The kinematic origin of the cosmological redshift

Emory F. Bunn*

Physics Department, University of Richmond, Richmond, VA 23173

David W. Hogg

Center for Cosmology and Particle Physics and Department of Physics, New York University, 4 Washington Place, New York, NY 10003

(Dated: February 28, 2009)

A common belief about big-bang cosmology is that the cosmological redshift cannot be properly viewed as a Doppler shift (that is, as evidence for a recession velocity), but must instead be viewed in terms of the stretching of space. We argue that the most natural interpretation of the redshift is in fact as a Doppler shift, or rather as the accumulation of many infinitesimal Doppler shifts. The stretching-of-space interpretation obscures a central idea of relativity, namely that of coordinate freedom, more specifically the idea that it is always valid to choose a coordinate system that is locally Minkowskian. We show that, in any spacetime, an observed frequency shift can be interpreted either as a kinematic (Doppler) shift or a gravitational shift by imagining a suitable family of observers along the photon's path. In the context of the expanding Universe, the kinematic interpretation corresponds to a family of comoving observers and hence seems to be the more natural one.

I. INTRODUCTION

Many descriptions of big-bang cosmology declare that the observed redshift of distant galaxies is not a Doppler shift but is instead due to the "stretching of space." The purpose of this paper is to examine the meaning of such statements and to assess their validity. We wish to make clear at the outset that we are not suggesting any doubt about either the observations or the general-relativistic equations that successfully explain them. Rather, our focus is on the question of interpretation: given that a photon does not arrive at the observer conveniently labeled "Doppler shift," "gravitational shift," or "stretching of space," when can or should we apply these labels?

Arguably the enlightened cosmologist never asks this question. In the curved spacetime of general relativity, there is no unique way to compare vectors at widely separated spacetime points, and hence the notion of the relative velocity of a distant galaxy is almost meaningless. Indeed, the inability to compare vectors at different points is the definition of a curved spacetime.[1–4] In practice, however, the enlightened view is far from universal. The view presented by cosmologists and astrophysicists, particularly when talking to nonspecialists, is generally that distant galaxies are "really" at rest, and that the observed redshift is a consequence of some sort of "stretching of space," which is physically completely distinct from the usual kinematic Doppler shift. In these descriptions, statements that are artifacts of a particular coordinate system are presented as if they were statements about the world, resulting in misunderstandings about the nature of spacetime in relativity. In this paper we will show that the redshifts of distant objects in the expanding Universe *can* be viewed as a kinematic shift due to relative velocities, and we will argue that – if we are forced to "interpret" the redshift – this interpretation is more natural than any other.

We begin with examples of the description of the cosmological redshift in the first three introductory astronomy textbooks chosen at random from the bookshelf of one of the authors.

- The cosmological redshift "is *not* the same as a Doppler shift. Doppler shifts are caused by an object's *motion through space*, whereas a cosmological redshift is caused by the *expansion of space*" (emphasis in original).[5]

- "A more accurate view [than the Doppler effect] of the redshifts of galaxies is that the waves are stretched by the stretching of space they travel through .... If space is stretching during all the time the light is traveling, the light waves will be stretched as well."[6]

- "Astronomers often express redshifts as if they were radial velocities, but the redshifts of the galaxies are not Doppler shifts .... Einstein's relativistic Doppler formula applies to motion through space, so it does not apply to the recession of the galaxies."[7]

More advanced textbooks often avoid this language; for instance, the books by Peacock[8] and Linder[9] give particularly careful and clear descriptions of the nature of the cosmological redshift. However, statements similar to those above can be found even in some advanced textbooks. For example, a leading text at the advanced undergraduate level says that Doppler shifts "are produced by peculiar and not by recession velocities."[10] In this paper, we will argue, as others have before us,[11–14] that statements such as these are misleading and foster misunderstandings about the nature of space and time in contemporary physics and cosmology.

From the point of view of general relativity, the "stretching of space" explanation of the redshift is quite problematic. Light is of course governed by the Maxwell equations (or their general relativistic generalization), which contain no "stretching of space term" and no information on the current size of the Universe. On the contrary, one of the most important ideas of general relativity is that spacetime is always locally indistinguishable from the (non-stretching) spacetime of special relativity, which means that a photon doesn't know what the scale factor of the Universe is doing.[15]

The emphasis in many textbooks on the stretching-of-spacetime interpretation of the cosmological redshift causes the reader to take far too seriously the stretching-rubber-sheet analogy for the expanding Universe. For example, it is sometimes stated as if it were obvious that "it follows that all wavelengths of the light ray are doubled" if the scale factor doubles.[10] While this statement is true, it is certainly not obvious. After all, solutions to the Schrödinger equation, such as the electron orbitals in the hydrogen atom, don't stretch as the Universe expands, so why do solutions to the Maxwell equations?

A student presented with the stretching-of-space description of the redshift cannot be faulted for concluding, incorrectly, that hydrogen atoms, the Solar System, and the Milky Way Galaxy must all constantly "resist the temptation" to expand along with the Universe. One way to see that this belief is in error is to consider the problem sometimes known as the "tethered galaxy problem":[16,17] A galaxy is tethered to the Milky Way, forcing the distance between the two to remain constant. When the tether is cut, does the galaxy join up with the Hubble flow and start to recede due to the expansion of the Universe? The intuition that says that objects suffer from a temptation to be swept up in the expansion of the Universe will lead to an affirmative answer, but the truth is precisely the reverse: unless there is a large cosmological constant and the galaxy's distance is comparable to the Hubble length, the galaxy falls towards us.[13,14] Similarly, it is commonly believed that the Solar System has a very slight tendency to expand due to the Hubble expansion (although this tendency is generally thought to be negligible in practice). Again, this can be shown by explicit calculation not to be true[18,19]: the tendency to expand due to the stretching of space is nonexistent, not merely negligible.

The expanding rubber sheet is quite similar to the ether of pre-relativity physics: although it is intuitively appealing, it makes no correct testable predictions, and some incorrect ones such as the examples described above. It therefore has no rightful place in the theory. (Some authors[20,21] have argued that considerations such as these do not refute the notion that space is "really" expanding. We agree with the calculations in these papers but have a difference of opinion regarding the most useful language to use to describe the relevant phenomena.)

In one set of circumstances, the proper interpretation of the redshift seems quite clear: when the curvature of spacetime is small over the distance and time scales traveled by a photon, it is natural to interpret the observed frequency shift as a Doppler shift. This is, for example, the reason that a police officer can give you a speeding ticket based on the reading on a radar gun; as far as we know, no one has successfully argued in traffic court that there is any ambiguity in interpreting the observed frequency shift as a Doppler shift.[22] In the expanding Universe, spacetime curvature is small over regions encompassing nearby objects, specifically those with $z = \Delta\lambda/\lambda \ll 1$. There should be no hesitation about calling the observed redshifts Doppler shifts in this case, just as there is none in traffic court. Surprisingly, however, many people seem to believe that the "stretching of space" interpretation of the redshift is the only valid one, even in this limit. We will examine the interpretation of redshifts of nearby objects more carefully in Section II.

Aside from low-redshift sources, there is another case in which spacetime curvature can be neglected in considering cosmological redshifts: low-density cosmological models. An expanding universe with density $\Omega = 0$ (often known as the Milne model[23]) is simply the flat Minkowski spacetime of special relativity, expressed in nonstandard coordinates.[24] In an $\Omega = 0$ universe, there are no gravitational effects at all, so any observed redshift, even of a very distant galaxy, must be a Doppler shift. Furthermore, for low but nonzero density ($\Omega \ll 1$), the length scale associated with spacetime curvature is much longer than the horizon distance. This means that spacetime curvature effects (i.e., gravitational effects) are weak throughout the observable volume, and the special-relativistic Doppler shift interpretation remains valid even for galaxies with arbitrarily high redshift.

The more interesting cases are when z and Ω are not small – that is, when the source is distant enough that gravitational effects matter over the photon's trajectory. The consensus is that the Doppler shift language must be eschewed in this setting. In Section III, we review a standard argument[8,13] that even in this case the redshift can be interpreted as the accumulation of infinitesimal Doppler shifts along the line of sight, and we further argue that there is a natural way to interpret the redshift as a single (non-infinitesimal) Doppler shift.

A common objection to this claim is that the coordinate velocity is not related to the redshift in accordance with the special-relativistic Doppler formula.[25,26] However, the velocity referred to in this claim is merely an artifact of a particular choice of time coordinate; specifically, it is the rate of change of the distance to the object with respect to the cosmic time coordinate, as measured at the present cosmic time. This is an unnatural thing to talk about, since it depends on data outside of the observer's light cone. The more natural velocity to talk about is the velocity of the object at the time it crossed our past light cone, relative to us today. This is itself a coordinate-dependent concept, but as we will show in Section III, the most natural way to specify this velocity operationally leads to a result that is consistent with the special-relativistic Doppler formula.

In Section IV, we widen our focus to consider frequency shifts in arbitrary curved spacetimes.

In general, in any curved spacetime, the observed frequency shift in a photon can be interpreted as either a kinematic effect (a Doppler shift) or as a gravitational shift. The two interpretations arise from different choices of coordinates, or equivalently from imagining different families of observers along the photon's path. We will describe this construction explicitly, and show that the comoving observers who are usually made use of to describe phenomena in the expanding Universe are precisely the ones that correspond to the Doppler shift interpretation.

II. REDSHIFTS OF NEARBY GALAXIES *ARE* DOPPLER SHIFTS.

We begin by returning to the Parable of the Speeding Ticket, mentioned in the previous section.

> A driver is pulled over for speeding. The police officer says to the driver, "According to Doppler shift of the radar signal I bounced off your car, you were traveling faster than the speed limit."
>
> The driver replies, "In certain coordinate systems, the distance between us remained constant during the time the radar signal was propagating. In such a coordinate system, our relative velocity is zero, and the observed wavelength shift was not a Doppler shift. So you can't give me a ticket."

If you believe that the driver has a legitimate argument, then you have our permission to believe that cosmological redshifts are not really Doppler shifts. If, on the other hand, you think that the officer is right, and the redshift can legitimately be interpreted as a Doppler shift, then you should believe the same thing about redshifts of nearby galaxies in the expanding Universe.

Why is the police officer right and the driver wrong? Assuming the officer majored in physics at the police academy, he might explain the situation like this: "Spacetime in my neighborhood is very close to flat. That means that I can lay down space and time coordinates in my neighborhood such that, to an excellent approximation, the rules of special relativity hold. Using those coordinates, I can interpret the observed redshift as a Doppler shift (since there is no gravitational redshift in flat spacetime) and figure out your coordinate velocity relative to me. The errors in this method are of the same order as the departures from flatness in the spacetime in a neighborhood containing both me and you; as long as I'm willing to put up with that very small level of inaccuracy, I can interpret that coordinate velocity as your actual velocity relative to me."

The principle underlying the officer's reasoning is, we hope, utterly uncontroversial. It is no different from the principle that lets football referees ignore the curvature of the Earth and use a flat coordinate grid in describing a football field.

We now return to the consideration of redshifts in the expanding Universe. As noted in Section I, it is instructive in this context (as in many others) to start with the zero-density expanding universe, also known as the Milne model[11]. The Milne universe consists of a set of galaxies expanding outward from an initial Big Bang, with the galaxies assumed to have negligible mass, so that the geometry is simply that of gravity-free Minkowski spacetime. This spacetime can be described either in the usual comoving coordinates of cosmology or in standard Minkowskian coordinates. Since there is no spacetime curvature and no gravity in this universe, it is clear that the observed redshifts should be interpreted as Doppler shifts. In comoving coordinates, however, the redshifts are easily seen to be the usual cosmological redshift. Clearly, in this case, there is no distinction between cosmological redshift and Doppler shift.

In our actual Universe, spacetime is not exactly flat, but we can approximate it as flat in a small neighborhood. One might be tempted to think that "approximating away" the curvature of spacetime is the same as approximating away the expansion altogether. (It seems to us that people who believe that cosmological redshifts cannot be viewed as Doppler shifts, even arbitrarily nearby, often believe something like this.) However, this belief is incorrect. When we approximate a small neighborhood of an expanding spacetime as flat, we make errors of order $(r/R_c)^2$ in the metric, where r is the size of the neighborhood and R_c is the curvature scale (generally the Hubble length). But the redshifts of galaxies in that neighborhood are of order (r/R_c), so they are not approximated away in this limit. (The Milne model corresponds to the limit $R_c \to \infty$, in which case no approximation at all is made.)

In comoving coordinates, the spacetime line element for the Robertson-Walker expanding universe is

$$ds^2 = -c^2dt^2 + [a(t)]^2(dr^2 + [S(r)]^2(d\theta^2 + \sin^2\theta\, d\phi^2)). \tag{1}$$

Here $S(r) = r$ for a flat universe; for a closed universe with positive curvature K, $S(r) = K^{-1}\sin(Kr)$, and for an open universe with negative curvature $-K$, $S(r) = K^{-1}\sinh(Kr)$. For realistic models, K^{-1} is at least comparable to the Hubble length, so $S(r) \approx r$ for nearby points in all cases.

Suppose that an observer at the origin at the present time t_0 measures the redshift of a galaxy at some comoving distance r. Let's assume that the galaxy is fairly nearby: $r/R_c \ll 1$ where R_c is the curvature length scale. Over such a distance scale, spacetime can be well approximated as flat. Let the observer lay down coordinates that approximate spacetime in her neighborhood as flat as

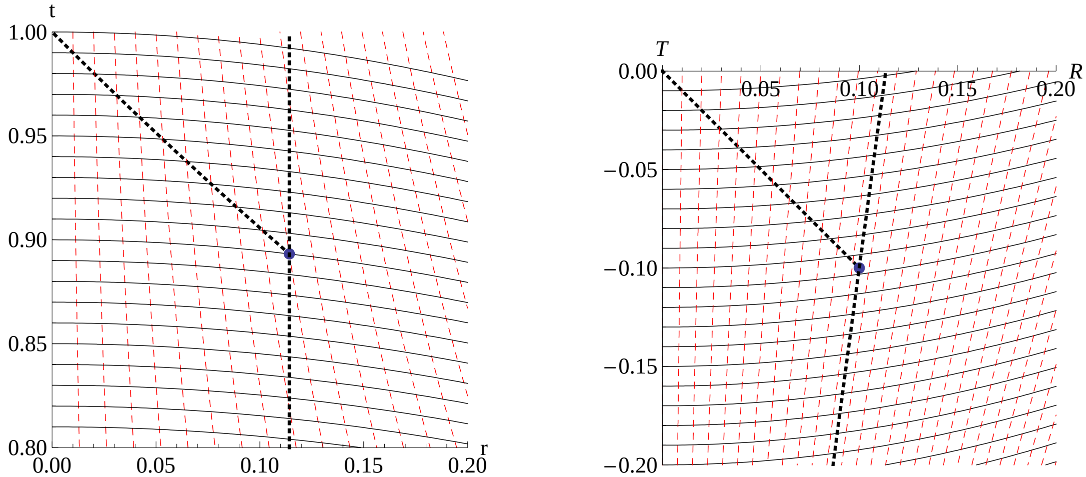

FIG. 1: Spacetime in different coordinate systems. The left panel is a plot of an expanding Robertson-Walker universe in the usual comoving coordinates (r,t). Coordinates are expressed in units of the Hubble length and time. The observer is located at $r=0$. The dotted curves show the world lines of a particular galaxy and of a light signal from that galaxy reaching the observer at the present time $(t=1)$. The dashed and solid curves are contours of constant Riemann normal coordinates (R,T) – that is, the coordinates that approximate flat spacetime as well as possible near the observer. In the right panel, the roles of the two coordinate systems are reversed: the coordinate axes are the Riemann normal coordinates (R,T), and the dashed and solid curves are contours of constant comoving coordinates (r,t). Note that as time passes, the galaxy moves to larger values of R – that is, it is not at rest in this coordinate system.

well as possible. That is, let her choose coordinates T,R,θ,ϕ such that

$$ds^2 \approx -c^2 dT^2 + dR^2 + R^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{2}$$

with the smallest possible errors in her neighborhood. There are several procedures for doing this, such as choosing Riemann normal coordinates. The resulting errors are of order $(r/R_c)^2$ in the metric. If we are willing to regard such errors as negligible, then we can legitimately approximate spacetime as flat. In such a coordinate system, a comoving galaxy does indeed have a velocity v relative to the observer, which is related to the observed redshift in the expected way to leading order:

$$z = (v/c) + O((r/R_c)^2). \tag{3}$$

Figure 1 illustrates the coordinate systems. In the left panel, the horizontal and vertical axes are standard comoving coordinates (r,t). The solid and dashed curves show contours of the coordinates (R,T) that best approximate spacetime as flat. World lines of a galaxy and of a light signal from the galaxy to the observer are also plotted. The panel on the right shows the view of spacetime in which the metric appears closest to flat spacetime – that is, it is the one to use when trying to use special relativistic language to describe our cosmic neighborhood. In this coordinate system, comoving galaxies are moving away from the observer, and the observed redshift is a Doppler shift.

Assuming that the observed galaxy is at a redshift much less than 1, errors in approximating spacetime as flat are small. As long as those errors are small enough to be neglected, the observer is in the same situation as the in the Parable of the Speeding Ticket. We don't merely say that the police officer is *allowed* to regard the observed redshift as a Doppler shift; we say that that is *the* natural interpretation of the shift. Precisely the same statement is true in the cosmological case: the only natural interpretation of the redshift of a nearby galaxy is as a Doppler shift.

III. REDSHIFTS OF DISTANT GALAXIES *CAN BE REGARDED* AS DOPPLER SHIFTS

For more distant galaxies, i.e., those with redshifts of order 1 or more, the light-travel distances involved are comparable to the curvature length scale, so the Riemann normal coordinate system described above cannot be used to approximate spacetime as flat with good accuracy. In this regime, then, one might imagine being required to drop the Doppler interpretation in favor of the "stretching of space" point of view. We will now argue that the Doppler interpretation is valid even in this regime.

Before examining questions of interpretation, let us begin by reviewing one standard derivation[8,13] of the cosmological redshift. Consider a photon that travels from a galaxy to a distant observer, both of whom are at rest in comoving coordinates. Imagine a family of comoving observers along the photon's path, each of whom measures the photon's frequency as it goes by. We assume that each observer is close enough to his neighbor that we can accommodate them both in one inertial reference frame and use special relativity to calculate the change in frequency from one observer to the next. If adjacent observers are separated by small distance δr, then their relative speed in this frame is $\delta v = H\,\delta r$, where H is the Hubble parameter. This is much less than c, so the frequency shift is given by the nonrelativistic Doppler formula

$$\delta\nu/\nu = -\delta v/c = -H\,\delta r/c = -H\,\delta t. \tag{4}$$

We know that $H = \dot{a}/a$ where a is the scale factor. We conclude that the frequency change is given by $\delta\nu/\nu = -\delta a/a$, or in other words that the frequency declines in inverse proportion to the scale factor. The overall redshift is therefore given by the rule

$$1 + z \equiv \frac{\nu(t_e)}{\nu(t_o)} = \frac{a(t_o)}{a(t_e)}, \tag{5}$$

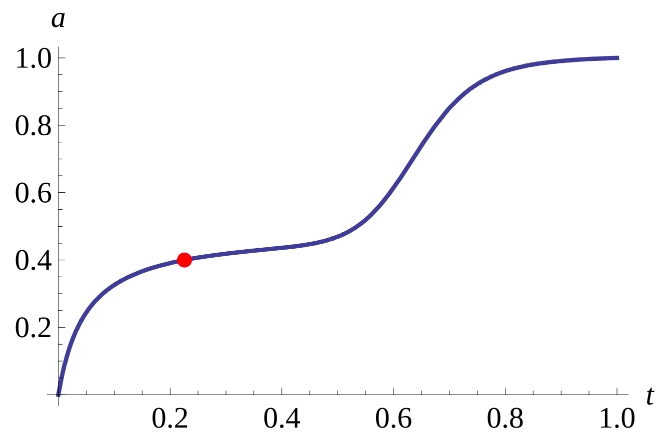

FIG. 2: The graph shows the scale factor of a hypothetical "loitering" universe, as a function of time (measured in units of the present time). At the time indicated by the dot, a galaxy emits radiation, which is observed at the present time. At the times of both emission and observation, the expansion is very slow; yet the galaxy's observed redshift is large.

where t_e, t_o refer to the times of emission and observation respectively.

In this derivation, we interpret the redshift as the accumulated effect of many small Doppler shifts along the photon's path. We now address the question of whether it makes sense to interpret the redshift as one big Doppler shift, rather than the sum of many small ones.

Figure 2 shows one common argument against such an interpretation. Imagine a universe whose expansion rate varies with time as shown in the figure. A galaxy emits radiation at the time t_e indicated by the dot, and that radiation reaches an observer at the present time t_0. The observed redshift is quite large: $z = a(t_0)/a(t_e) - 1 = 1.5$, which by the special-relativistic Doppler formula would correspond to a speed of $0.72c$. But at both the emission and absorption of the radiation, the expansion rate is very slow, and the speed $\dot{a}r$ of the galaxy is therefore much less than this value. Indeed, we can construct models in which both plateaus of $a(t)$ are arbitrarily close to flat. Since the redshift depends only on the values of a at those points, not on its slope, this would leave the redshift unaltered.

Upon closer examination, though, this argument is unconvincing, because the velocities being calculated are not the correct velocities. We should not compute velocities at a fixed instant of cosmic time (either $t = t_e$ or $t = t_0$); rather, we should compute the velocity of the galaxy *at the time of light emission* relative to the observer *at the present time*. After all, if a distant galaxy's redshift is measured today, we wouldn't expect the result to depend on what the galaxy is doing today, and certainly not on what the observer was doing long before the age of the dinosaurs.

In fact, when astrophysicists talk about what a distant object is doing "now," they often do not mean "at the present value of the cosmic time," but rather "at the time the object crossed our past light cone." For instance, when astronomers measure the orbital speeds of planets orbiting other stars, the measured velocities are always of this sort. There is an excellent reason for this: we never have any information about what a distant object is doing (or indeed if the object even exists) at the present cosmic time. Any statement in which "now" is used to refer to the present cosmic time at the location of a distant object is not talking about anything observable, since it refers to events far outside our light cone.

In summary, if we wish to talk about the redshift of a distant galaxy as a Doppler shift, we need to be willing to talk about $v_{\rm rel}$, the velocity of the galaxy *then* relative to us *now*. This is precisely the sort of thing that the enlightened cosmologist described in Section I refuses to do because, in a curved spacetime, there is not a unique way to define the relative velocity of objects at widely separated spacetime events. Determining the velocity of one object relative to another involves comparing the two object's velocity four-vectors. [To be specific, the dot product of these vectors equals $\gamma_{\rm rel} \equiv (1 - v_{\rm rel}^2/c^2)^{-1/2}$.] In order to do that, we have to transport one of the vectors to the location of the other. In a curved spacetime, the result of such "parallel transport" depends on the path along which the vector is transported.

Suppose that we wish to defy the purist and talk about $v_{\rm rel}$. We have to parallel transport the galaxy's velocity four-vector to the observer's location. The only natural path to choose for this parallel transport is the path followed by the light itself (that is, the null geodesic joining the emission and observation events). If we follow this procedure to determine the relative velocity, we find that the velocity obeys the rule

$$\sqrt{\frac{c + v_{\rm rel}}{c - v_{\rm rel}}} = \frac{a(t_o)}{a(t_e)}. \tag{6}$$

As we noted in eq. (5), the ratio of scale factors is equal to $1 + z$, so this is precisely the special-relativistic Doppler formula. In other words, the relative speed $v_{\rm rel}$ as defined by parallel transport is related to the observed redshift exactly as it should be if the redshift is a Doppler shift.

Figure 3 illustrates this process. The solid line is the world line of a galaxy, and the dashed line is the path of light traveling from the galaxy to the observer. The scale factor is as shown in Fig. 2. The short solid lines indicate the galaxy's velocity four-vector as it is parallel transported along the light path to the observer. During the periods when the universe is expanding slowly the direction of this vector doesn't change significantly; however, it does change when the universe expands more rapidly.

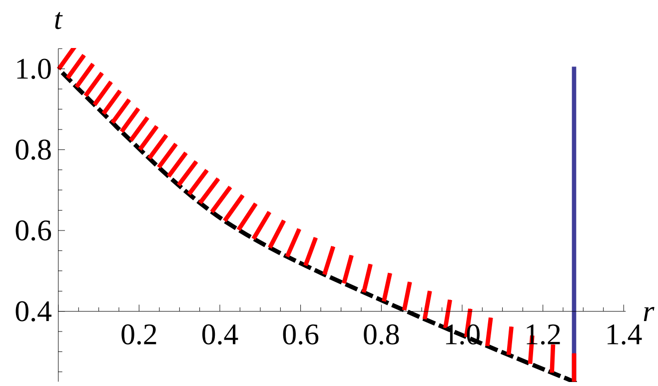

FIG. 3: Parallel transport of the velocity four-vector. The solid and dashed lines show the world line of a galaxy and the path of radiation from the galaxy to the observer in comoving coordinates. The short solid lines show the galaxy's four-velocity being parallel transported to the observer along the path of the light.

Equation (6) can be proved by a straightforward calculation using the rules for parallel transport, but the derivation becomes easier to see if we recast the statement in more physical terms. As we did at the beginning of this section, imagine many comoving observers stationed along the line from the observed galaxy to the observer. Each observer has a local reference frame in which special relativity can be taken to apply, and the observers are close enough together that each one lies in within the local frame of his neighbor. Observer number 1, who is located near the original galaxy, measures its speed v_1 relative to him and gives this information to observer 2. Observer 2 measures the speed u of observer 1 relative to him, adds this to the speed of the galaxy relative to observer 1 using the usual special-relativistic formula,

$$v_2 = \frac{v_1 + u}{1 + v_1 u/c^2}, \tag{7}$$

and interprets the result as $v_{\rm rel}$, the speed of the galaxy relative to him. He passes this information on to the next observer, who follows the same procedure, as does each subsequent observer. At each stage, the imputed velocity of the original galaxy relative to the observer will match the redshift of the galaxy in accordance with eq. (6). The procedure we have described operationally here is precisely equivalent to performing parallel transport of the galaxy's four-velocity and using that vector at each stage to compute $v_{\rm rel}$.

We can express this yet another way. Consider the "world tube" obtained by drawing a sphere of some small radius ϵ around each point on the path of the light ray from the galaxy to the observer. This region of spacetime can be considered as flat Minkowski spacetime, up to errors of order ϵ^2. (This is true for any small neighborhood around a geodesic.) Within this tube, the observed redshift can only be explained as a Doppler shift, since the spacetime is flat to arbitrary precision. In fact, this argument supplies the simplest proof we can think of that $v_{\rm rel}$ is related to the redshift by eq. (6).

We do not expect to have convinced the purist who refuses even to talk about $v_{\rm rel}$ to change his mind. He would dismiss $v_{\rm rel}$ as a mere "coordinate velocity," not an "actual velocity," and take no interest in it. If the purist has the courage of his convictions, he would furthermore say that it makes no sense to try to attach labels such as "Doppler" or "gravitational" to the observed redshift. This position is unassailable, and we have no wish to argue against it. We do claim, however, that *if* you wish to try to talk about $v_{\rm rel}$, then the definition proposed in this section is the most natural way to do so. Since this definition of $v_{\rm rel}$ results in the Doppler formula entirely explaining the redshift, we claim that you should either be a purist and refuse to try to label the cosmological redshift, or you should label it a Doppler shift.

It is interesting to contrast the position we advocate with another approach[13,27] that describes the cosmological redshift as a particular combination of Doppler and gravitational terms. In that approach, comoving coordinates are used to specify the velocity of the distant galaxy relative to the observer, and the observed redshift is decomposed into a Doppler term based on this velocity and a gravitational blueshift, which can be computed by considering the gravitational potential due to the matter in a sphere centered on the observer. The difference between the two approaches is in the way the velocity is defined. Since there is no unique specification of relative velocity for distant objects in curved spacetime, both approaches are correct. Our belief is that the approach based on defining $v_{\rm rel}$ through parallel transport is more natural than the one based on a particular choice of coordinates.

IV. KINEMATIC AND GRAVITATIONAL REDSHIFTS IN GENERAL SPACETIMES

Let us step back from the specific case of the expanding universe and consider the propagation of a photon in an arbitrary spacetime. A photon is emitted at some spacetime point $\mathcal{P}_e$ with a frequency ν_e and is absorbed at $\mathcal{P}_a$ with frequency ν_a. The two frequencies are measured by observers in local inertial reference frames at the two spacetime points. Assuming that there is a frequency shift ($\nu_e \neq \nu_a$), under what circumstances should we describe it as a Doppler shift or as a gravitational shift?

A natural way to answer this question is to populate the space between the photon's emission and observation with a dense family of observers. Let the first observer be at rest relative to the emitter of the photon at the emission event, and let the last be at rest relative to the absorber at the absorption event. Finally, let the velocities of observers vary smoothly along the photon's path. In this case, as in the previous section, we can explain the observed frequency shift by adding up the shifts as the photon passes from one observer to the next. We can construct two such families of observers: one in which each of the shifts is a Doppler shift, and one in which each is a gravitational shift. By reference to these families, we can interpret the observed shift as the accumulation of either many small Doppler shifts or many small gravitational shifts.

The Doppler and gravitational families are not the only possibilities, of course. We can construct other families with respect to whom both Doppler and gravitational shifts would be seen to contribute. As noted in the previous section, for example, one way of regarding the cosmological redshift[13,27] splits it into a specific combination of Doppler and gravitational terms. The Doppler and gravitational families provide two of many different ways of interpreting the observed shift. In any given situation, we can argue about which (if either) of the two families is natural to consider.

To construct the Doppler family of observers, we simply require all observers to be in free fall. In this case, within each local inertial frame, there are no gravitational effects, and hence the infinitesimal frequency shift from each observer to the next is a Doppler shift.

To construct the gravitational family of observers, we demand that each member be at rest relative to her neighbor at the moment the photon passes by, so that there are no Doppler shifts. Initially, it may seem impossible in general to satisfy this condition simultaneously with the condition that the first and last observers be at rest relative to the emitter and absorber, but in fact it is always possible to do so. One way to see this is to draw a small world tube around the photon path as in the previous section. Within this tube, spacetime is arbitrarily close to flat. We can construct "Rindler elevator coordinates," the special-relativistic generalization of a frame moving with uniform acceleration, within this tube, such that the velocities at the two ends match up correctly.[28] The members of the gravitational family are at rest in these coordinates. Since they are not in free fall, the members of the gravitational family all feel like they are in local gravitational fields. Since each has zero velocity relative to her neighbor when the photon goes by, each observer interprets the shift in the photon's frequency relative to her neighbor as a gravitational shift.

Since the two families exist for any photon path, we can always describe any frequency shift as either Doppler or gravitational. In some situations, one seems clearly more natural than the other. When we discuss the Pound-Rebka experiment, which measured the redshift of photons moving upward in Earth's gravitational field, we generally choose to regard observers fixed relative to the Earth (the gravitational family) as more natural than free-fall observers, and hence we interpret the measurement as a gravitational redshift. On the other hand, if you were falling past Pound and Rebka in a freely falling elevator as they performed the experiment, you might choose a Minkowskian inertial frame encompassing you and the entire experiment. Within this frame, you would, by the equivalence principle, interpret their results as a Doppler shift. In so doing, you would in effect be choosing to regard the Doppler family as the natural one, since this family is the one whose behavior is simplest to describe in your chosen frame. Furthermore, if you tried to beat a speeding ticket by claiming that the radar results were due to a gravitational redshift, you would in effect be considering a “gravitational family” of prodigiously accelerating observers, with one at rest relative to the radar gun and one at rest relative to the driver. Needless to say, the police officer who gives you a ticket regards this family as extremely unnatural.

In the cosmological context, we almost always work with freely falling comoving observers, a choice which corresponds precisely to the Doppler family. This is a natural choice, as it respects the symmetries of the spacetime and makes the mathematical description as simple as possible. On the other hand, the gravitational family is so unnatural in cosmology that as far as we know it has never been used for anything. Since the Doppler family is the by far the most natural family to work with, it is natural to interpret the cosmological redshift as a Doppler shift, and it is curious, to say the least, that this interpretation is generally frowned upon.

V. WHY THIS MATTERS

While there is no “fact of the matter” about the origin of the redshift – what one concludes is a function of the coordinate system or method of calculation – we believe it is instructive to analyze the differing interpretations of the cosmological redshift, partly to improve understanding of cosmology, but more importantly to improve understanding of general relativity. That analysis leads, in our view, to the conclusion that the most natural interpretation of the redshift is as having a kinematic origin.

One of the key ideas of general relativity is the importance of distinguishing between coordinate-independent and coordinate-dependent statements. Another is the idea that spacetime is always locally indistinguishable from Minkowski spacetime. Cosmology teachers, books (especially at the introductory level), and students often fall into the fallacy of reifying the rubber sheet – that is, of treating the expanding-rubber-sheet model of space as if it were a real substance. This error leads people away from both of these key ideas and causes mistaken intuitions such as that the Milky Way Galaxy must constantly “resist the temptation” to expand with the expanding Universe or that the "tethered galaxy" described in Section I moves away after the tether is cut.

The common belief that the cosmological redshift can "only" be explained in terms of the stretching of space is based on conflating the properties of a specific coordinate system with properties of space itself. This is precisely the opposite of the correct frame of mind in which to understand relativity.

Acknowledgments

We thank Sean Carroll, John Peacock, Jim Peebles, Dave Spiegel, and Ned Wright for helpful comments. This work was supported in part by NSF Grant AST-0507395 (EFB) and a research fellowship from the Alexander von Humboldt Foundation of Germany (DWH).

---

[*] Electronic address: `ebunn@richmond.edu`

[1] J. C. Baez and E. F. Bunn, "The meaning of Einstein's equation," Am. J. Phys. **73**, 644–652 (2005).

[2] S. Carroll, *Spacetime and Geometry: an Introduction to General Relativity* (Benjamin Cummings, San Francisco, 2003).

[3] B. F. Schutz, *A First Course in General Relativity* (Cambridge University Press, Cambridge, UK, 1985).

[4] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W.H. Freeman, New York, 1973).

[5] W. J. Kaufmann and R. Freedman, *Universe*, 5th edition (W.H. Freeman, New York, 1999).

[6] A. Fraknoi, D. Morrison, and S. Wolff, *Voyages Through the Universe*, 3rd edition (Thomson Brooks/Cole, Belmont, CA, 2004).

[7] M. A. Seeds, *Foundations of Astronomy*, 10th edition (Thomson Brooks/Cole, Belmont, CA, 2007).

[8] J. Peacock, *Cosmological Physics* (Cambridge University Press, Cambridge, UK, 1998).

[9] E. V. Linder, *First Principles of Cosmology* (Addison-Wesley, Reading, MA, 1997).

[10] E. Harrison, *Cosmology*, 2nd edition (Cambridge University Press, Cambridge, UK, 2000).

[11] M. Chodorowski, "Is space really expanding? A counterexample," (2006). Eprint arXiv:astro-ph/0601171.

[12] M. J. Chodorowski, "A direct consequence of the expansion of space?" Mon. Not. Royal Astron. Soc. **378**, 239–244 (2007).

[13] J. A. Peacock, "A diatribe on expanding space," (2008). Eprint arXiv:0809.4573.

[14] A. B. Whiting, "The expansion of space: free-particle motion and the cosmological redshift," The Observatory **124**, 174–189 (2004).

[15] One might argue against this claim by noting that the general relativistic Maxwell equations contain covariant derivatives, which "know about" the expansion. However, covariant derivatives are locally

indistinguishable from ordinary derivatives (because a local Minkowski frame can always be chosen).

[16] E. R. Harrison, "Mining energy in an expanding universe," Astrophys. J. **446**, 63–66 (1995).

[17] T. M. Davis, C. H. Lineweaver, and J. K. Webb, "Solutions to the tethered galaxy problem in an expanding universe and the observation of receding blueshifted objects," (2001). Eprint arXiv:astro-ph/0104349.

[18] M. Sereno and P. Jetzer, "Evolution of gravitational orbits in the expanding universe," Phys. Rev. D **75**(6), 064031 (2007).

[19] F. I. Cooperstock, V. Faraoni, and D. N. Vollick, "The influence of the cosmological expansion on local systems," Astrophys. J. **503**, 61–66 (1998).

[20] L. A. Barnes, M. J. Francis, J. B. James, and G. F. Lewis, "Joining the Hubble flow: implications for expanding space," Mon. Not. Royal Astron. Soc. **373**, 382–390 (2006).

[21] M. J. Francis, L. A. Barnes, J. B. James, and G. F. Lewis, "Expanding space: the root of all evil?" Pub. Astron. Soc. of Australia **24**, 95–102 (2007).

[22] If you try this, let us know how it works out for you.

[23] E. A. Milne, "A Newtonian expanding universe," Quarterly Journal of Mathematics **5**, 64–72 (1934).

[24] This is an occasional source of confusion, because a low-density universe is an open universe, in which *space*, as opposed to spacetime, does have negative curvature. In the present context, what matters is the spacetime curvature, i.e., the deviation of the overall spacetime metric from that of special relativity, not the spatial curvature, which describes the geometry of a "slice" through the spacetime at constant time in a particular coordinate system.

[25] T. M. Davis and C. H. Lineweaver, "Superluminal recession velocities," in "Cosmology and Particle Physics," edited by R. Durrer, J. Garcia-Bellido, and M. Shaposhnikov, volume 555 of *American Institute of Physics Conference Series*, pp. 348–351 (2001).

[26] T. M. Davis and C. H. Lineweaver, "Expanding confusion: common misconceptions of cosmological horizons and the superluminal expansion of the Universe," Pub. Astron. Soc. of Australia **21**, 97–109 (2004).

[27] H. Bondi, "Spherically symmetrical models in general relativity," Mon. Not. Royal Astron. Soc. **107**, 410–425 (1947).

[28] W. Rindler, *Essential Relativity*, 2nd edition (Springer-Verlag, New York, 1977).