR Argument 37: "Special relativity deals with weak gravity, GR deals with strong gravity"

37.1. SR/GR compartmentalisation

According to this argument, special relativity applies where the effects of gravity are sufficiently weak, while GR1916 applies elsewhere.

If we subscribe to the "compartmentalist" approach, then special relativity's flat-spacetime model is never *wrong* (as such), merely sometimes *inappropriate*. We use it whenever we can get away with it, but when the problems become too much, we switch to general relativity. In the case of particle curvature, we know that SR *must* be wrong, but we treat this as a technicality – the fields of particles must surely be to absurdly small to bother with.

When we find that this is not a defensible position (section 40), we pass the buck on to general relativity, saying that if particles *do* have curvature, they must be modelled using our theory of relativity in curved spacetime – GR. When it then turns out that GR1916 *also* can't cope with particles with curvature (because it uses the flat SR shift equations, and because GR doesn't agree with QM), we pass the buck again, and say, well this is then a problem that requires quantum gravity.

Since we *do not have* a theory of quantum gravity, [150] this is effectively a way of saying that we accept that a problem is insoluble under current theories, but still do not accept that anything might be wrong.

37.2. Attempted parameterisation

If the Doppler equations for special relativity can't apply to moving black holes (section 29.1), then surely they are still essentially correct for cases where moving bodies have insignificant gravitation? Perhaps we can continue making a distinction between the realm in which SR is good enough to apply, and the realm where gravitational effects become too large to ignore? Now that we know the exact degree to which relativity diverges from SR for "extremal" gravitational bodies (one complete additional Lorentz redshift), we might be able to parameterise, and write something like,

$$E'/E = [SR] \times (1- v^2/c^2)^{\mathrm{grav}/2}$$

, where [SR] is the SR Doppler prediction, and "grav" is a measure of the gravitational field differential to the surface of a mass when the mass is not moving, with values ranging from "zero" (in which case the equation gives special relativity), to "one", for a gravitational horizon, in which case we have full gravitomagnetic dragging and the equation gives $(c-v)/c$ for recession. To most intents and purposes, we'd then have special relativity applying to all everyday objects, and would only have to start worrying about the gravitomagnetic factor for cases where the gravitational field of a mass had a terminal velocity that was a sensible fraction of the speed of light.

Unfortunately, this doesn't work.

37.3. A metric requires the Doppler equations to be universal

The requirement that classical light-propagation should behave as a purely local effect, depending only on the properties of the region (and for extreme high-energy nonlinear cases, also on the properties of the light), gives us compatibility with wave theory and the idea of a metric. We then require precisely the same Doppler relationships to apply to all moving objects, regardless of their gravitational field strengths. If we look at black hole or a neutron star, floating in deep space ten lightyears away, and there is a single hydrogen atom stationary with respect to the star, only one lightyear away (but in the same line of sight), then if we take a photograph of the atom-and-star, then change our state of motion along the line of sight and take a second photograph, we require the two sets of photographed signals to alter between the two images by precisely the same ratio. By the time the two signals have reached us, they are effectively a single signal stream, and when we change velocity, we are changing how we interact with that signal stream, rather than with the original distant bodies that generated it some time ago (no retroactive causality).

The existence of a metric requires the two signal sources to change frequency identically. [i]

37.4. The gravitomagnetic paradox

If the relative velocity of the moving star's field causes its Doppler relationship to diverge from the SR prediction, then the atom, somehow, must show precisely the same deviation as the star … and so must every other mass in the universe.

At first sight we seem to have a paradox – the form of the equations ***must*** vary as a function of gravitational field strength, but also ***must***, somehow, be identical for every mass. We therefore require every mass to show the same gravitational field intensity at its surface, which seems impossible, especially since different bodies can be immersed in different external gravitational fields.

There only seem to be two ways out of this:

Solution One ("SR everywhere") says that we stay with special relativity, and assume the absence of all gravitomagnetic effects. Unfortunately this means dismissing the results of the Gravity Probe B experiment, which successfully showed the existence of the Earth's rotational gravitomagnetic field, and also discarding any other gravitomagnetic behaviour predicted by general relativity. Since relativity pretty much needs the speed of gravity to equal the speed of light, $c_g=c$ (section 8), and gravity plus a finite speed of gravity gives gravitomagnetism, the universal application of special relativity would pretty much mean assuming that our universe contains no objects with detectable gravitational fields, a conclusion that is rather contradicted by the available evidence.

Solution Two ("Horizons everywhere") says that if the black hole case generates $(c\text{-}v)/c$, then all other moving bodies must generate $(c\text{-}v)/c$, meaning that every mass must behave like a black hole. This seems like a ridiculous suggestion, since the Earth and a tennisball and a grain of sand are most emphatically ***not*** black holes. To make this work, we have to hypothesise that all our everyday objects are made of smaller fundamental massed particles (which we might call "atoms"), and to assign the "atom" (or the part of the "atom" responsible for the emission and absorption of light) a horizon surface. We then need all fundamental massed particles to have identical surface properties.

As far as fundamental massed particles are concerned, the principle of relativity (and the idea of a metric) require aspects of a particle's gravity to be **quantised**. We do not have a smooth range of options for the gravitational differential between a (physical) observer and the functional electromagnetic surface of a particle. We only have two possible options – either the particle has no field (SR solution), or the particle has an effective horizon.

---

i To make the though-experiment even more exacting, we could suggest that by some highly-improbable state of affairs, the two signals travelling along the same path just happen to have ended up with the same frequency and phase, making it impossible to separate them.

> Just as Einstein's general theory forced relativity to intrude into gravitational theory, **relativistic gravitation** requires relativity to start merging with fundamental particle physics and quantum mechanics.

37.5. All or nothing

Could some less extreme departure from SR work? Apparently not – if all fundamental massed particles had an identical gravitational field that was *any weaker* than the horizon case, then when many of these particles were grouped together, and their fields overlapped, the surface field strength of the particles would increase, and this group of particle would then obey a different motion-shift law. The only way to have a surface field strength that is independent of any environmental field effects from neighbouring masses is to either say that masses have no fields at all (special relativity) or to *identify* the surface of a fundamental massed particle with its effective horizon surface, so that this surface extends when the particle in placed in a more intense background field and contracts again when it is removed.

Prior to the 1970s, this would have seemed like an obviously absurd idea: if we assigned horizons to particles, how can they possibly emit light? It turns out that while the SR solution is the unique solution that fits flat spacetime, the "$(c\text{-}v)/c$" solution is the unique solution that, when applied to gravitational fields, makes horizons observer-dependent, relative, and "acoustic", [23] making the emission of light by a particle a classical Hawking radiation event.

37.6. Corroboration from the principle of relativity

We will be understandably reluctant to "throw SR under a bus" based on a single unfamiliar argument to do with metrics and wave theory, but when we now turn to the principle of relativity, it concurs that there can only be one Doppler relationship, which must apply everywhere.

If an atom exchanges signals with a dense star, and the two bodies have relative motion, the principle of relativity requires that the signal frequencies received by the two bodies are not affected by whether we choose to calculate the motion shifts by assuming that it is the star that is "really" moving, or the atom. An atom is required to be able to replicate the dragging effects of any object that it could ever meet and exchange signals with, up to and including a black hole, and this is only possible if the atom (or the part of the atom responsible for emitting and absorbing light) has surface field properties that match that of a black hole.

37.7. Summary

> Compartmentalisation is not a valid option in relativity theory.

> In real life, "strong-gravity" and "weak-gravity" objects are allowed to exchange signals, as in order for us to be unable to tell who is "really" moving, both classes of object must obey precisely the same equations of motion.

> There are only two quantised solutions for relativistic gravitation: Either all massed particles making up bodies have exactly zero gravity, or they all have maxumum gravity. The principle of equivalence rules out the zero gravity solution.

38.SR Argument 38: "Metric arguments generate special relativity"

38.1. The "faceting" approach

It tends to be assumed that as soon as we assume a spacetime metric, we are dealing with special relativity. When gravitational fields are involved, spacetime is curved, whereas Minkowski spacetime is flat: we get around this by ***faceting*** – we say that any curved surface can be approximated arbitrarily well by a *faceted* surface, where we can always improve the accuracy of the match by making the facets smaller, and where special relativity applies within each individual flat facet. As we define smaller and smaller regions, special relativity's approximation becomes increasingly accurate, and at the limit at which each facet is shrunk to an infinitesimally small pointlike region, special relativity becomes, in effect, an exact theory (of nothing!).

In this (discontinuous) approximation of continuous classical physics, special relativity can describe the physics that is going on *within* each facet, with gravitational theory describing the effects of the angled disjoints *between* the facets. These disjoint edges can be considered to be the curvature that belongs to the region, swept out of each individual fact to make it flat, and exiled to the facet boundaries. As the number of facets is increased, the curvature is shared out amongst more angular edges and the angles become shallower and flatter, until the edge angles become imperceptible.

We then have what appears to be a geometrical proof that gravitational (curved spacetime) physics reduces to the (flat) physics of special relativity over small enough regions.

38.2. Shortcomings of the faceting approach

If we look at the region surrounding a star, it will be curved. If we zoom in on a small region we should eventually be able to define a tangential facet whose centre touches the actual surface, and whose dimensions are small enough that the the facet and the adjacent surface are effectively indistinguishable. We then feel entitled to say: "*we know that the internal physics of this flat region is governed by special relativity*".

But this reasoning does not sit well with the principle of equivalence of inertial and gravitational mass, because once we assign particles their own fields, the internal physics of a facet cannot include the physics of moving masses – it becomes a precondition of its flatness that it does not contain any massed particles. A region containing a pair of particles can be described as a pair of gravity-wells (with a relative tilt, if the particles have relative motion), the shape of which then has to be approximated by subdividing the region into a significant number of even *smaller* facets.

If we now want to derive the physics of our two particles using geometry, although the region separating them will be tiled by an arbitrarily large number of facets each of which *nominally* give SR *in isolation*, the full description of how light and forces propagate through the region involves not just the facets, but also angles between the facet boundaries. We cannot ignore the additional physics and geometry of the boundaries. Although we can make the effects *per edge* arbitrarily weak, by using an arbitrarily large number of facets, our signals then have to cross an arbitrarily large number of these edges.

> The relativistic geometry of moving masses is not described by flat spacetime. It is described by necessary *departures* from flat spacetime.

> Small regions can only conform to SR physics if they are matter-free. If we have SR then we do not have moving matter, and if we have moving matter we do not have special relativity.

> Special relativity is a physics that depends on the number of masses being zero. If there are masses, the geometry (and the resulting equation-set, and the physics) is different.

38.3. Is the assumption of a metric equivalent to the assumption of the principle of relativity?

The arguments of section 37 present an intriguing possibility. In section 37.6 we derived a result for relativistic gravitation from the assumption that the principle of relativity held for disparate gravitational bodies, while in section 37.3 we had derived the *same* result by assuming that the containing region obeyed the condition of wave-theory-compatibility that must apply to metrics.

Might it be that by assuming a metric we are automatically assuming the principle of relativity?

Perhaps. The principle of relativity assumes that there are no absolute references, that nothing is “nailed down”, and that matter “swishes around” freely according to whatever else is going on in its region. A “free” metric, untethered to any form of absolute externally-imposed reference, and with no particulate properties that allow us to identify absolute motion, is, similarly, a self-contained, free-standing system.

The principle of relativity says that the same laws of physics apply regardless of the state of motion of the observer. A metric requires the same laws of geometry to apply, and the same intrinsic interrelations of features to be the same, regardless of what translations, projections or topological transformations we choose to apply to the geometry, and to hold for projections made from the point of view of any massed particle in a region. It would not be entirely surprising if the most efficient laws that we could derive for the first case turned out to be identical to the most efficient laws that we could derive for the second.

38.4. Understanding the laws for metric theories

If it *does* turn out that assuming a “free” metric means assuming the principle of relativity, then it would seem that if we assume a gravitational metric, with no prior geometry, we have the apparent return of the *general principle* of relativity.

This is where textbook geometrical physics breaks down.

On the one hand, surely a covariant metric allows us to transform between any sets of arbitrary coordinates, and must support the general principle of relativity?

On the other hand, when we try some exercises with arbitrarily-selected coordinates and states of motion, we tend to find that the results *do not* fully correspond to the general principle of relativity (Schild, Moller), at least, not as far as we might expect. We tend to find that the fields associated with the selected sate of motion are distinguishable from “real” gravitational fields, in that they can be made to go away by a suitable alternative choice of coordinate system.

This has left geometrical physics with the problem, that our equations are considered valid because they have been selected for their support for the GPoR, and yet when we test them we keep finding that our examples *do not* support the GPoR. We find the type of solutions described by Moller and Schild, in which “extended SR” applies … and “extended SR” is geometrically incompatible with the general principle of relativity.

The GPoR *surely must* hold, and yet it appears not to.

38.5. Solving the metric paradox

The solution to our little paradox is to realise that a workable gravitational metric simultaneously supports two classes of solution, and two types of coordinate transformation. **Category 1** must obey the GPoR, and **Category 2** needn't (and probably shouldn't). Category 2 transformations are far more common, which is why we are likely to get a Category 2 when we apply a random, arbitrary transformation.

Th significance of the distinction between the two categories is that the GPoR-compatible Category 1 is the smaller set of transformations that correspond to actual physics, and Category 2 are the larger set of transformations that do not.

We start to understand this distinction when we remember that all physical masses must have associated curvature, and that all *moving curvature sources* must have associated gravitomagnetic distortions. If we look at a region of spacetime populated by massed particles, then every particle has a worldline, which appears as a curvature "streak" marking the particle's path through spacetime. This limited set of identifiable, *special* worldliness represent the only possible positions from which physical observations can be made, and the only locations at which physical matter can be seen. If we transform between the view of the metric from one of these "special" worldlines and another, the comparison must obey the general principle. These comparisons give us the actual observer-physics of the metric.

If one of these "physical" worldlines shows a *physical* acceleration, then the metric surrounding the path will show a corresponding gravitomagnetic field through which it sees its environment, and through which its environment sees it. This curvature is intrinsic, and real for everyone. If we are doing the exercise properly, we can also point out that the rest mass and velocity of particles should *also* be associated with intrinsic curvature, both rest curvature and gravitomagnetic. If a massed particle exists for one person, it exists for everyone that can see that region of spacetime.

On the other hand, if we just draw a curved path though the spacetime block at random, the apparent curvatures seen by the fictitious observer are not intrinsic and can be distinguished from real (Møller: "*permanent*") fields. These are non-GPOR-compliant. [i]

> We then have a conceptually simple distinction between our two classes of geometrical transformation. There is a **preferred set** of worldlines which correspond to the paths of actual particles, whose experiences generate phsyical law, and which agree with the GPoR … and we have a far larger set of worldlines that identifiably *do not* correspond to real masses, whose different geometrical behaviour *cannot* generate exact physical law, and which *do not* obey the GPoR.

> It appears the the laws of geometrical physics, which are *only obliged* to apply to actual physical observers, are somewhat "picky" – the refuse to apply to mathematical "ghosts" and other non-physical entities, and *only apply* to observers that actually physically exist.

This distinction is so specific (and so obstinate!) that it suggests that we *may* be dealing with the shape of fundamental, exact and potentially final physical law.

---

i An objection here is that we have skipped over the subject of rotating bodies. Surely defining phsyical behaviour based on worldines rather than frames is inadequate, because although a worldine can support velocities and accelerations, it cannot support rotation as a property? Neither (arguably) does a simple point-particle. Our response is that although rotating *bodies* are not usually fundamental particles, their *constituent* particles (with horizons) will have worldlines that describle circling paths in spacetime, whose distortions (individually and combined) then, again, generate real physical deviations from flatness, in accordance with the general principle.

38.6. The physicalisation of spacetime

The realisation that the same metric can simultaneously support both physical and unphysical transformations, obeying different rules, is both enlightening and annoying. How do we identify the “physical” solutions? Does this require new math?

- Given that the arguments elsewhere in this paper indicate that the GPoR requires fundamental massed particles to be horizon-bounded, our test for whether a path is physical or unphysical would seem to involve dividing the metric into two types of region separated by horizons: a main “bulk” region, and the remaining networks of tubelike horizon-bounded regions, each section of which encloses a “physical” worldline. [i]
- Minkowski-like arguments suggest that any point (“event”) in spacetime can be described as lying on multiple possible intersecting worldlines, any of which can be considered legal as long as they remain within the Minkowski lightcone. The metric then supports a continuum of possible alignments of the space and time axes for the point, as long as all the time axes remain within the cone. For a GPoR-compatible system, we have to instead consider whether the worldline for a given fundamental particle represents a preferred local time axis, and an absolute orientation of time for any point on the worldline.
- We then have to consider what it means to have space and time axes applied to the intermediate regions of spacetime, where (by definition) there are no observers. Do we revert to a Minkowskian approach, and say that we are free to align space and time coordinates however we like (within cone limits), since there is no observer present to prove us wrong?
  Might it be correct for some “unphysical” laws that only hold in the absence of observers to nevertheless still give a correct geometrical model for the propagation of electromagnetic fields in regions that are genuinely otherwise empty? Should we distinguish between “observer” physics (which requires massed particles), and a separate layer of physics

Some of these issues and questions are difficult, and will require further analysis.

38.7. Summary

If a fully covariant metric automatically supports the general principle of relativity, then since the GPoR is incompatible with SR physics (Schild, 1960 [46]), we have a situation in which the existence of a gravitation-compatible metric would automatically *invalidate* SR as physics.

This would be in general agreement with the arguments of sections 37.6 and 37.3, that the principle of relativity applied to gravitation, and gravitational bodies embedded in a metric, both eliminate special relativity as a viable possibility. If we have a metric curved by the existence of gravitational bodies with relative motion, the requirement that we must be able to describe the same geometry from the point of view of any body, and explain the same underlying intrinsic geometry, would seem to be equivalent to agreeing to the general principle of relativity.

> While textbooks argue that assuming a metric makes SR compulsory, we actually find that assuming a consistent metric (plus gravitation) makes SR *physics* impossible.

---

i The physical worldlines will form multiple networks: if a gamma-ray creates a particle-antipartile pair that then self-annihilates, then this will give a single “toruslike” loop of tube, which will not be connected to other physical worldlines.

> Metric arguments, applied naturally to gravitational problems, say that the relativistic relationships of observations made by differently moving masses are not described by flat spacetime, and are therefore not correctly described by special relativity.

> The Moller/Schild argument that the general principle of relativity is not correct is based on examining the properties of a GPoR-compliant metric *that do not relate to the motion of actual matter.* If we restrict ourselves to *only* considering comparisons that refer to matter, the GPoR seems to be fine.

The properties of a GPoR-compliant metric seem to be either those of a relativistic acoustic metric, or to something incredibly similar. The metric needs to be nonlinear, it needs to describe the propagation of gravitational waves as nonlinear, and its limits (horizons, velocities) need to be established dyamically rather than prescriptively.

Going further to address the breakdown of Einstein's theory when black holes move, we find that the extended horizon associated with a receding black hole needs to be an acoustic horizon, and the exterior physics of a moving black hole needs to be acoustic. These requirements are incompatible with special relativity.

> A metric implementation of general relativity seems to require an *acoustic* metric.
> Acoustic metrics are not SR-compatible.

> If matter is associated with curvature, a valid metric cannot generate SR relationships for moving matter.
>
> The textbook definition of a metric theory as being a theory that **(i)** has a metric, and **(ii)** supports SR physics, is incompatible with the general principle of relativity and the PoE.

It would seem that the second clause imposing special relativity **(ii)** was only necessary because if we did *not* put in this manual override, gravitational metric theories would default to *invalidating* special relativity and therefore also invalidating Einstein's general theory. In order to protect SR/GR1916, theorists simply override the geometry, and added a (geometrically impossible) manual condition that declared SR to be compulsory.

> Any theory that conforms to the SR-centric definition of a metric theory is almost ertainly invalid.

39.SR Argument 39: "Observer-specific effects are irrelevant"

It is tempting (and understandable, and to some extent in "primitive" physics, even good practice) to try to eliminate the role of the observer from our analyses. We want to be able to reassure ourselves that what we are describing is not an "artefact", and relates to some sort of physically real objective reality that would still exist if we were not actively looking at it.

When we get to more advanced models, where our task is to establish the rules for how matter communicates with matter via light, this approach is no longer trustworthy: the nature of the interaction of a physical observer-particle via light is no longer an inconvenient complication to be defined away so that we get get to the "real" underlying physics – it *is* the real physics.

When the relative speed of bodies is a significant fraction of the speed of light, and Doppler effects significantly alter the energies of received signals, then these are not "artefacts", they are a fundamental aspect of physical law.

Similarly, for real observers (with mass), the choice of whether or not to model the observer's own field is important, because under a relativistic system, the properties that we assign to observers are also the default properties for all other bodies. The gravitomagnetic effect of the relative motion of masses physically alters the properties of lightsignals sent between them. Remove the observer's own field, and things can start to go horribly wrong.

39.1. The observer problem

The issue of the observer's own gravitational field is touched on by Møller:

> Møller (1955), [47] page 290: " *Strictly speaking, the particle itself will create a gravitational field which should also be described by the functions g_{ik}. In the present sections we assume, however, that this field is weak in comparison with the external field so that its influence on g_{ik} may be neglected. …* "

This approach is dangerous, and can lead to the inadvertent destruction of the general principle of relativity. According to the GPoR, *all* apparent fields are to be considered *legitimate* gravitational fields, and we must not make a physical distinction between the "fictitious forces" of old Newtonian theory and "real" gravitational forces.

39.2. "Correct physics" according to the general principle of relativity:

If two masses or two massed systems have a relative *rotation*, this causes a "twisting up" of the intermediate region of spacetime, and this field distortion or geometrical distortion is "real" in the sense that it exists in the frames of both masses, in all intermediate frames, and also in all other frames. The field cannot be eliminated with a clever choice of coordinate system – it is physically there for everyone, inertial or noninertial.

If two masses or massed systems have a relative *acceleration*, then again, the requirement that the same laws of physics operate for both observers means that if a mass feels geeforces due to its relative acceleration to the background stars, then background observers must in turn feel an effect due to the relative acceleration of the mass (Wheeler, "democratic principle" [48]). In a GPoR-compliant theory, forcibly and physically accelerating a mass so that it experiences geeforces causes a dragging effect on nearby matter and light, as a backreaction – the region between the two masses or two systems is physically distorted and has an *intrinsic curvature* that exists for both systems and for any other onlookers, and again, cannot be made to go away by an inspired choice of coordinate system.

In both exercises, the general principle requires that when a physical observer experiences a gravitational field, the distortion effect is not merely fictitious – it is a physically tangible and corporeal effect, associated with an identifiable intrinsic distortion of spacetime – the existence of the field is unambiguously a part of physical reality and the region's geometry, for everyone.

39.3. How to screw up general relativity, and blame it on the theory

The general-relativistic logic is so clear and simple that one might wonder how on earth we could possibly mess it up. This is achieved by doing what Moller did on the previous page – by assuming that the mass-field of the observer could be treated as insignificant and therefore set to zero.

If the supposed observer has *no* mass, and nominally rotates or accelerates relative to other matter, then its own relative motion is NOT altering the physical shape of the metric (because the **pseudo-observer** does not physically exist). We then have distortion-fields being experienced by the pseudo-observer that are NOT seen by background observers, and a mathematician can point out that there *is indeed* a physical difference between "real" gravitational fields that exist for everyone, and "merely fictitious" fields. Our mathematician can then conclude, quite wrongly, that fields due to relative rotation or acceleration are, similarly, not "real", that we can distinguish between these fields and "real" fields, and that the general principle of relativity is therefore wrong (Schild 1960 [46]), or that it needs to be downgraded from physical law to something more akin to a useful rough heuristic guide.

But this apparent proof of the invalidity of the GPoR is entirely dependent on our initial *bad*, *unphysical*, *illegal* decision to set the observer's own field to zero. The general principle *appears* to break because we broke it ourselves (by breaking the principle of equivalence) in our initial assumptions as to how to tackle the problem. If the field is positive (as it must be), then the observer's acceleration or rotation again warps spacetime, the existence of the warpage is agreed by everyone, and the general principle is fine. [i]

We can see here that Einstein and the relativity community underestimated the *specificity* of the general principle. While we constructed a set of mathematical tools and hoped that the result would apply to matter, it turns out that the GPoR *only* applies to matter, and does not apply to frames, or to massless pseudo-observers.

> The general principle of relativity only applies to masses and massed bodies and systems that actually exist. It does *not* apply to masses that do *not* exist.

> It does not apply to arbitrary transforms between fictitious worldlines that do not correspond to worldlines of actual matter.

> The Schild and Moller disproofs of the general principle are based on an inappropriate use of massless observers.

In short,

> The observer's own gravitational field is not something that we can ignore without disastrous consequences – the general principle of relativity doesn't work without it. [ii]

---

i If we are doing general relativity *properly*, the simple relative motion of massed particles also warps spacetime.

ii If I watch a moving star, its moving field causes additional distortion, and its equations must be non-SR. If I move, and declare myself to have no field, then my motion will NOT distort the metric, and my equations will agree with SR. When I exchange signals with the star, I will then get different physical outcomes depending on who is "really" moving, breaking the principle of relativity.

39.4. Re-engineering general relativity

Einstein's vision for a general theory was of a geometrical system in which the distinction between inertial and noninertial frames disappeared. **The same laws of physics apply to observations made in any frame.**

> Einstein 1916: "*The general laws of nature are to be expressed by equations which hold good for all systems of coordinates, that is, are c-variant with respect to any substitutions whatsoever (generally co-variant).*"

What we have learnt from this exercise is that Einstein's statement needs modification. If we have a gravitation-compatible metric, and laws that let us express its geometry as seen from *absolutely any* location and state of motion, and convert this information into the view from any other location and state of motion, then these laws will be a *superset* of the actual physical laws, and will include behaviours that are "not physics". The vast majority of solutions will not be physical – the only physical solutions will be the ones that map one actual *physical* particle's worldline to another. These solutions are the subset in which the general principle of relativity applies.

In a valid general theory of relativity, we have to discard the "frame" abstraction (as far as derivations are concerned), and go back to fundamentals. The new (old) rule is that **the same laws of physics apply to observations made by any *<u>mass</u>*.**

Again in the 1916 paper, Einstein says:

> Einstein (1916): "*… there is no immediate reason for preferring certain systems of co-ordinates to others, that is to say, we arrive at the general principle of co-variance.* "

This, again, probably needs modification. While we should indeed be able to use geometry to transform between any arbitrary system of coordinates, the comparatively tiny proportion of paths through spacetime that correspond to the worldlines of actual particles have a special significance, as these are the only paths that correspond to real observers. These paths can be identified by the "streaks" of the particles' associated curvatures.

Transformations that map between worldlines of actual moving masses conform to the general principle of relativity, and generate exact laws of physics. The distortions seen from one physical worldline will exist for all physical worldlines (and will exist as intrinsic curvatures).

Transformations that map between views corresponding to *arbitrary* lines drawn through the region that *do not* correspond to particle worldlines, will generate apparent, "fictitious" fields for the unphysical pseudo-observer, and perhaps other artefacts. These transforms do not have to agree with the general principle of relativity, or generate correct physical laws. But this does not matter, because, *by geometrical definition*, no masses capable of making observations are moving along these paths.

> The Schild and Moller arguments against the general principle of relativity depend critically on there being no actual masses travelling along the paths specified: they represent geometrical laws of "unphysics". [i]

i A perfectly engineered machine, whose design depends on perfection, can be halted by a single grain of sand in the gears. As we close in on a final, *exact* theory of physics, the structure becomes more sensitive to inappropriate unphysical assumptions. The GIGO principle (Garbage In, Garbage Out") becomes stronger. The good thing about this lack of fault-tolerance is that we immediately know when we have made a mistake, because the theory breaks. Avoiding approximations, and strict adherence to the general principle ought to guide us to the exact final theory.

39.5. Observers that don't exist, don't exist

Relativistic gravitation requires observers to have gravitational fields. If we want to explore how a region would look by defining the position and state of motion of an observer, to get an exact agreement requires that an actual mass be present at the location, moving in that way. We cannot simply draw arbitrary lines across a surface and say: "*The intersection of the geometry with these lines describes how the region would look to any observers following these paths.*" These are **pseudo-observers**. If the paths describe accelerations, then the acceleration of a *physical* observer will further distort spacetime around the paths, and the altered lightbeam geometry may change how things look to that observer, who is viewing their surroundings through these distortions.

We may use the idea of "test particles" to model an adequate approximation of what one would see in real life – after all, the interaction of two mutually orbiting neutron stars will not be significantly changed by the presence or absence of a single background hydrogen atom. But we cannot safely use massless test particles as part of a *derivation of physical law.*

39.6. Voting rights

In a "democratic" universe (Wheeler [48]), this principle can be expressed by saying "*If you don't exist, you don't get a vote*". If we are doing relativistic *physics*, we do not care if an observer that doesn't exist (according to spacetime geometry) supposedly sees the principle of relativity being broken, or the local speed of light to have the wrong value, or has their feelings hurt by our ignoring them. They cannot communicate their disappointment to us. [i] If the hypothetical breakdown (as seen by a hypothetical massless pseudo-observer) has no consequences in the physical world, then we should not lie awake at nights worrying about it. [ii]

If we carefully craft a theory of how moving masses (with associated distortions thanks to the principle of equivalence) interact and communicate, why should we care if those laws then turn out not to apply to illegal "curvature-free" observers whose very existence would violate the principle of equivalence? [iii]

39.7. Nulls and zeroes

It is natural when confronted with the problem of observer-properties to be reluctant to add further complexity to our models, and, instead of adding extra detail to deal with what these properties might be, to try to produce a simpler, and hopefully more theory-neutral approach in which we delete the properties of the observer altogether.

But this approach is *not* neutral.

> To assign the observer "no properties" is to assign the *property* of having no properties.
> It is an additional postulate, that can change the resulting geometry and physics.

---

i It is also a current feature of voting systems in the field of politics that human beings that do not physically have a corporeal existence are not allowed to vote. We are not entitled to register our "imaginary friends" as voters.

ii It would of course be *nice* to have a system of relativity that works for both corporeal and non-corporeal observers … but this does not seem to work. In the current system we have compromised our ability to properly model the interactions of matter in order to have a system that also applies to the purely imaginary.

iii The closer we get to a final theory of physics, the tighter the constraints are likely to be, and the more likely it is that unphysical assumptions *ought* to lead to nonsensical results. While finding that a system *only* works for realistic matter and fails for "fantasy particles" may annoy a mathematician (who will often want a solution that is as general as possible), for a physicist it is an indicator that perhaps we might be closing in on a final solution.

39.8. Adjunct theories

In an "adjunct theory" (*e.g.* Wheeler and Feynman, 1949 [165]), we try to model a particle while ignoring its own complicating field and self-interactions, by saying that the motion of a particle (and its fields) is defined purely by the sum of background fields and background geometry – by taking into account everything in the universe *except* the particle's own fields and distortions.

But the "adjunct" approach has problems: suppose that we aim two singularity-particles with different masses at each other at high speed, and they undergo a near miss, swing around each other by a large angle at the closest point (exchanging momentum), and then fly away from the encounter at two very different angles to those of their initial paths. If we are modelling one particle by assuming that everything else in the experiment is "fixed", then we will say that the other particle is initially travelling in a straight line, *and will continue moving in a straight line* over the course of the experiment, because that is what was initially defined.

We cannot therefore get an exact answer for the path of a particle that interacts strongly with other particles by modelling only the field that would exist if the particle was not there, plotting a geodesic and then saying that a particle would then follow that geodesic. The existence of the particle's field needs to be part of its physics.

> Einstein and Infeld (1949): [73] *"In such a solution the same two world-lines would appear together that before appeared singly. Therefore the field with its linear laws cannot imply any interaction between the singularities. Thus only non-linear field equations can provide us with equations of motion"*

Extreme nonlinearities are annoying and inconvenient, but do seem to be part of reality.

> To get the correct answer, we have to take into account that as a particle moves, it is affecting the spatial geometry of the region it moves through, and that this geometrical interaction with other masses can then in turn affect its own path.

> In the relativistic physics how how objects and observers interact, the observer is part of the physics. Over-idealised observers lead to over-idealised physics.

39.9. Summary

What does all of this mean for special relativity?

The Schild/Moller arguments show that we cannot derive a legitimate general theory of relativity using field-less pseudo-observers. The GPoR requires all observers to be real, with nonzero mass, and to leave corresponding "footprints" in the metric, and it applies to moving masses, not frames. [i]

This makes a working general theory incompatible with special relativity, which relies on observers not having fields, and supports frame-based logic.

> This is yet another way of saying that a valid general theory (which requires particles to have gravitational mass) cannot reduce to the physics of a special theory that requires particles *not* to have gravitational mass.
>
> Allowing "curvature-free" observers in order to accommodate special relativity wrecks relativistic gravitation.

---

i Every time we invoke a frame argument (section 34), we are saying that the observer's own interaction isn't relevant to the situation being described. This can be fine for generating external views of how two *other* systems interact … but the observer cannot be one of the systems. "Frame" logic is not normally suitable for derivations.

40. SR Argument 40: "We know that curvature associated with small masses is far too small to have any meaningful effect"

40.1. The "gravity is weak" argument

This argument says that although we know that *in principle* special relativity cannot genuinely be a perfect limiting case of gravitational theory (because the principle of equivalence doesn't allow a massed particle to have *absolutely zero* curvature), the gravitational fields of small masses are so absurdly small that in practice we can forget about them. If the ratio of the strength of an electron's gravitational field to its electric field horribly small, making the field undetectably puny, then any *modifications* of that tiny field (which we'd expect to be even smaller) will be so weak as to effectively be zero.

We can then treat particle physics as being *effectively* a flat spacetime problem, and consider any resulting errors to be too small to ever be of interest to humankind. SR will be *effectively* correct to all practical intents and purposes.

However, when we calculate the gravitomagnetic shift on light emitted or received by a particle, we do not care what the particle's influence is on a particle on another particle a metre, or even a centimetre away: we care how strong its curvature effects are at its electromagnetic surface, *at effectively zero separation.* In an abstract exercise, where we treat a particle mass as being effectively pointlike, there will be a finite positive distance at which the mass has an associated curvature horizon, and at this distance, the particle's mass-field is sufficient to totally drag light, and (if the particle is allowed to be complex enough to support rotation), to totally drag light around with it if it rotates – its influence on inertia at this distance is able to overwhelm the combined effect of the entire outside universe.

40.2. The "strong gravitomagnetism" scenario

If the particle's horizon is identified with its effective surface for light absorption and emission, then the gravitomagnetic effect is not vanishingly small, it is dominant, and if an atom is moving, the gravitomagnetic "Doppler shift extinction" argument from section 36.4 suggests that when we try to measure a moving massed particle's traditional Doppler shift, what we are actually measuring is its gravitomagnetic shift.

Further, if stationary particles really have no detectable influence on lightspeed, then there should be no such thing as refractive index, and if the motion of particles has no effect on light, then we should not have the Fizeau effect.

The belief that gravitational and gravitomagnetic effects are too small to have ever been seen in the laboratory does not reconcile with the fact that we have been exploiting refractive index to make lenses for perhaps two and a half thousand years, and confirming the existence of the Fizeau effect since 1850.

It may be that rather than particle-curvature effects being too small to measure, they are *so large* that we fail to notice them due to overfamiliarity. [i]

---

i We do sometimes miss the obvious. In some Nineteenth-Century works, the authors discuss the hypothetical curvature of space, and wonder whether there might be some way to detect the consequences of irregular curvature … and occasionally even suggest that perhaps there might be some small-scale particulate behaviour that we already know about, caused by curvature, but which we don't *recognise* as being due to curvature … and miss the retrospectively-obvious connection between curvature and gravitation.
Similarly, Newcomb expressed scepticism in the 1890s about the idea of the existence of a fourth dimension, apparently not noticing the (retrospectively obvious) application of the idea to time coordinates.

40.3. Inertia and gravitation

The relativity princople does not allow us to blame a body's inertial behaviour on the existence of absolute space or some absolute set of references. If we take away all background matter, and all possible external references that might let us tell whether a body is or is not rotating or accelerating (making these behaviours "unphysical"), then the body's inertial resistance to induced rotation or acceleraiton should disappear (because making an "unphysical", purely hypothetical change to a system should not require any physical effort).

If we then apply the relativistic Mach-Einstein principle that the inertia of a body is therefore partly (or wholly) a function of its interaction with background matter, inertia becomes a result of a field interaction between our test body and background mass.

The immediate properties of the inertial field are that intensifying the field in a region increases inertial effects in that region (slowing timeflow), and that if the field varies across space, light and matter are deflected by Huyghen's principle towards the region of greatest field-intensity.

> The inertial field is the gravitational field.

The range of effects that we traditionally *think* of as being gravitational is far more limited than the range that can be modelled using curvature concepts. If anybody wants to object that gravitational effects of small objects are far too small to have measurable consequences, we can ask, does this mean that lab-scale bodies' *inertial masses* are too small to measure?

40.4. Small deviations from SR don't work

A consistent theory of relativistic gravitation turns out not to permit "insignificant" deviations from SR. Relativity requires our choices regarding the application of gravitational effects to fundamental massed particles to be *binary*: the it requires either *absolutely zero* deviation from SR, or a full-blown Lorentz deviation to the red (curvature horizon).

Relativity does not allow "small" or "insignificant" departures from Minkowski spacetime: we must either have *perfect* Minkowski spacetime (and no such thing as gravitational mass, even for stars), or we must jump to straight to a different theory of relativity (by default, based on a relativistic acoustic metric), whose equations, definitions and principles are substantially different to those of the 1905 theory.

> *There is no such thing as a fundamental massed particle with a tiny gravitational field.*
>
> We either have a *maximally strong* field, or we have no field at all.

In a universe that supports any form of gravitation, the correct starting point for deriving the relativistic laws of inertial physics is then not the hypothetical geometry of a spacetime where bodies have zero curvature – it is the geometry of a spacetime where all masses show the maximum curvature possible.

40.5. Summary

> For the problem of gravitational/inertial physics, there appear to be two – and only two – relativistic solutions – one for universes in which there is no gravity, and one for which the electromagnetic interfaces of all fundamental massed particles are represented by horizons.
>
> There is no such thing as a weak-gravity solution to relativity theory, only a no-gravity solution and a full-gravity solution.

41.SR Argument 41: "SR is right because major physics theories are never really wrong"

41.1. Infallibility

According to this argument (which is apparently being taught in some physics classes), major theories of physics are never found to be wrong, they are merely superseded by more advanced theories that are *even more right*, and in which the older theory lives on as a limiting case or useful first approximation.

The physicist can then say: "*Newtonian theory was never **disproved**, it instead became a limiting case of special relativity, and special relativity was never overturned, it in turn became a limiting case of GR1916. Similarly the laws of the conservation of mass and of energy were never overturned, they became limiting cases of a larger, more inclusive law of the conservation of massenergy*".

Physics is portrayed to the student as a sort Victorian-style single-track concept of evolution in which the subject always progresses upwards and onwards along a single illuminated path, and in which the current state is an inevitable step toward future perfection, with no wrong turnings or dead ends. According to this worldview, special relativity will not and *cannot* ever be overturned, because That's Not How Science Works. [i]

41.2. Previous breakdowns

This worldview is, of course, fiction.

Newton's original system had major flaws that were deleted or changed when the theory was redefined in the early Nineteenth Century, and inverted some critical relationships.

Similarly, Einstein's general theory, considered either as a principle theory or as a geometrical theory, was a failure.

> The failure of major theories is a natural part of science, and should not be seen as something to be dreaded. It is a natural part of the process of scientific progress.

> If we say that our three biggest theories of classical physics are Newtonian theory and Einstein's special and general theories, then two out of these three have already failed. If we recognise invalidation as being the norm rather than something that never happens, it becomes less outrageous to ask whether the third theory might also have issues.

41.3. The "cheerleading" problem

In order to solve a theory's problems (or produce a better theory) we first have to acknowledge that problems exist. This means admitting that our current state of knowledge is fallible. This is something that the physics community is not always good at:

> Misner, Thorne and Wheeler "MTW", **Gravitation** (1974) [53] §44.2, page 1199:
> " *No inconsistency of principle has ever been found in Einstein's geometric theory of gravity* ". [163]

Given that MTW *cite* the 1960 Schild paper, and refer to gravitational collapse (§44.1) as "*the greatest crisis in physics of all time*", it is difficult to take the "boxed" statement above seriously.

i This is of course slightly problematic from the Popperian viewpoint that a theory has to be falsifiable in order to be considered "scientific".

41.4. Science vs. theology

A belief that orthodoxy cannot be wrong is damaging to science. Science is supposed to be different to, say, theology, in that scientist is supposed to question *everything*. If we teach students not to question standard beliefs, and to simply extrapolate and continually add further layers of complexity, then what they are doing is not scientific but theological. [i] While science has a theological *aspect*, the scientist differs from the theologian in that they are expected to ask "what if everything I've been taught is simply wrong, and I need to start over?" [ii]

The theologian does not have this luxury – a priest is not expected to ask "What if the Church is not a force for good?" or "Perhaps we should investigate whether Buddhism might be a superior system?". Religious organisations are quite capable of carrying out other activities associated with science, such as data-collection and modelling (the Vatican has an astronomical observatory) – what makes religions different from science is their non-negotiability of received beliefs.

If physics students are being taught that the correctness of mainstream science is not open to question, then this training is essentially religious-theological rather than scientific. If there *is* something fundamentally wrong with current mainstream theory (as there was in the early C18th), these students will have been trained not to see the problem, and not to believe that there *can* be a problem, even when it is pointed out to them. [iii]

41.5. Summary

While one could be forgiven from reading educational/promotional matterials that major accepted physics theories are never wrong, the reality seems to be that some members of the community are simply very good at pretending.

> Most major successful classical theories up until the Twentieth Century turned out to be wrong in some way. If special relativity should *also* turn out to be wrong, this would not be the unthinkable and unprecedented breakdown of a major theory – it would be entirely normal behaviour.

---

i The behaviour of the scientific method can be likened to that of an **Abelian Sandpile**. If we take conical pile of sand whose sides slope at the critical angle, and drop a new grain of sand at the peak, the grain will either roll down the slope towards the bottom, reducing the angle, or snag on grains near the top, increasing the angle. As we pile on more and more grains, a local region eventually becomes too top-heavy and collapses. The pattern of collapses is chaotic, and fractal, with collapses more common at smaller scales, and progressively rarer for larger regions.
In the scientific model, we start with a base of hypotheses designed to support a certain amount of data, and pile new data and hypotheses on top of it. Eventually the logical structure becomes top-heavy, the base of the region is no longer the most efficient system to explain the data it supports, and the region collapses into a new state, supported by new hypotheses more suited to cope with the additional data. The largest-scale collapses or iterations – significant *scientific revolutions* – occur perhaps once a century. To someone living between major revolutions, it may seem as if the large-scale structure of the sandpile is solid and stable: but a collapse may happen at any time.

ii Scientific method. The theologian will say: "*Given these initial fundamental truths, what are the consequences?*" The scientist, similarly, will ask: "*Given these initial assumptions, facts and principles, what are the consequences?*" Both will then construct elaborate frameworks of logic and secondary rules. Where the scientist differs is that they will also ask, "*Might the system be better if we used different initial facts and principles, or allowed the possibility that some of the data being relied on might be wrong or misleadingly incomplete?*"

iii The idea that the 1916 general theory might have a short shelf-life was understood by Einstein. The Ptolemaic system had lasted for over a thousand years, Newton's had lasted a few hundred, special relativity's reign as the main theory of relativity had only lasted a little over a decade … if the pattern continued, GR1916's spell in the limelight might be very brief indeed.

42. SR Argument 42: "Anyone who questions Einstein's approach with special and general relativity has not understood it"

42.1. The original structure

According to Einstein's 1919 characterisation of general relativity for *The Times* newspaper (London) [50], special relativity provided a flat-spacetime foundation for the system, with the general theory then providing additional phenomena related to gravitational effects.

> Einstein, 1919, " ... *I must observe that the theory of relativity resembles a building consisting of two separate stories, the special theory and the general theory. The special theory, on which the general theory rests, applies to all physical phenomena with the exception of gravitation; the general theory provides the law of gravitation and its relations to the other forces of nature.* "

42.2. Einstein's 1950 article

By 1950, Einstein was querying the validity of starting with SR and then trying to build on top of it it – perhaps a general theory ought to be constructed only from components that were certified from the outset as being compatible with the general principle of relativity, and perhaps it might require new ideas that couldn't necessarily be derived incrementally? [164]

> Einstein, "On the Generalized Theory of Gravitation" (1950) [164] " … *all attempts to obtain a deeper knowledge of the foundations of physics seem doomed to me unless the basic concepts are in accordance with general relativity from the beginning. This situation makes it difficult to use our empirical knowledge, however comprehensive, in looking for the fundamental concepts and relations of physics, and it forces us to apply free speculation to a much greater extent than is presently assumed by most physicists. …* "

… he then criticised the two-stage approach that he'd taken with GR1916, of trying to deal with "gravitational" and "non gravitational" physics separately, in the hope that the two would then turn out to be compatible. There was no reason to assume that GR-type principles *only* applied to gravitation, and didn't also apply in the realm usually dealt with by SR …

> " *… I do not see any reason to assume that the heuristic significance of the principle of general relativity is restricted to gravitation and that the rest of physics can be dealt with separately on the basis of special relativity, with the hope that later on the whole may be fitted consistently into a general relativistic scheme. …* "

… he then characterised the two-stage approach as being something "*historically understandable*" (as in, "the best that could be managed at the time"), but no longer justifiable with the benefit of hindsight (from the "modern" perspective of 1950). The "SR plus GR" approach was now out-of-date. Einstein argued that he no longer considered it legitimate to try to model inertial physics separately without taking into account curved-spacetime arguments.

> " *… I do not think that such an attitude, although historically understandable, can be objectively justified. The comparative smallness of what we know today as gravitational effects is not a conclusive reason for ignoring the principle of general relativity in theoretical investigations of a fundamental character. In other words, I do not believe that it is justifiable to ask: What would physics look like without gravitation?* "

Although Einstein's 1950 piece stops short of suggesting that the actual *equations* of SR might be wrong, the article's thrust – embracing rather than rejecting curvature effects, even within the realm of inertial physics – leads inevitably to the invalidation of special relativity.

It is interesting to speculate on how Einstein might have responded if he had lived another ten years, had had enough time to properly develop these ideas, and had been around for the 1960 crisis. When we found that SR and the GPoR appeared to be incompatible, the community's reaction was to try to save the existing structures and adopt the defensive position of declaring that special relativity was being incapable of being wrong. Einstein's 1950 position, applied to the 1960 situation, would have instead tended to have give the GPoR precedence even if this meant discarding special relativity. [i]

42.3. Summary

The idea that we mustn't question the structure of current physics, because it's so obviously right that to do so betrays one's ignorance, is slightly undermined when we find out that *Einstein himself* seriously queried the structure's validity.

The argument emerging against special relativity in these pages is that SR cannot logically coexist with relativistic gravitation in the same universe – special relativity in isolation is internally consistent, but a "mash-up" of special relativity and the general principle of relativity is not. While Einstein had entirely understandable pragmatic and personal reasons for wanting his general theory to reduce to SR, [ii] [iii] the decision was logically, geometrically and architecturally questionable.

It is to Einstein's great credit that, as the one person who arguably had most incentive to *want* to believe that GR1916 was free from design problems, he was still open-minded enough to be able to look at his own theory critically, and decide that it still appeared (to him) to have unresolved architectural issues [iv] – perhaps special relativity hadn't (yet?) proved itself to be worthy of a place within general relativity.

> Einstein's 1950 article undermines the usual narrative that the structure of Einstein's general theory with respect to SR is perfect, and obvious, and that only a stupid person would query it. Einstein was clearly neither stupid, nor ignorant of how his general theory had been constructed.

There is a German saying that, with laws and with sausages, it is easier to be enthisiastic about the final product if one has not seen the process by which they are made. Some of the decisions that go into attempted derivations of physical law are not just undocumented, they are taken before the theory becomes solidified by language and notation. Since the form of the theory can define what some words will then mean within the context of the theory, some of these decisions will be made intuitively and instinctively, at the prelinguistic level, and may defy later attempts to explain them. Einstein's 1950 querying of the 1916 structure would have been informed by his unique understanding of the decisions (explicit and implicit) that had gone into building his 1916 theory.

---

i The correct scientific approach, of course, would have been to properly explore both options.

ii Once Einstein had made the critical breakthrough (1911 [12]) of identifying that curvature had to be applied to both space and time coordinates, and not just space. Einstein needed to finish the theory quickly, or risk someone else producing a general theory first. Grafting special relativity into the general theory as a limiting case, "as-is", allowed him to finish the theory more quickly (researching the curved-spacetime version of inertial physics might have taken him many more years).

iii Einstein would also have had an understandable emotional attachment and loyalty to special relativity, as the theory was one of the 1905 papers that allowed him to stop self-funding and finally get a paid job in academic research.

iv It's also possible that part of Einstein's original insistence that SR should live on within GR may have been a defensive response to teasing by some of his contemporaries who were not SR enthusiasts, "congratulating" him on his work to get rid of SR and replace it with a better theory built on different principles.

43. SR Argument 43: “SR must be right, because physicists aren’t stupid”

43.1. Humans are well known for doing stupid things

Part of our reason for believing that special relativity *must* be correct is because it is familiar to us, we are invested in it, and the idea of it *not* being right can make us feel insecure and uncomfortable. Physicists are clever people, surely that many clever people could not be wrong? To some, the idea that special relativity might be wrong is professionally and personally insulting, and anyone suggesting such a thing must be stupid, ignorant, delusional or malicious. To those with a strong psychological investment in the theory, criticisms of the theory can be seen as a personal attack deserving of retaliation in kind.

To assess the likelihood of a widespread community mistake, we must turn our powers of scientific analysis away from natural phenomena and onto the physics community itself. Mainstream theoretical physics is supposed to have a perfect track record (section 41). Why would we believe that an apparently flawless system would suddenly go off the rails?

The truth is that the rationality of progress in theoretical physics has always been something of a fiction. While the field is sometimes held up (usually by physicists) as a shining example of logic and rational thought, we can also write an alternative history of theoretical physics in which developments are often governed by chance and dumb luck, by accident and happenstance, and by ambition, conformity, and the less sunny side of human nature, and in which breakthroughs sometimes happen despite the efforts of the community rather than because of them.

43.2. Historical stupidity

- **In the C17th** we had the younger Newton being accused of experimental fraud by some continental experts for his “irreproducable” work validating the theory that light came in a continous range of colours and wavelengths, [193] and setting aside natural philosophy for years in disgust at the behaviour of his community. If he had not been coaxed back into the field by his friends, we would not have had *Opticks* and the *Principia*. The wider community screwed up, badly, and for a while discouraged one of its greatest talents.
- **In the C18th** we had a hundred years of English physicists failing to notice that Newton had accidentally inverted the relationship between the energy and wavelength of light, and mocking those who queried the Newton result. Instead of Huygens’ principle being used to *correct* Newtonian theory, wave theory was derided as nonsense.
- **In the C19th** we *should* have had a revolution in gravitational theory after John Michell’s 1784 publication of the theoretical prediction of gravitational shifts, leading to a general theory in the mid-C19th. But Michell had dutifully cited the offending passage in Opticks where Newton had said that light was attracted to regions of *faster* lightspeed. The “polite veil” drawn over the episode by physicists in the C19th, ostensibly to protect Newton’s reputation (but also their own) required the suppression of Michell’s result, possibly retarding some aspects of gravitational theory by another century.

We also had Lord Kelvin (section 11.8), ridiculing and intimidating Darwinists and dismissing geologists’ calculations that certain geological formations and deposits would have taken billions of years to form.

With hindsight, this was all very, *very* stupid.

Modern physicists will tend to say, "Yes, but these are very old examples, modern physics does not still operate like this"

- **In the C20th** we had Einstein's German lecturers deciding that Einstein was insufficiently respectful of his elders and betters to deserve a place in the research community, giving him such bad references that he found himself unable to get a university position Minkowski's verdict had been "*a lazy dog who will achieve nothing*" and who "*never bothered about mathematics at all*". Weber's view was that Einstein had a personality defect in that he "*wouldn't be told*". Einstein is supposed to have applied to and been either rejected or ignored by every major physics university in Europe, with the result that Einstein's "wonder year" work of 1905 had to be done with Einstein supporting himself financially as a patent clerk, estranged from his wife and child. When presented with the gift of one of the most important analytical minds of the Twentieth Century, European academia essentially said "no thank you".
- The C20th also saw German science in the 1930s attempt to reinstate traditional German research traditions and values via the *Deutsche Physik* movement, which effectively banned most work into quantum mechanics and "Jewish science". *Deutsche Physik* was popular with second-division researchers who hoped to better their own career prospects by "eliminating" the wave of bright young researchers who had entered academia because antisemitism made it difficult for them to get work in law, medicine or German industry. On a more positive note, this damage to German theoretical research may have helped prevent Nazi Germany from developing viable nuclear weapons.

43.3. The Gamow affair

Special relativity is an average: [8] simple signal-timelag effects make approaching bodies look bluer and more stretched, and receding bodies look redder and compacted. For a receding body, assuming a speed of light fixed in the observer's frame gives $E'/E = len'/len = c/(c+v)$, and lightspeed fixed in the emitter's frame gives $E'/E = len'/len = (c-v)/c$. [8] Dividing one prediction by the other tells us how much they disagree by $(1-v^2/c^2)$. Special relativity replaces the two conflicting predictions with their "geometric mean", and SR's averaged, *intermediate* predictions then necessarily differ from both by the square root of this disagreement, which is $\sqrt{1-v^2/c^2}$. An SR-compliant photograph will show bodies to be redder and shorter than we would expect by assuming that $c=c_{OBSERVER}$, but *less* red and *less* short than we would expect assuming $c=c_{EMITTER}$. The SR Lorentz factor does not necessarily describe what one *sees*, it is SR's *modification* of what one *might expect* to see, depending on whether we believe the speed of light to be globally fixed in one frame or another.

These are not difficult concepts, and the mathematics is elementary. However, for a significant part of the Twentieth Century, it seemed that our physics community were unable to do these calculations properly, because the *correct* calculations disagreed with their expectations.

In the popular "Mr Tompkins" books [167] (and another book published in 1958 [168]), physicist **George Gamow** explained special relativity's behaviour by saying that we see an approaching cyclist's dimensions squashed in their direction of motion, a description that is compelling, memorable ... and quite wrong. Gamow's description (with a helpful cartooned illustration) became part of standard SR teaching, and was perpetuated by mathematician **Jacob Bronowski** in the prestigious BBC/Time-Life television series "*The Ascent of Man*" [169] (since corrected), which originally showed a moving observer's view of a passing street of buildings and tramcars, some approaching and some receding, all uniformly Lorentz-contracted.

> Bronowski (1973): " *He sees the other two trams tall and thin, because both are moving at high speeds. One tram looks bluer because it is moving towards him, and the other looks redder because it is moving away; but these are not relativity effects* "

What the "Ascent" team didn't seem to realise was that the standard "Gamow" description had already been shown to be "bad" over a decade earlier. [i] James Terrell had pointed out the mistake, and after repeated rejections had finally managed to get the correction past peer review in 1959 [26], [27] (Terrell's experience with the community was such that he later quit physics). Terrell's piece was followed by a small flurry of papers from other researchers, but as late as 1994, this author was visiting various physics institutions and finding that they still seemed to be teaching the faulty "Gamow" description of SR (and saying that we knew that this "bad" account was correct because "*otherwise particle accelerators wouldn't work*").

The mistake is not entirely Gamow's fault, as if we look at Einstein's 1905 electrodynamics paper, we find him writing (apparently echoing an idea promoted by Lorentz and Poincaré):

> Einstein (1905). §4: " *For v=c all moving objects – viewed from the stationary system – shrivel up into plain figures.* "

The Gamow affair shows that it is possible for good physicists to produce terrible work when they want to force calculations to agree with something that they are convinced (though teaching and social conditioning) *must* be the right answer, [ii] and that a body of researchers are then capable of convincing themselves that they can *get* that wrong answer using legitimate mathematics. We are taught when learning special relativity that the theory is not intuitive and that we must learn to suspend disbelief and concentrate on learning how to calculate the right answers. Unfortunately, it seems that users of special relativity, when taught the *wrong* answers can become adept at abusing mathematics, reflexively, in order to obtain a "desired" faulty result.

The interesting thing about the Gamow episode is that even when a simple mistake was mathematically obvious, counted as user-error, and *harmed the theory being defended*, parts of the SR community were still highly resistant to the correct calculations, to accepting that their authority-figures could have been wrong, or to countenancing the idea that they themselves might have been unwittingly mistaught the subject, without noticing. Special relativity was being taught by people who insisted that the theory was simple, but who did not understand it and *did not understand that they did not understand it*. It appeared to be in nobody's interests for the Gamow misinterpretation to persist – there was no evil cabal of conspiring scientists suppressing a correct version of special relativity and propagating a bad version for some nefarious ends, or for money, or advancement … this was a plain and simple case of widespread self-perpetuating human stupidity, flying in the face of clearly contradicting mathematics, spanning decades. [iii]

---

i The offending episode was digitally remastered before the series was (eventually) released on DVD in 2007, but the mistake is still visible in the 1973 book of the series, on pages 249-251. [169]

ii Prospective physics students are supposed to be able to calculate signal-propagation effects correctly in their teens, as introductory physics … and yet somehow the relativity community seemed to lose this ability when the correct calculations disagreed with how they had been taught SR.
Richard Dawkins has said that for good people to do bad things requires religion. Perhaps for good physicists to do bad physics requires special relativity.

iii In economics, the idea that individuals act in their own selfish interests, and in the interests of their group is sometimes referred to as the *Homo economicus* fallacy – one objection being that, even if individuals *fully intend* to act selfishly in the interests of themselves and their group, they often don't know how to, as they often do not understand their environment, their situation, or the consequences of their actions.
In other words, contrary to economic modelling that assumes that people act efficiently and rationally for maximum advantage, we (individually and collectively) have a tendency do dumb things.

Still, Gamow's characterisation *was* overturned in the 1959, and a note by Roger Penrose to the effect that Gamow's description doesn't correspond literally to SR's physical predictions, *does* appear in the foreword of modern editions:

> Penrose (1992): [170] "*Gamow's descriptions of flattened bicycles and city blocks are intuitively helpful, but they do not represent what an observer would actually see.*"

, so some *might* argue that this counts as a "success story", in which the physics community does (eventually) correct its own mistakes, even if that information has trouble propagating to the educational sector. The Terrell correction *was* over half a century ago, and perhaps nowadays we are more careful (and have a larger scientific population, and the internet to help us).

So here's a more ambitious challenge: can we find another basic, provable, recent or even *current* mistake in modern fundamental theoretical physics, that has never been corrected?

43.4. Time-asymmetry

It is a notable feature of our universe that it is not symmetrical with respect to time – if we require physical theories to correspond reasonably well to reality in order to be considered credible, we will want them to show similarly asymmetrical behaviour. While philosophers seem have been fond of discussing the nature of time for as long as we have *had* philosophers, the physics community's involvement with the fact that time is asymmetrical seems to stem (based on the citation record), from a lecture given by Eddington in 1927, published in 1928. [162] This then seems to have been treated as the basis for pretty much all further work on the subject, and is cited around 300 times. Eddington stated that while special relativity had been criticised for lacking an arrow of time, it was the same as previous theory in this regard, and merely made the existing time-symmetry obvious. [i]

> Eddington (1928), [162] "***Time's Arrow*** *… Objection has sometimes been felt to the relativity theory because its four-dimensional picture of the world seems to overlook the directed character of time. The objection is scarcely logical, for the theory is in this respect no better and no worse than its predecessors, The classical physicist has been using without misgiving a system of laws which do not recognise a directed time; he is shocked that the new picture should expose this so glaringly.* "

Statements of position (apparently inspired by Eddington's assurance), saying that "all classical theory is time-symmetrical" appear scattered throughout subsequent literature:

> Thomas Gold (1958), [172] " *Newton's laws of gravitation and dynamics single out no sense of the time coordinate. If somebody recorded the motion of the planets and reversed the record of the time coordinate, this would leave it an example of a dynamical system that is as much in accord with Newton's laws as the actual. The change from Newton's laws to Einstein's did not affect this symmetry.* "

> Halliwell, Pérez-Mercader, and Zurek (1996), [173] " *Newton's laws, quantum mechanics, electromagnetism, Einstein's theory of gravity, etc., make no distinction between past and future - they are* ***time-symmetric****.*"

If one searches Google for the phrase "*make no distinction between past and future*", one finds a large number of papers and books using identical or almost identical wording. Since it should be impossible to create time-asymmetrical behaviour from time-symmetrical laws, this has left

i It would be interesting to know *who* had already been criticising SR for time-symmetricality by 1927 … Unfortunately, Eddington's piece mentions no names.

theoretical physics with something of a headache, [ii] and some of our best minds have tried and failed to come up with a convincing answer (*see:* section 45) to the resulting paradox.

What none of these researchers seem to have checked is whether Eddington's initial declaration was actually correct.

And in fact, **Eddington got it wrong.** As we can see from section 16.2, the Doppler equations of Newtonian theory (and the equations of any hypothetical relativistic theory other than special relativity) *are lop-sided* with respect to time: they generate an energy-loss in forward time and an energy-gain in reversed time. The Doppler equations only become time-symmetrical under relativity theory with the introduction of the Lorentz-Einstein equations. This is a result that an entire research community seems to have missed, apparently because of a culture in which we do not sufficiently question fundamental statements of fact from our authority-figures. If Eddington stated a thing, confidently, then *obviously* he would have checked whether what he was saying was true … and if it wasn't true, *someone, somewhere, surely would have noticed and pointed it out.*

However, a scan of the abstracts and some of the full papers of the citation-listings suggests that Eddington's statement may never have been fact-checked – an entire branch of (rather unsuccessful) physics theory then appears to have been partly built on a statement that was untested, and is *provably* not true, if we are paying attention.

43.5. Summary

Humans are social animals, and trusting one's colleagues, friends and superiors is generally considered to be an admirable personality-trait. But theoretical physicists do not have this luxury: in the case of time-asymmetry, a body of work covering decades has been potentially compromised, simply because physicists didn't feel the need to question and fact-check the work of other physicists who supposedly knew more than they did. Even researchers fluent in advanced quantum mechanics and string theory have worked on this problem and failed to see a simple mistake in the definitions because they've accepted what they've been told, at face value.

Human enterprises are fallible, humans do stupid things, both individually and as groups, and physicists are human. Given that the physics community has a long history of failing to notice problems that we (with hindsight) consider obvious, it is difficult to argue that the current generation of researchers are somehow immune to the problems that beset our predecessors, and aren't similarly failing to see the whole picture. We may believe that our generation is brighter, more disciplined and smarter than previous generations, and therefore less likely to mess up … but previous generations seem to have believed the same thing, and this seems to have been partly responsible for their mistakes.

Most human societies and groups try to improve their track record over time by learning from their past mistakes – past examples of failure are taught as cautionary tales, as a way of avoiding similar errors in the future. In the physics community, we teach that established physics has never *made* any serious mistakes (no Newtonian crisis, no 1960 GR breakdown), and are therefore more vulnerable to large-scale errors, because we don't believe them to be possible.

> Is it really conceivable that if some important aspect of current theory was wrong, we could have somehow failed to notice it? On the basis of the Gamow and time-asymmetry cases, yes … it seems to be an entirely credible possibility.

ii … except as local statistical fluctuations, which can occur with either orientation. But statistical fluctuations have fluctuations-within-fluctuations, and timeflow in our section of universe appears uniformly forward-oriented.

44. SR Argument 44: Modern cosmology requires Einstein's general theory, and therefore special relativity

44.1. Is GR1916-based cosmology any good?

A number of texts sing the praises of general relativity in how wonderfully it deals with cosmology. We are told that GR1916 is "*the gold standard*", has "*passed all tests with flying colours*", and produces an amazingly good match to cosmological data … once we take into account dark matter, dark energy, and inflation. [i] Some of these statements claims are so gushingly effusive that one may start to suspect that perhaps we are not supposed to peer too closely at the detail.

Let's take a look. While the system of modern cosmology has been based on painstaking empirical work, a consistent foundational *logical structure* for the system does not seem to exist. Attempts to make the system at least nominally compatible with the 1916 theory have left cosmology inconsistent and incoherent, [121] and incompatible with basic geometrical principles. The root cause for this is, yet again, GR1916's attempt to incorporate the flat-spacetime principles of special relativity into what was supposed to be a theory of curved spacetime.

44.2. The Hubble expansion redshift

Observations and analysis of redshift data for galaxies by **Georges Lemaître** (1894-1966) [175] and **Edwin Hubble** (1889-1953) [176] in 1927 and 1929, respectively, indicated an odd pattern. We could use the temperature of the stars' light to work out how bright they ought to be, and then compare this with their actual brightness to work out their distances. We could also work out how much their light has been shifted along the spectrum by looking for known spectral lines in the starlight, and seeing how far these were out of the normal positions. This work told us that, as an averaged trend, there was a redshift in a star's light that increased with its distance. It seemed that the further away a star was, the faster it was receding. Assuming that there was nothing uniquely special about the Earth's location, this suggested that *everything in the universe* was moving away from everything else (the usual analogy being with points marked on the surface of an inflating balloon). The entire universe appeared to be expanding. [ii]

This effect had been missed by Einstein, who had invented an extra term within general relativity, the **cosmological constant**, specifically to explain why the universe was "hovering", neither collapsing under its own gravity or expanding. Einstein used the constant to argue that the universe was *not* changing in size, shortly before Lemaître and Hubble showed that it *was.* Einstein famously referred to this as the biggest blunder of his life – if he'd predicted a size-varying universe, the experimental confirmation just a few years later would have been one of the biggest theoretical successes in the history of science. Einstein eventually accepted the expansion idea (partly because his own solution had also turned out to be unstable [177], [178], [179]), and removed the constant from the theory, but the fact that expanding-universe arguments had never played a part in the *design* of his general theory meant that we were left with some basic definitional conflicts between GR1916 and geometrical cosmology, and some unresolved issues.

---

i Which of course, have all been invented specifically to occupy the gap between what GR1916 predicts and what we actually see. If one is allowed to invent *three* new categories of thing as corrections, then the final agreement *better* be almost perfect.

ii A second interpretation might have been that perhaps the distance-dependent redshift was evidence of an apparent gravitational field increasing with strength with distance from the Earth. This would make sense if the observational time delay was revealing that the universe's gravitational field density had been stronger in the past and was steadily decreasing over time. But this would be equivalent to suggesting that that the universe had been expanding, so it's really the same thing.

44.3. Which shift law? Option "A"

If space in our universe is considered as the three-dimensional equivalent of a spherical surface, and the surface is expanding, then we can imagine a situation in which a line of equidistant stars are each receding from their immediate neighbours at the same speed,

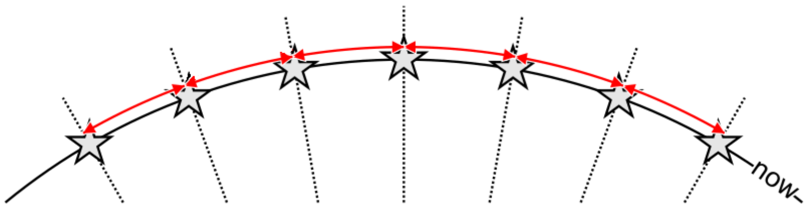

*Figure 9: **A line of mutually-receding stars** in an expanding universe.*

, and each part of the universe is expanding at the same rate.

If each star recedes from it's neighbours at $0.1c$, then a signal can move from Star 1, to Star 2, to Star 3, and lose the same proportion of its energy each time, but no matter now many steps it makes, the signal is never redshifted all the way down to zero frequency.

This exercise suggests that there is *no* cosmological horizon: if the universe "wraps around", and we stood in one place for long enough and used a sufficiently powerful telescope, then in theory we'd be able to see the back of our own head one universe circumference away (Clifford, 1879 [180]) and a second image and a third (*etc.*) at multiples of the distance.

If the expansion effect (a redshift as a function of recession velocity) can be expressed as a Doppler relationship, we will then be tempted to use the recession relationship **A**, from section 25.5,

$$E'/E = c/(c+v)$$

, as this gives the result that the redshift never goes to zero until the recession velocity is infinite.

Unfortunately, this doesn't work. ...

44.4. A finite universe age suggests a horizon

As we look at parts of the universe that are further and further away, we are seeing them as they were in the more distant past. If we look sufficiently far, we might expect to see all the way back to the hypothetical Big Bang event, and then (assuming that the event is "transparent") perhaps even further. But if we are seeing information from before our universe, then technically this counts as part of our universe, too. At this point things become unsatisfactory – however there's another argument that says that we cannot see all the way back even as far as the Big Bang. ...

44.5. The gravitational cosmological horizon

A signal that takes billions of years to reach us will have originated in a noticeably smaller universe, in which the background density of the field would have been significantly higher. The signal has effectively moved from a denser region of spacetime to a more rarefied region and has crossed a density gradient, and is therefore expected to show a gravitational redshift.

If our gravitational equations support horizons, we can then say that the hypothetical initial singularity is censored by a gravitational horizon, and we *do* have a cosmological horizon after all.

Unfortunately, this doesn't work, with **A**, either. ...

44.6. Duality of cosmological and gravitational shifts

Are the gravitational shift and the cosmological shift separate or dual?

- **If they are separate**, then our current cosmology is wrong, because we haven't been taking into account a separate cosmological *gravitational* shift, and have only been calculating the cosmological *expansion* shift.
- **If the two are dual**, then current cosmology could be broadly right, but both classes of effect would need to use the same Doppler shift relationship.

The **strip argument** (Baird, 2019 [121]) says that if we follow the path of a lightpulse (or a consecutive pair of light-pulses) during their cosmological-timescale journey, we can look at just the path marked out in spacetime by the pulses, and a thin surrounding strip or cylinder of spacetime, and (if the interaction of light and spacetime is purely local physics) calculate the final redshift effect from just the curvature properties of the strip, without considering anything else.

But this exercise does not let us judge whether the strip's curvature is "cosmological" or "gravitational", as finding this out involves investigating the wider context of the region *outside* the strip. We therefore only have a single curvature shift effect, that cannot be broken down into separate effects for cosmological curvature and gravitational curvature: we have only a single cause for redshifting – curvature – and therefore also only a single shift relationship.

This rules out Doppler equation **A**: if the cosmological shift *was* c/(c+v), then we'd need the gravitational shift to also be $c/(c+v)$ (so no gravitational horizons), and since the gravitational shift is calculable as a motion shift, $c/(c+v)$ would also have to be the recession Doppler shift relationship for simply-moving matter, which … would allow infinite energy machines (section 25.3), and also *does* seem to be strongly ruled out by SR experimental testing.

At this point, it is probably safe to assume that **A** is not a viable candidate.

44.7. Which shift law? Option "B", Special relativity

Our next option is to ask whether cosmological shifts might obey special relativity. This would agree with GR1916's use of an SR shift relationship for gravity, and with the Schwarzschild metric, [97] but one thing that cosmology texts tend to be quite emphatic about is that cosmological shifts ***do not*** obey the SR shift law, as the nature of a spacetime expansion curvature shift has different characteristics to an SR recession redshift in flat spacetime.

An advantage of the SR relationship is that it nominally generates a horizon. This horizon is required to be absolute, suggesting that, when distant objects are rushing away from us at the speed of light, the horizon remains stationary for us, and sweeps though those objects at local lightspeed. This would make cosmological horizons "absolute", as no signal or body further away from us than the horizon could cross it in our direction without travelling at more than local lightspeed.

… Unfortunately, if we revisit the "line of stars" argument and Figure Error: Reference source not found, we find that for a signal sent between any chain of stars, the shift is again still finite – we do not have a physical horizon, and since the SR shift is supposed to be a property of undisturbed spacetime, we should be able to take away the line of stars and still get the same answer. We then have the (unwanted!) result that there is no *actual* cosmological horizon in an SR-based model, as the SR velocity addition law puts the effective lightspeed horizon at what, without the SR velocity addition law, would be v=infinity. This makes the SR equations seem unworkable.

This should not be a surprise: if there was any easy way to integrate the SR relationships into cosmological shifts, the community would have done this many years ago.

Associating the SR Doppler law for motion shifts with cosmological curvature would also mean associating relative velocity with curvature, which tends to break special relativity.

It would seem that option **B** (special relativity) is also not a viable candidate.

44.8. Which shift law? Option "C", Newtonian relationships

Our third option ("**C**") from section 25.5 is the Newtonian Doppler relationship, $E'/E=(c-v)/c$.

The application of this third relationship is interesting: the "**C**" law gives us a cosmological horizon when there is no intervening matter, but also supports the "chain of stars" argument that says that we should be able to see arbitrarily far using masses as relay stations – it manages to reconcile the apparently irreconcilable (solving the **cosmological shift paradox**) by saying that we can dissolve a horizon by applying an NM velocity-addition law (section 19.6) … but only when there is actual physical intervening matter in the signal path, to provide a *physical* justification for dividing up a velocity into smaller stages.

The cosmological horizon then becomes an *acoustic* horizon, the cosmological background radiation becomes the counterpart of acoustic Hawking radiation, and physics spanning a cosmological horizon becomes described by "acoustic" physics, described using an acoustic metric. [i]

Since the only relativistic equations that support classical Hawking radiation are the $(c-v)/c$ set *exactly*, [121] if we want this acoustic solution, we have to use *exactly* this redder Doppler equation.

44.9. Full shift equivalence

It would seem that if we want to have a consistent expanding-universe model in a relativistic universe, we must use the Newtonian Doppler equations for Hubble shifts across regions of empty space, with a (deterministic, derived) corrections to express the modification of the expected shifts according to the light-dragging properties of any matter alongside the signal path.

Our "strip" argument then requires that the same non-SR equations have to apply to gravitational shifts, and if gravitational shifts obey this relationship, then motion shifts need to as well. [ii]

We then have a model in which all three classes of shift must obey the same Doppler equations, and must in fact be interchangeable, giving a pleasing "compaction" of physical law. [iii] Current theory already lets us calculate gravitational shifts as velocity-shifts, but under GR1916 the relationship is one-way. In an acoustic physics, bodies with relative velocity v are associated with

---

i Although we've said elsewhere that acoustic metric do not require a particulate medium or metric, they have obvious advantages when we *do* have a particulate medium, which is the case in cosmology, where we have billions of stars whose gravitational fields cannot be reasonably ignored, and where flat spacetime is not a reasonable suggestion.

ii We can, of course, also apply a topological inversion to our view of the universe, to turn it into a different description in which we are at the periphery looking inward, and the cosmological horizon is in the centre, facign outward, and censoring a central Big Bang singularity. This remapping turns the cosmological horizon into a conventional grvaitaitonal horizon, or turns a black hole horizon into a cosmological horizon. The two types of horizon then need to obey identical laws. We are not allowed to override the geometry and say "but we know the two things are different!" – if the exercise doesn't work, it means that our current theory can't cope with topological remappings, and is not a proper geoemtrical theory.

iii A theory's ability to merge *two* physical effects that were previously considered to be separate into one is usually considered impressive. In this case we are compacting *three* different effects, in other words, not just achieving *duality* but *triality*.

a gravitomagnetic curvature with an equivalent velocity-differential *also* equal to v, so we can calculate the same shift *either* in the time domain as a motion shift, *or* in the curvature domain as a gravitomagnetic shift, based on the curvatures that appear frozen into the region's spacetime in a Δt=0 "snapshot" of the geometry. Shift equivalence also solves Zeno's paradox against motion: relative motion doesn't "disappear" in a frozen image of a situation, it remains frozen into the image as equivalent gravitomagnetic curvature.

In the new scheme, if we have a distant redshifted galaxy moving at high speed with respect to its neighbours, we no longer have to worry about breaking its apparent recession rate into different components for cosmological recession, gravitomagnetic effects, "real" relative recession, the variation in cosmological curvature due to the huge moving mass, and so on. We can ignore the different possible interpretations of *why* the galaxy is redshifted, and instead, just convert the visible redshift into an equivalent velocity value, and interpret this value any way we like.

This trick doesn't work with SR-based physics, as associating curvature with relative velocity gives us gravitomagnetism, which breaks the SR/Minkowski assumption that lightbeam geometry remains unchanged regardless of the relative velocities of bodies.

44.10. Summary

It is an unfortunate accident of history that cosmological redshifts were not discovered until just a few years after Einstein has already finalised the architecture of his general theory of relativity, as, if we were to decide to create a general theory from scratch in 1930, with the benefit of knowing about Hubble shifts, the result would not have been Einstein's general theory, and would not have included the physics of special relativity.

If we include an expanding-universe cosmology in general relativity at the design stage, and treat it as a further class of effect that needs to be absorbed into relativistic theory, along with inertial physics, acceleration, rotation and gravity, then the additional constraint restricts us to only one possible set of equations of motion for moving bodies – which are not those of special relativity.

> Modern cosmology is not a triumph of GR1916, it is crippled by it. The SR content of the 1916 theory prevents cosmology from being able to apply proper geometrical arguments and rules, and leaves it a mess of inconsistent arguments, in which we cannot even safely derive a straightforward equation for the energy-change of light for a given distant receding galaxy without worrying about how the visible shift ought to be divided up into different components, obeying different rules.

> Cosmology would be better off (and more consistent, and would be a more powerful predictive system) if we replaced Einstein's general theory with a *proper* general theory, based on an acoustic metric.

–

The cosmology of an expanding hyperspherical universe is not a flat-spacetime problem, and is not correctly described by the SR equations. The classical physics we get from QM is also not a flat-spacetime system, and is not correctly described by the SR equations. Further, the classical physics of moving gravitational masses is not a flat-spacetime problem, and is not correctly described by the SR equations (and *all* masses are gravitational masses). The specific divergences from special relativity demanded by all three of these exercises, gravitomagnetic, quantum-gravitational, and cosmological, *are precisely the same.*

45. SR Argument 45: Total energy conservation – the one valid theoretical argument in favour of special relativity

45.1. Energy conservation

There is one (and apparently *only* one) legitimate argument for special relativity being the correct theory of relativity, and for our wanting to incorporate SR into a general theory of relativity:

Special relativity is the only relativistic system that offers total energy conservation.

Almost everything that has been written in papers and textbooks over the last century about special relativity's chances of being right can be thrown away as irrelevant and replaced with one simple observation: if the principle of relativity is correct, and total energy conservation is correct, then special relativity is inescapable.

If we are serious about wanting to know whether or not SR is correct, we have to consider how physical law changes when we depart from traditional energy conservation.

45.2. Energy-conservation under special relativity

Consider the case of a signal sent between two opposing walls of a laboratory, via a transponder moving directly along the signal beam (in either direction), with velocity v :

Under special relativity, the signal will undergo two successive Doppler shifts, one where v is positive, and one where v has the same magnitude but is negative. The result is:

$$\mathbf{E'/E} = \sqrt{\frac{c-v}{c+v}} \times \sqrt{\frac{c-(-v)}{c+(-v)}}$$

$$= \sqrt{\frac{c-v}{c+v}} \times \sqrt{\frac{c+v}{c-v}} = \mathbf{1}$$

Changing the sign of the velocity value in SR's Doppler equation (swapping plusses and minusses) inverts the relationship, exactly. The signal arrives at the far side of the laboratory with precisely the same final energy, regardless of the presence of the moving transponder in the signal path. We can add as many intermediate masses as we like, and the SR result will still be the same – the presence of any intermediate stages has zero effect on the final predictions.

This is not true for any other system of relativistic physics.

If we accept that the principle of relativity *itself* is correct, then we can argue that all potential relativistic solutions are separated by Lorentzlike factors. For any competing relativistic theory, a given velocity gives the same Lorentzlike divergence from SR regardless of direction, so if a different relativistic system diverges from SR with a Lorentzlike deviation to the *blue*, then both of the component shifts in the above calculation will have an additional blueshift, and the result will be greater than unity (energy-gain). If a relativistic system diverges from SR with a Lorentzlike deviation to the *red*, then both components in the above example will have an additional *red*shift, and the final result will be less than unity (energy-loss).

If we insist that the energy out must precisely equal the energy in, then we are forced to use the SR Doppler equation, which (via the relativistic ellipse exercise in section 4.3) then gives us the rest of SR's geometry. [24] If we repeat the "ellipse" exercise multiple times with different velocities, and intersect the results, putting the ellipses on suitably tilted planes, the intersecting ellipses will reconstruct a Minkowski lightcone and Minkowski spacetime.

45.3. Energy-loss under non-SR theories

If we now repeat the exercise with the "Newtonian" Doppler relationships (representing the maximum possible deviation from SR, to the red), we obtain;

$$\mathbf{E'/E} = \frac{c-v}{c} \times \frac{c-(-v)}{c} = 1-v^2/c^2$$

A little counter-intuitively, it turns out to be the *Newtonian* relationships that give a Lorentz-squared redshift after two frame transitions. [i] For other theories intermediate to SR and NM, we expect an intermediate Lorentzlike energy-loss, in the range E'/E=1 to E'/E="Lorentz-squared".

45.4. The energy-loss question

A proper analysis and discussion of whether special relativity is *correct* relativity theory can therefore discard almost every metaphysical aspect of SR – coordinate systems, Minkowski spacetime, reinterpreted lengths and times – and ask just one fundamental question: might it be reasonable to get less energy out of a complex system than we put in?

45.5. Implications of a universe with energy-loss

Thermodynamics

If we are constantly bleeding energy out of a system, then the system's attempt to reach equilibrium will result in a bias towards exothermic reactions (those that give out energy) at the expense of endothermic ones (those that take it in).

Arrow of time

The bias towards exothermicity, under the condition of time-reversal, turns into a bias towards *endo*thermicity (just as the lossy range (**B, C**] in section 25.5 converts under time reversal to the range (**B, A**]. Physics then looks different in forward and reversed time.

Hubble redshifts

If a "lossy" set of equations are applied to gravitational shifts, then, if light is sent through a gravity-well (and if the blueshift and redshift are calculated from opposing velocities of the same magnitude), we get a round-trip Lorentzlike energy-loss
(if $v=c$, we have a gravitational horizon, and the energy-loss is total [ii]).

If we now send energy across cosmological distances, then in a universe in which matter is roughly evenly distributed, we should expect to see an avergae energy-loss, roughly as a function of distance.

Universe expansion

Lossy systems need to be gravitomagnetic, and gravitomagnetic theories need to associate the relative velocities of physical masses with relative curvature, with "velocity-shift" and "curvature-shift" descriptions then being interchangeable. Under a gravitomagnetic system, if we can't differentiate between a velocity shift and a curvature shift, then if a distance-dependent redshift makes it *look* as if the universe is expanding, that universe *is*

---

i This is also apparent from a slight modification of the NM velocity addition formula from section 19.6 to cope with different-sign velocities. The result of "adding" two equal and opposite velocities $+v$, $-v$ is to create an equivalent recession velocity of $v_3 = v^2/c$. Applying the recession shift formula then gives E'/E = (c - v_3)/c, or 1 - v^2/c^2 .

ii This suggests that if we aim a signal through a transparent mass that undergoes gravitational collapse, SR-based gravitation describes the signal energy as unchanged until a horizon forms (sudden cutoff), while in an NM-based description, the energy of the emerging beam fades smoothly towards zero during the collapse.

expanding.

45.6. A quick scoresheet

If physics is "lossless", i we have no thermodynamic arrow of time, physics looks the same backwards as forwards, and we will be encouraged to use a similar cosmology to the one originally presented by Einstein for general relativity, [90] in which the universe is immortal and looks the same at all times.

If physics is "lossy", we have a thermodynamic arrow of time, physics looks different forwards and backwards, and we can predict Hubble redshifts and an expanding universe – a prediction that Einstein famously failed to make on the basis of GR until after Hubble had already presented the effect.

Given these two options, it would seem that a "lossy" universe is more in agreement with the world we see around us. We can also note that the principle of equivalence and general principle of relativity both require *some* form of deviation from special relativity, suggesting energy-loss.

45.7. The lossy universe

"Lossiness" is already a feature of expanding-universe cosmologies due to Hubble redshift (so total energy conservation is already known *not* to be a law of physics). Cosmologists say that this violation of energy conservation should not be considered important, because conservation assumes a constant environment, and we know (now) that the universe's properties are not constant over time. However, just because we can explain why the universe *should* lose energy, this does not mean that we have answered the question, "*where does the energy **go**?*". Cosmological energy-loss is associated with an increase in total distance: if we wanted to restore conservation laws, we could hypothesise that this is not a coincidence, and describe the missing energy as having (in some respect) been converted into space (so that "mass-energy" conservation is replaced with some larger law of "mass-energy-space" conservation). ii In such a system, energy-losses are not an incidental consequence of expansion, they also *drive* expansion.

Once we have such a law, we can apply it to the case of energy-losses at smaller scales, to again explain "*where does the energy go?*". Energy that disappears in complex systems would either be bound up in the curvature associated with that system, or be radiated into less-dense regions containing fewer exothermic processes, to drive the expansion process (section 30.9).

45.8. Cosmological arrow

Since an "arrow of time" is a desirable feature, there have been some attempts to argue that the cosmological arrow of time and the *cosmological* bias towards exothermicity due to Hubble shift somehow creates a similar bias for smaller-scale physics, given that stars and galaxies will always be radiating more energy into the void than they get back. Stars will therefore be encouraged to favour internal exothermic reactions over endothermic ones

This optimistic "trickle down" approach does not seem to work: if we pump more energy *into* a

i Simon Newcomb (1893): [92] " *We must also remember that the discovery of what could **not** be done has been an important element in progress. We are met at every step by the iron law of the conservation of energy.* "

ii Incorporating space into conservation laws allows a compactification of physical law similar to the compactification that happened when $E=mc^2$ merged the separate conservation laws of mass and energy. This compression of physical law is not possible in an SR-based system.
Similarly, the compaction of principles that results from a universal shift equivalence principle is not available in SR-based physics, because it has to associate motion shifts with curvature, contradicting SR's position that inertial physics can be treated as a flat-spacetime process.

system, we do not seem to be able to persuade time in that region to run in reverse. The Earth, by receiving energy from the Sun, is able to support systems that build up complexity, which *is* thematically associated with the opposite of decay … but, regardless, time on Earth refuses to run backwards with respect to the rest of the universe (or we would see the universe to be contracting).

It is more realistic to assume that time-asymmetry is not imposed from above, but is already built into particle-level physics, [i] with cosmological time-asymmetry then being the larger-scale bulk consequence. It is less credible to suggest that an atom can somehow sense outside its immediate environs in order to be able to decide whether the larger universe is expanding or contracting (and how quickly), so that it can adjust its behaviour accordingly. [ii]

45.9. Reinstating conservation laws

What we do not have (yet) is a specification for how energy converts to space (or to some other quantity associated with space), along the lines of Einstein's $E=mc^2$ equation.

There will be ways of deriving candidate equations for this law, for instance, relating apparent length-changes to energy-changes when we move through our environment at speed.

It will be important to tread carefully when deriving this equation: given the unfamiliarity of the context we have to be sure (for instance) that we are not deriving a conversion factor that explains why conventional energy is *not* conserved, while still inadvertently using elements of special relativity, which assumes that conventional energy *is* conserved.

45.10. Summary

Special relativity is the principle of relativity, plus total traditional energy conservation.

Individual conservation laws are not sacrosanct, and have a habit of being replaced by more general laws. In times past it may have seemed absurd that atoms (or rather, fundamental particles) could be created or destroyed. Being discrete, the total number of these particles in the universe was a definite aspect of the universe that one could (in theory) assign a integer value to – why would one want to sacrifice such a tidy feature? Nevertheless, $E=mc^2$, replacing the old separate conservation laws of mass and energy with a single law was a big step forward.

> If we want to show confidence that special relativity *really is* correct, we are obliged to carry out some sort of study of how physics changes if traditional energy conservation is *not* correct, and massenergy is constantly being lost or converted to some other property, even if we consider such a thing desperately unlikely. If the only relativistic alternative to SR is a "lossy" universe, then this possibility has to be studied, if only to eliminate it.

> Since NM-based physics is time-asymmetrical and includes a small-scale "arrow of time" missing from current theory, it is also worth studying this non-SR physics – even if we still believe in SR – as a way of obtaining wider contextual view of the "arrow" problem.

---

i The subject of time-symmetry was damaged by Eddington's famous 1928 work, [162] which set the tone for subsequent work by stating, incorrectly, that classical physical laws before SR had no arrow of time.
If we want to do serious work in "deep" theoretical physics, it is a better rule to trust nothing and nobody. As the motto of the Royal Society has it, "*Nullius in verba*" ("*take nobody's word for it*").

ii While we could hypothesise an additional field parameter that lets an atom "taste" whether the universe is expanding or contracting (similarly to how we could explain matter-antimatter asymmetry with an additional chirality field), the begetting of additional independent parameters is to be avoided where possible.

46. Conclusions

We are standing on the threshold of a major revolution in theoretical physics. However, we have now been timidly standing on this same threshold for more than half a century, unwilling to take the next step, or to even to *explore the possibility* of taking the next step, since it involves cutting ourselves free of special relativity, a theory that we have convinced ourselves *cannot possibly* be wrong.

During the special theory's "run" of over a hundred years, our conceptual vocabulary and our ideas about which properties a physical theory ought to support have advanced. We no longer believe that global lightspeed constancy or traditional energy conservation are laws of Nature, we understand more about spacetime curvature, and we have realised that special relativity conflicts with a surprising range of basic principles and theories, showing deep incompatibilities with the principle of equivalence, the general principle of relativity, gravitation, gravitomagnetism, quantum mechanics, and current cosmology. The 1905 theory, based on unrealistic and impossible idealisations, is the rogue element of modern theory that prevents all the other components from working together. During this time, the community appears to have produced no comparative review or study of the wider and more general characteristics of relativity theory, making our opinions neither informed or scientific, but *theological*. Without knowing the surrounding logical landscape of other potential solutions, we are effectively trusting in blind faith or good luck that the theory we have just happens to be the right one.

A review of the evidence for special relativity shows that many of the supporting arguments are based on comparisons that are at best highly selective, and at worst, simply wrong. Community mechanisms that are supposed to identify and correct errors have not worked properly. Theoretical studies have been compromised by a need to correspond with what we believe to be the overwhelming experimental evidence, while experimental testing has in turn been profoundly compromised by bad theory (such as the misbelief that transverse redshifts are somehow unique to SR). Neither the theoretical or experimental communities seem to have carried out basic background checks, and researchers often seem not to have realised that the "classical theory" predictions that SR is commonly compared against are not those of Newtonian mechanics.

Meanwhile, the principle of equivalence and the principle of relativity both logically require that there be at least *some* form of velocity-dependent deviation from SR when real matter is involved, and for the resulting physics to still obey the principle of relativity, the necessary deviation must be Lorentzlike. For over a century we have studiously avoided the question of how large this Lorentzlike deviation must be, invoking a compartmentalisation between gravitational and non-gravitational physics … a compartmentalisation that is forbidden by the principle of relativity. Our hope that this deviation be vanishingly small is undermined by the relativity principle's requirement that the "extremal" deviation that we calculate for a moving black hole must then apply to all other moving matter. This maximum deviation turns out also to be the same solution required for compatibility with quantum mechanics, gravitomagnetism, and cosmology. Relativistic gravitation requires not just that special relativity be wrong, but that it be *significantly* wrong.

It is easy to look back at the physics of the C17th and be incredulous that so many experimenters reported being unable to replicate Newton's result that light was composed of an continuous range of distinct colours, apparently due to the strength of their conviction that this could not be true. [166] It is also difficult to avoid incredulity at the ~300 experiments with N-rays that were published in the 1900s, concerning an effect that experimenters *could* see, but which did *not* exist. [111]

It may be that future generations may have similarly incredulous reactions to SR experimentation in the Twentieth Century, in which experimenters using bad logic, bad test theory and bad historical information, failed to notice a deviation from special relativity.

References

1. Albert Einstein, "*Zur Elektrodynamik bewegter Körper*", **Annalen der Physik**, **vol. 322**, issue 10 (1905), pages 891-921. https://doi.org/10.1002/andp.19053221004
   translated and reprinted as:
   Albert Einstein, "*On the Electrodynamics of Moving Bodies*", **The Principle of Relativity** (Methuen 1923) pages 35-65. ISBN 0486600815
2. Albert Einstein, "Die Grundlage der allgemeinen Relativitätstheorie", Annalen der Physik vol. 49 no. 7 (1916), pages 50- . https://doi.org/10.1002/andp.19163540702
   translated and reprinted as:
   Albert Einstein, "*The Foundation of the General Theory of Relativity*", **The Principle of Relativity** (Methuen 1923) pages 109-164. ISBN 0486600815
3. Clifford M. Will, **Was Einstein Right? Putting General Relativity to the Test** (Basic Books, 1986), Appendix "*Special Relativity: Beyond a Shadow of a Doubt*". ISBN 0465090885
4. Clifford M. Will, **Was Einstein Right? Putting General Relativity to the Test, second edition** (Basic Books, 1993). ISBN 0465090869
5. Eric Baird, **Relativity in Curved Spacetime** (Chocolate Tree, 2007), sections "*5: The Newtonian Catastrophe*" pages 51-60. ISBN 0955706807
6. Eric Baird, "*Newton's aether model*" (2000) arXiv:physics/0011003 https://arxiv.org/abs/physics/0011003
7. H.A. Lorentz, "*Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern*" ("*Attempt of a Theory of Electrical and Optical Phenomena in Moving Bodies*"), Leiden: E.J. Brill, Leiden (1895).
8. Eric Baird, "*Special Relativity Considered as an Average of Earlier Theories*" (17 February 2020). https://doi.org/10.13140/RG.2.2.35879.70566
9. H.A. Lorentz "*Electromagnetic phenomena in a system moving with any velocity smaller than that of light*" **Proceedings of the Acadademy of Sciences of Amsterdam**, **vol. 6** (1904), pages 809-831.
10. Hermann Minkowski, "*Raum und Zeit*", **Jahresbericht der Deutschen Mathematiker-Vereinigung vol. 18** (1909), pages 75–88.
    translated and reprinted as:
    Hermann Minkowski, "*Space and Time*", **The Principle of Relativity** (Methuen 1923), pages 73-91. ISBN 0486600815
11. John Michell, "*On the means of discovering the distance, magnitude, &c. of the fixed stars, in consequence of the diminution of the velocity of their light …*", **Philosophical Transactions of the Royal Society of London vol. 74** (1784), pages 35-57. https://doi.org/10.1098/rstl.1784.0008
12. Albert Einstein, "*Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes*", Annalen der Physik vol. 35 (1911), pages 898-908. https://doi.org/10.1002/andp.19053221004
    translated and reprinted as:
    Albert Einstein, "*On The influence of gravitation on the propagation of light*", **The Principle of Relativity** (Methuen 1923) pages 97-108. ISBN 0486600815
13. Albert Einstein, "*Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?*" **Annalen der Physik vol. 18**, no. 13 (1905), pages 639-641. https://doi.org/10.1002/andp.19053231314
    translated and reprinted as:
    Albert Einstein, "*Does the Inertia of a Body Depend Upon Its Energy Content?*", **The Principle of Relativity** (Methuen 1923), pages 67-71. ISBN 0486600815
14. Eric Baird, "*Two exact derivations of the mass/energy relationship, $E=mc^2$*" (2000/2020) arXiv: physics/0009062v2
15. Isaac Newton, **Opticks: Or, a Treatise of the Reflections, Refractions, Inflections and Colours of Light**, **fourth edition** (1730)
16. T.R. Sandin, "*In Defense of Relativistic Mass*", **American Journal of Physics vol. 59**, no. 11 (1 November 1991) pages 1032-1036. https://doi.org/10.1119/1.16642
17. P.A. Cerenkov, "*Visible Emission of Clean Liquids by Action of Gamma Radiation*", **Doklady Akademii Nauk SSSR (DAN SSSR) vol. 2** (1934), page 451.
18. Oliver Heaviside, "*Electromagnetic waves, the propagation of potential, and the electromagnetic effects of a moving charge*", **The Electrician** (various issues, 1888-1889), *and in*
    Electrical Papers vol. 2 (Macmillan, 1894).

19. Albert Einstein, "*On the Principle of Relativity*", **Vossische Zeitung**, 26 April 1914, Morgen-Ausgabe, no.209, pp.[1-2] (26 April 1914) , translated and republished in, **The Collected Papers of Albert Einstein: Volume 6: The Berlin Years: Writings, 1914-1917 (English translation supplement)** (Princeton University Press, 1997) pages 3-5. ISBN 9780691017341 https://einsteinpapers.press.princeton.edu/vol6-trans/15
20. S. W. Hawking, "*Black hole explosions?*" Nature vol. **248**, pages 30–31 (1st March 1974) https://doi.org/10.1038/248030a0
21. S. W. Hawking, "*Particle creation by black holes*" **Communications in Mathematical Physics**, **vol. 43**, issue 3 (1975), pages 199-220 . https://doi.org/10.1007/BF02345020
22. Kip S. Thorne, **Black Holes and Time Warps: Einstein's outrageous legacy**, (Papermac, 1995) ISBN 0333639693
23. Eric Baird, "*The Doppler equations of Newtonian optics as a unique solution to quantum gravity*" (May 2018). https://doi.org/10.13140/RG.2.2.21885.67045
24. William Moreau, "*Wave Front Relativity*", **American Journal of Physics vol. 62**, no. 5 (1 May 1994), pages 426-429. https://doi.org/10.1119/1.17543
25. Eric Baird, **Relativity in Curved Spacetime** (Chocolate Tree, 2007), section 8: "*Aberration of Angles / The Relativistic Ellipse*", pages 77-84. ISBN 0955706807
26. James Terrell, "*Invisibility of the Lorentz Contraction*", **Physical Review vol. 116**, no. 4 (15 November 1959), pages 1041-1045. https://doi.org/10.1103/PhysRev.116.1041
27. James Terrell, "*The Terrell Effect*", **American Journal of Physics vol. 57**, no. 1 (1 January 1989), pages 9-10. https://doi.org/10.1119/1.16131
28. Roger Penrose, "*The Apparent Shape of a Relativistically Moving Sphere*", **Mathematical Proceedings of the Cambridge Philosophical Society**, **vol. 55**, no. 1 (January 1959) pages 137-139 https://doi.org/10.1017/S0305004100033776
29. Roy Weinstein, "*Observation of Length by a Single Observer*", **American Journal of Physics vol. 28**, no. 7 (1 October 1960), pages 607-610. https://doi.org/10.1119/1.1935916
30. Mary L. Boas, "*Apparent Shape of Large Objects at Relativistic Speeds*", **American Journal of Physics**, **vol. 29**, no. 5 (1 May 1961), pages 283-286. https://doi.org/10.1119/1.1937751
31. Robert Resnick, **Introduction to Special Relativity**, (Wiley, 1968), "*1.7 ...the Aether-Drag Hypothesis*" pages 28-33, "*2.7 … Aberration and Doppler Effect …*", "*3.4 Alternative Views of Mass in Relativity*". ISBN 047171724X
32. Ray d'Inverno, **Introducing Einstein's Relativity** (OUP 1992). ISBN 9780198596868
33. Richard A. Mould, **Basic Relativity** (Springer-Verlag, 1994). ISBN 9780387952109
34. Wolfgang Rindler, **Special Relativity: Special, General and Cosmological, second edition** (OUP, 2006). ISBN ISBN 0198567316
35. Oliver Lodge, **The Ether of Space** (Harper, 1909), "*Grating theory*" pages 134-136.
36. Eric Baird, **Relativity in Curved Spacetime** (Chocolate Tree, 2007), Calculations: 3: "*Aberration Redshift*" pages 362-353. ISBN 0955706807
37. Eric Baird, "*Special relativity without time dilation or length contraction*" (February 2020) https://doi.org/10.13140/RG.2.2.14036.55688
38. Edwin F. Taylor, and John Archibald Wheeler, **Spacetime Physics, 2nd Edition** (W.H. Freeman, 1992), Box 3.4, "*Does a moving clock **really** 'run slow'?*" pages 76-77 . ISBN 0716723271 http://www.eftaylor.com/spacetimephysics/03_chapter3.pdf
39. Albert Einstein, **The Collected Papers of Albert Einstein: Volume 7: The Berlin Years: Writings, 1918-1921 (English translation supplement)** page 141 *circa* ~1920- "*31: Fundamental Ideas and Methods of the Theory of Relativity, Presented in their Development*" 18: "*Influence of Gravitational Fields upon the Rate of Clocks. Redshift of Spectral Lines*" https://einsteinpapers.press.princeton.edu/vol7-trans/157
40. Albert Einstein, May 1921 "Stafford Little" Princeton lectures, published as **The Meaning of Relativity** (Meuthen, 1922), fourth section, "*The General Theory of Relativity (continued)*".
41. Cranshaw, T. E., J. P. Schiffer, and A. B. Whitehead. "*Measurement of the Gravitational Red Shift Using the Mössbauer Effect in* Fe^{57}" **Physical Review Letters** vol. **4**, no. 4 (15 February 1960): 163–64. https://doi.org/10.1103/PhysRevLett.4.163

42. Hay, H. J., J. P. Schiffer, T. E. Cranshaw, and P. A. Egelstaff. “*Measurement of the Red Shift in an Accelerated System Using the Mössbauer Effect in Fe^{57}*”, P**hysical Review Letters** vol. **4**, no. 4 (15 February 1960): pages 165–66. https://doi.org/10.1103/PhysRevLett.4.165
43. R.V. Pound, and G.A. Rebka, “*Gravitational Red-Shift in Nuclear Resonance*” **Physical Review Letters**, **vol. 3**, no. 9 (1 November 1959), pages 439-441. https://doi.org/10.1103/PhysRevLett.3.439
44. R.V. Pound, and J.L. Snider, “*Effect of Gravity on Nuclear Resonance*”, **Physical Review Letters**, **vol. 13**, no. 18 (2 November 1964), pages 539-540. https://doi.org/10.1103/PhysRevLett.13.539
45. Robert V. Pound, “*Weighing Photons, I*”, **Physics in Perspective**, **vol. 2**, no. 3 (September 2000), pages 224-268. https://doi.org/10.1007/s000160050045
46. Alfred Schild, “*Equivalence Principle and Red-Shift Measurements*” **American Journal of Physics** vol. **28**, 778 (1960), https://doi.org/10.1119/1.1936000
47. C. Møller, **The Theory of Relativity** (Oxford, 1955) *esp.* “*VII The Foundations of the General Theory of Relativity / 83. The principle of equivalence*”, and “*X: The Influence of Gravitational Fields on Physical Phenomena / 110. Momentum and mass of a particle. Gravitational force*”.
48. John Archibald Wheeler, **A Journey into Gravity and Spacetime** (Scientific American Library, 1999), “7: The Boundary of a Boundary” pages 108-121; “*Gravity’s Next Prize: Gravitomagnetism*” pages 232-233. ISBN 0716760347
49. C.W.F. Everitt, D.B. DeBra, B.W. Parkinson, J.P. Turneaure, J.W. Conklin, M.I. Heifetz, G.M. Keiser, et al., “‘*Gravity Probe B: Final Results of a Space Experiment to Test General Relativity*”, **Physical Review Letters vol. 106** (22): 221101 (2011). http://arxiv.org/abs/1105.3456 https://doi.org/10.1103/PhysRevLett.106.221101
50. Einstein, **The Times** newspaper (London) (November 28, 1919)
51. Karl R.Popper, Philosophy of Science: A personal report, in
In J. H. Muirhead (ed.), British Philosophy in the Mid-Century, (George Allen and Unwin 1957). pages 182-183 (1957)
reprinted in
Karl R. Popper, Conjectures and Refutations
52. Max von Laue, **Das Relativitätsprinzip** (Vieweg, 1913).
53. Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler (“MTW”), **Gravitation** (Freeman, 1973). ISBN 0716703440
54. J.K. Bailey, F. Borer, H. Combley, F. Drumm, F. Krienen, E. Lange, Picasso, et al., “*Measurements of Relativistic Time Dilatation for Positive and Negative Muons in a Circular Orbit*” **Nature**, **vol.268**, no. 5618 (July 1977), pages 301–305 . https://doi.org/10.1038/268301a0
55. Pierre Millette, “*On the Question of Acceleration in Special Relativity*”, **Progress in Physics** ISSN 1555-5534 **vol.13** (28 September 2017), pages 213–217 .
56. C. O. Alley, "*Relativity and Clocks*", Proceedings, 33rd Annual Symposium on Frequency Control, U.S. Army Electronics Research and Development Command, Fort Monmouth, N.J., pp 4 - 39A (1979).
57. Alex Harvey, and Engelbert Schucking, “*A Small Puzzle from 1905*”, **Physics Today vol. 58**, no. 3, pages 34-36 (12 January 2007). https://doi.org/10.1063/1.1897562
58. Edmund Bertschinger, and Edwin F. Taylor, “*Gravitational waves*” (2017), chapter 16 of **Exploring Black Holes, Second Edition** http://www.eftaylor.com/exploringblackholes/Ch16GravWaves171018v1.pdf
59. Albert Einstein, “*On the principle of relativity*”/“*Vom Relativatäts-Prinzip*" **Vossische Zeitung** (26th April 1914).
translated and reprinted in, **The Collected Papers of Albert Einstein, Volume 6: The Berlin Years, Writings, 1914-1916** (Princeton, 1997), pages 3-5. ISBN 0691017344
60. Isaac Newton, **Philosophiæ Naturalis Principia Mathematica** (1687), **Principia** translation from Latin to English (1728).
61. Arthur S. Eddington. **Report On The Relativity Theory Of Gravitation: The Physical Society Of London**, (Fleetway Press, 1920). http://archive.org/details/reportontherelat028829mbp
62. Peter Coles, “*Einstein, Eddington and the 1919 Eclipse*”, in **Historical Development of Modern Cosmology, ASP Conference Proceedings vol. 252** (2001) “*Historical Development of Modern Cosmology*”, pages 21-. ISBN 1583810927
63. Johann Georg von Soldner, “*Ueber die Ablenkung eines Lichtstrals von seiner geradlinigen Bewegung*”, Berliner Astronomisches Jahrbuch, 1804, pp. 161-172
“*On the Deflection of a Light Ray from its Rectilinear Motion*” https://en.wikisource.org/?curid=755966

64. Clifford M. Will, "*Henry Cavendish, Johann von Soldner, and the Deflection of Light*", **American Journal of Physics vol. 56** no. 5 (1988), pages 413-415. https://doi.org/10.1119/1.15622
65. Albert Einstein, **Relativity: The Special and the General Theory** (Meuthen, 1920).
66. Irwin I. Shapiro, "*Fourth Test of General Relativity*", **Physical Review Letters vol. 13**, no. 26 (28 December 1964), pages 789-791. https://doi.org/10.1103/PhysRevLett.13.789
67. Clifford M. Will, **Was Einstein Right? Putting General Relativity to the Test** (Basic Books, 1986), chapter 6 "*The Time Delay of Light: Better Late Than Never*". ISBN 0465090885
68. O. Chwolson, "Über Eine Mögliche Form Fiktiver Doppelsterne", **Astronomische Nachrichten vol. 221**, no. 20 (1924), pages 329-330. https://doi.org/10.1002/asna.19242212003
69. W.K. Clifford, "*On the Space-Theory of Matter*" (lecture abstract, 21 February 1870), **Proceedings of the Cambridge Philosophical Society (1864-1876)**, **vol. 2**, pages 157-158.
70. Eric Baird, "*Reduction to special relativity may be unphysical*" (August 2018) https://doi.org/10.13140/RG.2.2.14377.70241/1
71. Albert Einstein, "*Do gravitational fields play an essential part in the structure of the elementary particles of matter?*" **Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)**. (1919), pages 433-
translated and reprinted in, **The Principle of Relativity** (Dover, 1923): 189-198 . ISBN 0486600815
72. Albert Einstein., Leopold Infeld, and Banesh Hoffmann. "The Gravitational Equations and the Problem of Motion". Annals of Mathematics 39, no. 1 (1938): 65–100. https://doi.org/10.2307/1968714.
73. Einstein, A., and L. Infeld. "*On The Motion of Particles in General Relativity Theory*", **Canadian Journal of Mathematics**, **vol. 1**, no. 3 (June 1949), pages 209-241. https://doi.org/10.4153/CJM-1949-020-8
74. Richard Feynman, "*Seeking New Laws*", lecture from the lecture series "**On the Character of Physical Law**" (1964). https://www.youtube.com/watch?v=phfaDMEMFaU
75. Albert Einstein, "*Geometry and Experience*", An expanded form of an Address to the Prussian Academy of Sciences in Berlin on January 27th, 1921, translated and reprinted in:
Albert Einstein, **Sidelights on Relativity** (Meuthen, 1922).
76. Steven D. Stellman, "*A Spherical Chicken*" **Science**, **vol. 182**, Issue 4119, (28 Dec 1973) pp. 1296 https://doi.org/10.1126/science.182.4119.1296-b
77. Bertrand Russell, "*V. Mathematics and the Metaphysicians*", in **Mysticism and Logic and Other Essays** (George Allen and Unwin, 1917) pages 74-93 . https://en.wikisource.org/wiki/Mysticism_and_Logic_and_Other_Essays/Chapter_05
78. Albert A. Michelson, and Edward W. Morley, "*On the relative motion of the earth and the luminiferous Æther*", **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science**, **vol. 24** (1887) pages 151, and 449-463. https://doi.org/10.1080/14786448708628130
79. Oliver Joseph Lodge, "*XV. Aberration Problems.—A Discussion Concerning the Motion of the Ether near the Earth, and Concerning the Connexion between Ether and Gross Matter; with Some New Experiments*", **Philosophical Transactions of the Royal Society of London (A)**, (1893) pages 727-804, and pages 31-32 (figures). https://doi.org/10.1098/rsta.1893.0015
80. Scott Walter, "*Hermann Minkowski's Approach to Physics*", **Mathematische Semesterberichte vol.55** (1 September 2008), pages 213–235. https://doi.org/10.1007/s00591-008-0044-4
81. Albert Einstein, "*Aether and the Theory of Relativity*", an address delivered on May 5th, 1920, at the University of Leyden. Translated and published in:
Albert Einstein, **Sidelights on Relativity** (Meuthen, 1922).
82. Herbert E. Ives, and George R. Stilwell, "*An experimental study of the rate of a moving atomic clock*" **JOSA vol. 28**, no. 7 (July 1938), pages 215-226.
83. J. Huchra, M. Gorenstein, S. Kent, I. Shapiro, G. Smith, E. Horine, and R. Perley. "*2237 + 0305: A New and Unusual Gravitational Lens*", **The Astronomical Journal vol. 90** (1 May 1985), pages 691–696 . https://doi.org/10.1086/113777
84. Karl Friedrich Gauss, "*Disquisitiones generales circa superficies curvas*" (1827)
Carl Friedrich Gauss, "*General Investigations of Curved Surfaces*", translated by James Cadall Morehead and Adam Miller Hildebeitel, in English, **Project Gutenberg**, 1902) http://www.gutenberg.org/files/36856/36856-pdf.pdf
85. Bernhard Riemann, "*Ueber die Hypothesen, welche der Geometrie zu Grunde liegen*", **Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen**, **vol.13** (1868), pages 133-150 http://www.deutschestextarchiv.de/riemann_hypothesen_1867/7

86. Bernhard Riemann, "*On the Hypotheses which lie at the Bases of Geometry. Translated by William Kingdon Clifford*" **Nature vol. 8**, 14–17 (1st May 1873) https://www.emis.de//classics/Riemann/WKCGeom.pdf
87. William Kingdon Clifford, **On the Common Sense of the Exact Sciences, Second Edition** (1886), Chapter IV: §19: "*On the Bending of Space*".
88. Nikolai I. Lobachevsky, "*On the foundations of geometry*", **Kazan Messenger vol. 25** (1829), pages 28- .
89. Johann Karl Friedrich Zollner, **Über die Natur der Cometen: Beiträge zur Geschichte und Theorie der Erkenntnis** (Engelmann, 1872).
90. Albert Einstein, "*Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*", **Sitzungsberichte Preussischen Akad. d. Wissenschaften, Berlin (Math.Phys.)** (February 1917), pages 142-152.
Bibcode: 1917SPAW.......142E
translated and reprinted in:
Albert Einstein, **The Principle of Relativity** (Methuen 1923) "*Cosmological Considerations on the General Theory of Relativity*", pages 175-188. ISBN 0486600815
91. Cormac O'Raifeartaigh, Michael O'Keeffe, Werner Nahm, and Simon Mitton, "*Einstein's 1917 Static Model of the Universe: A Centennial Review*", **The European Physical Journal H 42**, no. 3 (1 August 2017), pages 431-474. https://doi.org/10.1140/epjh/e2017-80002-5
92. Simon Newcomb, "*Modern Mathematical Thought*" (address given to the New York Mathematical Society, December 1893), **Bulletin of the New York Mathematical Society vol. 3**, no. 4 (January 1894), pages 95-107.
93. Simon Newcomb, "*The Philosophy of Hyperspace*" (address given on 27th December 1897), **Bulletin of the American Mathematical Society vol. 4**, no. 5 (February 1898), pages 187-195.
94. Lewis Carroll, **Sylvie and Bruno** (Macmillan, 1890) Chapters 21 and 23.
95. Albert Einstein, and Marcel Grossman, "*Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation*" [*Outline of a Generalized Theory of Relativity and of a Theory of Gravitation*], **Zeitschrift für Mathematik und Physik**, **vol. 62** (1913), pages 225-261.
96. David E. Rowe, "*Einstein Meets Hilbert: At the Crossroads of Physics and Mathematics*" **Physics in Perspective vol. 3**, no. 4 (1 November 2001), pages 379-424. https://doi.org/10.1007/PL00000538
97. Karl Schwarzschild, "*On the Gravitational Field of a Point-Mass, According to Einstein's Theory*", **Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, S**. (1916) pages 189–196
Translated and reprinted in
The Abraham Zelmanov Journal, **vol. 1** (2008), pages 10-19. ISSN 1654-9163
98. S. Chandrasekhar, "*Einstein and General Relativity: Historical Perspectives*", **American Journal of Physics vol. 47**, no. 3 (1 March 1979), pages 212-217. https://doi.org/10.1119/1.11666
99. W.L Fadner, "*Did Einstein Really Discover 'E=mc2' ?*" **American Journal of Physics vol. 56** (1 February 1988): 114-122. https://doi.org/10.1119/1.15713
100. Wolfgang Rindler, "*General Relativity before Special Relativity: An Unconventional Overview of Relativity Theory*". **American Journal of Physics vol. 62**, no. 10 (1 October 1994), pages 887-893. https://doi.org/10.1119/1.17734
101. Augustin Fresnel, "*Lettre d'Augustin Fresnel à François Arago sur l'influence du mouvement terrestre dans quelques phénomènes d'optique*", **Annales de Chimie et de Physique vol. 9** (1818).
102. H. Fizeau, "*Sur les hypothèses relatives à l'éther lumineux*", **Comptes Rendus. vol. 33** (1851), pages 349–355.
103. H. Fizeau, "*The Hypotheses Relating to the Luminous Aether, and an Experiment which Appears to Demonstrate that the Motion of Bodies Alters the Velocity with which Light Propagates itself in their Interior*", **Philosophical Magazine**, **vol. 2** (1851), pages 568–573.
104. Reginald Victor Jones, "*'Fresnel Aether Drag' in a Transversely Moving Medium*", **Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences vol. 328**, no. 1574 (13 June 1972), pages 337-352. https://doi.org/10.1098/rspa.1972.0081
105. Albert Einstein, "*Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*" ("*Investigations on the theory of Brownian Movement*"), **Annalen der Physik vol. 322**, no. 8, pages 549-560. https://doi.org/10.1002/andp.19053220806
106. Albert Einstein, "*Über Einen Die Erzeugung Und Verwandlung Des Lichtes Betreffenden Heuristischen Gesichtspunkt*", **Annalen Der Physik vol. 322**, no. 6 (1905), 132–48. https://doi.org/10.1002/andp.19053220607
107. Hasselkamp, D., E. Mondry, and A. Scharmann. "*Direct Observation of the Transversal Doppler-Shift*", **Zeitschrift Für Physik A Atoms and Nuclei vol. 289**, no. 2 (1 June 1979), pages 151-155. https://doi.org/10.1007/BF01435932

108. Darrell Huff, **How to Lie with Statistics** (Norton, 1954). ISBN 0575004207
109. George A. Akerlof, "*The Market for "Lemons*", **Quality Uncertainty and the Market Mechanism, The Quarterly Journal of Economics**, **Volume 84**, Issue 3 (August 1970), Pages 488-500. https://doi.org/10.2307/1879431
110. James T. Cushing, "*Electromagnetic Mass, Relativity, and the Kaufmann Experiments*", **American Journal of Physics vol. 49**, no. 12 (1 December 1981), pages 1133-1149. https://doi.org/10.1119/1.12561
111. Mary Jo Nye, "*N-Rays: An Episode in the History and Psychology of Science*", **Historical Studies in the Physical Sciences vol. 11**, no. 1 (1 January 1980), pages 125-156. https://doi.org/10.2307/27757473
112. Do, Tuan, Aurelien Hees, Andrea Ghez, Gregory Martinez, Devin Chu, Siyao Jia, Shoko Sakai, *et. al.* "*Relativistic Redshift of the Star S0-2 Orbiting the Galactic Center Supermassive Black Hole*" **Science vol. 365** (25 July 2019), pages 664-668 . https://doi.org/10.1126/science.aav8137
113. Monwhea Jeng, "*A Selected History of Expectation Bias in Physics*" **American Journal of Physics vol. 74**, no. 7 (1 July 2006), pages 578-583. https://doi.org/10.1119/1.2186333
114. Norris S. Hetherington, "*Just How Objective is Science?*", Nature vol. 306 (22/29 December 1983) pages 727-730.
115. John Bernhard Stallo, "*The Primary Concepts of Modern Physical Science III – The Assumption of the Essential Solidity of Matter*", **Popular Science Monthly vol. 4** (December 1873).
116. Luca Consoli, "*Scientific Misconduct and Science Ethics: A Case Study Based Approach*", **Science and Engineering Ethics vol. 12** (1 August 2006), pages 533-541. https://doi.org/10.1007/s11948-006-0051-6
117. Joseph Priestley, **The History and Present State of Discoveries Relating to Vision, Light, and Colours** (*a.k.a.* **History of Optics**) (J. Johnson, 1772).
118. Madge G. Adam "Astronomical evidence for the gravitational redshift" The New Scientist (29 September 1960) pages 863-866.
119. N.S. Hetherington, "Sirius B and the Gravitational Redshift - an Historical Review", **Quarterly Journal of the Royal Astronomical Society vol.21** (1 September 1980) pages 246-252. Bibcode: 1980QJRAS..21..246H
120. J.B. Holberg, "*Sirius B and the Measurement of the Gravitational Redshift*" **Journal for the History of Astronomy vol. 41** (2010) pages 41-64. https://doi.org/10.1177/002182861004100102
121. Eric Baird, "*Cosmological redshifts vs. gravitational redshifts*" (June 2018) https://doi.org/10.13140/RG.2.2.26042.29128
122. A.P. French, "*The Strange Case of Emil Rupp*", **Physics in Perspective vol. 1**, **no. 1** (1 March 1999), pages 3-21 . https://doi.org/10.1007/s000160050002
123. P.P. Ewald (1912) Dispersion und Doppelbreehung von Elektronengittern (Kribtallen). Thes, Univ. of Munich, 1912) Gottingen, Dieterischsche Universitats - Buchdruckerei, page 45-.
translated in:
P.P. Ewald, "*On the Foundations of Crystal Optics. Part 1. Dispersion Theory. Part 2. Theory of Reflection and Refraction*": **Fort Belvoir, VA: Defense Technical Information Center**, 26 August 1970. https://doi.org/10.21236/AD0717697 https://apps.dtic.mil/dtic/tr/fulltext/u2/717697.pdf
124. C. W. Oseen, "*Über Die Wechselwirkung Zwischen Zwei Elektrischen Dipolen Und Üer Die Drehung Der Polarisationsebene in Kristallen Und Flüssigkeiten*", **Annalen Der Physik**, **vol. 353**, no. 17 (1915), pages 1-56. https://doi.org/10.1002/andp.19153531702
125. Masud Mansuripur, "*The Ewald-Oseen Extinction Theorem*", **Optics and Photonics News vol. 9**, no. 8 (1 August 1998), pages 50-55. https://arxiv.org/abs/1507.05234 https://doi.org/10.1364/OPN.9.8.000050
126. G.H. Keswani, *Fizeau experiment, aberration of starlight and Einstein: Relativity (issues)* Appendix 2 *Extinction Theorem*, printed in **Physical Interpretations of Relativity Theory: Proceedings of International Scientific Meeting PIRT-2003 Moscow: 30 June – 03 July**, (2003) *Eds.* M.C. Duffy, V.O. Gladyshev, A.N. Morozov pages 316-322. ISBN 5893440064
127. J.G. Fox. "*Experimental Evidence for the Second Postulate of Special Relativity*". **American Journal of Physics vol. 30** (1 April 1962), pages 297-300. doi: 10.1119/1.1941992
128. George Biddell Airy, "*On a Supposed Alteration in the Amount of Astronomical Aberration of Light …*", **Proceedings of the Royal Society of London vol. 20** (1872), issue 130-138, pages 35-39 https://doi.org/10.1098/rspl.1871.0011
129. G. D. Scott, and H. J. van Driel. "*Geometrical Appearances at Relativistic Speeds*", **American Journal of Physics vol. 38**, no. 8 (1 August 1970), pages 971-977. https://doi.org/10.1119/1.1976550

130. Steve Carlip, "*Aberration and the Speed of Gravity*", **Physics Letters A 267**, no. 2–3 (March 2000), pages 81-87. https://doi.org/10.1016/S0375-9601(00)00101-8
131. Clifford M. Will, "*The Confrontation between General Relativity and Experiment*", **Living Reviews in Relativity 9**, no. 1 (27 March 2006): 3. https://doi.org/10.12942/lrr-2006-3 https://arxiv.org/abs/gr-qc/0510072v2
132. The Nobel Prize in Physics 1993: Russell A. Hulse and Joseph H. Taylor, Jr. https://www.nobelprize.org/prizes/physics/1993/press-release/
133. Lutendo Nyadzani, and S. Razzaque. "*Polarization of Gravitational Waves from Binary Systems*", in **Proceedings of Science (PoS) High Energy Astrophysics in Southern Africa – HEASA2018 (2019)** https://doi.org/10.22323/1.338.0029 https://inspirehep.net/literature/1747211
134. Pierre Azoulay, Christian Fons-Rosen, and Joshua S. Graff Zivin, "*Does Science Advance One Funeral at a Time?*" **American Economic Review vol. 109**, no. 8 (August 2019), pages 2889-2920. https://doi.org/10.1257/aer.20161574
135. George Orwell, **Nineteen Eighty-Four** (1949) Appendix: The Principles of Newspeak. https://orwell.ru/library/novels/1984/english/en_app.
136. C.M. Will, **Theory and Experiment in Gravitational Physics, revised edition** (CUP, 1993) "*2.2 Basic criteria for the viability of a gravitation theory*". ISBN 9781107117440
137. C.M. Will, **Theory and Experiment in Gravitational Physics, second edition** (CUP, 2018) "*2.1 The Dicke Framework*". ISBN 9781107117440
138. Di Casola, Eolo, Stefano Liberati, and Sebastiano Sonego, "*Nonequivalence of Equivalence Principles*", **American Journal of Physics vol. 83**, no. 1 (January 2015), pages 39–46. https://doi.org/10.1119/1.4895342
139. American Bar Association, **Handbook on the Antitrust Aspects of Standards Setting** (American Bar Association, 2005). ISBN 1590314123
140. Eugene P. Wigner, "*The Unreasonable Effectiveness of Mathematics in the Natural Sciences*", in **Communications in Pure and Applied Mathematics**, **vol. 13**, No. 1 (February 1960).
141. Albert Einstein and Leopold Infeld, **The Evolution of Physics** (CUP, 1938).
142. Abraham Robinson, "*Non-standard analysis*", **Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings, series A**, **vol. 64** (1961), pages 432-440.
143. Abraham Robinson, **Non-Standard Analysis** (North-Holland, 1961).
144. Willem de Sitter, "*A Proof of the Constancy of the Velocity of Light*", **Proceedings of the Royal Netherlands Academy of Arts and Sciences vol. 15** (1913), pages 1297-98.
145. Willem de Sitter, "*On the Constancy of the Velocity of Light*", **Proceedings of the Royal Netherlands Academy of Arts and Sciences**, **vol. 16** (1913), pages 395-396.
146. Kenneth Brecher, "Is the Speed of Light Independent of the Velocity of the Source?" **Physical Review Letters**, **vol. 39**, no. 17 (24 October 1977), pages 1051-1054. https://doi.org/10.1103/PhysRevLett.39.1051
147. R.S. Shankland, "*Conversations with Albert Einstein*", **American Journal of Physics**, **vol. 31**, no. 1 (1 January 1963), pages 47-57. https://doi.org/10.1119/1.1969236
148. John Preskill, "*Do Black Holes Destroy Information?*", in **Black Holes, Membranes, Wormholes and Superstrings: Proceedings of the International Symposium** (World Scientific, 1994) ISBN 9789814554084 http://arxiv.org/abs/hep-th/9209058 (1992)
149. Leonard Susskind, "*Black Holes and the Information Paradox*", **Scientific American**, April 1997. https://doi.org/10.1038/scientificamerican0497-52
150. Steve Carlip, "*Quantum Gravity: A Progress Report*", **Reports on Progress in Physics vol. 64**, no. 8 (July 2001), pages 885-942. https://doi.org/10.1088/0034-4885/64/8/301
151. Gerard 't Hooft, "*Dimensional Reduction in Quantum Gravity*" arXiv:gr-qc/9310026 (1993). http://arxiv.org/abs/gr-qc/9310026
152. Leonard Susskind, "*The World as a Hologram*", **Journal of Mathematical Physics**, **vol. 36**, no. 11 (1 November 1995), pages 6377-6396. https://doi.org/10.1063/1.531249
153. Jacob D. Bekenstein, "*Information in the Holographic Universe*", **Scientific American** (April 2007) https://doi.org/10.1038/scientificamerican0407-66sp
154. Kip S. Thorne, Richard H. Price, and Douglas A. MacDonald, **Black Holes: The Membrane Paradigm** (Yale, 1986). ISBN 9780300037692

155. Richard H. Price, and Kip S. Thorne, "The Membrane Paradigm for Black Holes", Scientific American, Vol. 258 no. 4 (April 1988), pages 69-71. https://doi.org/10.1038/scientificamerican0488-69
156. Werner Heisenberg, **Encounters with Einstein : And Other Essays on People, Places, and Particles** (Princeton University Press, 1989) pages 113-114 . ISBN 0691024332
157. Niels Bohr, "*Über die Serienspektra der Elemente*" **Zeitschrift für Physik vol. 2**, no. 5 (1 October 1920), pages 423–69 . https://doi.org/10.1007/BF01329978
158. J.R. Nielsen, **Niels Bohr Collected Works. Vol. 3. The Correspondence Principle (1918 – 1923)** (Elsevier, 2013). ISBN 0080871011
159. Jan Faye "*Copenhagen Interpretation of Quantum Mechanics*", **The Stanford Encyclopedia of Philosophy (Winter 2019 Edition)**, Edward N. Zalta (ed.) https://plato.stanford.edu/archives/win2019/entries/qm-copenhagen/
160. Khavtgain Namsrai, "*Space-Time Structure near Particles and Its Influence on Particle Behavior*", **International Journal of Theoretical Physics**, **vol. 23**, no. 11 (1 November 1984), pages 1031-1041. https://doi.org/10.1007/BF02213415
161. Stephen W. Hawking, "*Information Preservation and Weather Forecasting for Black Holes*" (22 January 2014). http://arxiv.org/abs/1401.5761
162. Arthur Stanley Eddington, **The Nature of the Physical World** (Macmillan, 1928), §4: "The Running-Down of the Universe", pages 68-69. https://doi.org/10.5962/bhl.title.5859
163. Harry G. Frankfurt, **On Bullshit** (Princeton University Press, 2005). ISBN 0691122946
164. Albert Einstein, "*On the Generalized Theory of Gravitation*", **Scientific American vol. 182**, no. 4 (April 1950) pages 13-17. https://www.jstor.org/stable/24967425
165. J.A. Wheeler, and R.P. Feynman. "*Classical Electrodynamics in Terms of Direct Interparticle Action*", **Reviews of Modern Physics vol. 21**, no. 3 (July 1949), pages 425-433 . https://doi.org/10.1103/RevModPhys.21.425
166. Simon Schaffer, "*Glass works: Newton's prisms and the uses of experiment*", in **The Uses of Experiment: Studies in the Natural Sciences**, edited by David Gooding, Trevor Pinch, and Simon Schaffer (CUP, 1985), pages 67-104. ISBN 0521337682
167. George Gamow, "*Mr. Tompkins in Wonderland*", **Discovery: The Magazine of Scientific Progress**, **vol.1**, pages 431- (Cambridge: At the University Press, 1938)
reprinted as:
George Gamow, Mr. Tompkins in Wonderland, Or Stories of c, G, and h (Cambridge: At the University Press, 1939), and rewritten 1962.
168. George Gamow, **Matter, Earth and Sky** (Prentice-Hall, 1958) "*7: Modern Views on Space, Time, and Motion: Space-Time Transformation*", pages 183-190.
169. Jacob Bronowski, **The Ascent of Man** (BBC, 1973), pages 249-251. ISBN 0563170646
170. George Gamow, and Roger Penrose, **Mr Tompkins in Paperback** (CUP 1993), page x. ISBN 0521447712
171. John D. Norton, "*General Covariance and the Foundations of General Relativity: Eight Decades of Dispute*", **Reports on Progress in Physics**, **vol. 56**, no. 7 (1993) pages 791-858. https://doi.org/10.1088/0034-4885/56/7/001
172. Thomas Gold, "The Arrow of Time", printed in pages 81-95 of the proceedings of the 1958 Solvay conference *Instituts Solvay Conseil de physique*, **La Structure Et L'évolution de L'univers: Rapports Et Discussions**. (1958). http://www.solvayinstitutes.be/pdf/Proceedings_Physics/1958.pdf
173. J.J. Halliwell, J. Pérez-Mercader, and W. H. Zurek. **Physical Origins of Time Asymmetry**, (Cambridge University Press 1996), cover description. ISBN 0521433282
174. David Merritt, "*Cosmology and Convention*", **Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics vol. 57** (1 February 2017), pages 41-52. https://doi.org/10.1016/j.shpsb.2016.12.002
175. Georges Lemaître, "Un Univers Homogène de Masse Constante et de Rayon Croissant Rendant Compte de La Vitesse Radiale Des Nébuleuses Extra-Galactiques", **Annales de La Société Scientifique de Bruxelles vol. 47** (1927), pages 49-59.
176. Edwin Hubble, "*A relation between distance and radial velocity among extra-galactic nebulae*", **PNAS vol. 15**, no.3 (March 15, 1929), pages 168-173. https://doi.org/10.1073/pnas.15.3.168

177. A. Friedmann, **Zeitschrift fur Physik**, **vol. 10** (1922), pages 377-386.
translated as:
A. Friedmann, "*On the Curvature of Space*", **General Relativity and Gravitation vol. 31**, no. 12 (1 December 1999), pages 1991-2000. https://doi.org/10.1023/A:1026751225741

178. A.S. Eddington, "*On the Instability of Einstein's Spherical World*", **Monthly Notices of the Royal Astronomical Society vol. 90**, no. 7 (9 May 1930), pages 668-678. https://doi.org/10.1093/mnras/90.7.668

179. Harry Nussbaumer, "*Einstein's Conversion from His Static to an Expanding Universe*", **The European Physical Journal H 39**, no. 1 (1 February 2014), pages 37-62. https://doi.org/10.1140/epjh/e2013-40037-6

180. William Kingdon Clifford, **Lectures and Essays** (Macmillan, 1879) "*III: The Postulates of the Science of Space*"

181. Matt Visser, "*Acoustic Black Holes: Horizons, Ergospheres, and Hawking Radiation*" **Classical and Quantum Gravity vol. 15**, no. 6 (1 June 1998), pages 1767-1791. https://doi.org/10.1088/0264-9381/15/6/024 https://arxiv.org/abs/gr-qc/9712010

182. William G. Unruh, "*Experimental Black-Hole Evaporation?*" **Physical Review Letters vol. 46**, no. 21 (25 May 1981), pages 1351-1353. https://doi.org/10.1103/PhysRevLett.46.1351

183. William G. Unruh, "Experimental Black-Hole Evaporation" **Physics Today** (9 September 2016). https://doi.org/10.1063/PT.5.2047

184. Carlos Barceló, Stefano Liberati, and Matt Visser, "*Analogue Gravity*", **Living Reviews in Relativity vol.14**, no. 1 (2005 / 11 May 2011), article 3. https://doi.org/10.12942/lrr-2011-3

185. Matt Visser, Carlos Barceló, and Stefano Liberati, "*Analogue Models of and for Gravity*", **General Relativity and Gravitation vol. 34**, no. 10 (1 October 2002), pages 1719-1734. https://doi.org/10.1023/A:1020180409214

-==-